Post-Graduate Degree Programme (CBCS)
in
GEOGRAPHY

Semester-I

Self Learning Material

Paper Code: GEO/CC/T-101
Paper: Geotectonics, Geomorphology and Hydrology

DIRECTORATE OF OPEN AND DISTANCE LEARNING (DODL)

UNIVERSITY OF KALYANI

Kalyani, Nadia West
Bengal, India

Course Materials Compiled by:

Dr. S. Choudhary and Dr. A. Rudra

Directorate of Open and Distance Learning (DODL)

University of Kalyani, Kalyani-741235, Nadia, West Bengal

December 2021

Published by the Directorate of Open and Distance Learning (DODL)
University of Kalyani, Kalyani-741235, Nadia, West Bengal

Disclaimer: This self-learning material is based on different books, journals and web-sources.

Syllabus

SEMESTER-I						
Paper Code	**Paper**	**Theory/ Practical**	**Internal Assessment/ Evaluation**	**Examination/ Report/ Viva-Voce**	**Credit**	**Marks**
GEO/CC/T-101	Geotectonics, Geomorphology and Hydrology	Theory	10	40 (Semester-end Examination)	4	50

Unit-1: Plate tectonics as a unified theory of global tectonics

Unit-2: Tectonic geomorphology: Influence of tectonics in landscape evolution

Unit-3: Concepts in Geomorphology: spatial scale, temporal scale, systems, feedback, equilibrium and threshold

Unit-4: Catchment process and fluvial processes; Factors regulating entrainment, transportation and deposition

Unit-5: Adjustment of channel forms and patterns to morphodynamic variables

Unit-6: Coastal morphodynamic variables and their influence in evolution of landforms

Unit-7: Periglacial processes and landforms

Unit-8: Elements of slope and different approaches to study slope development

Unit-9: Concept of basin hydrology and run off cycle; Unit hydrograph, rating curve and their applications

Unit-10: Storm water and flood management: storm water management, design of drainage system, flood routing through channels and reservoir, flood control and reservoir operation

Unit-11: Drought management**:** drought assessment and classification, drought analysis techniques, drought mitigation planning

Unit-12: Methods of water conservation: rainwater harvesting and watershed management

Mode of Internal Evaluation: Class test

Contents

1.1 INTRODUCTION

Geomorphology is the area of study leading to an understanding of and appreciation for landforms and landscapes, including those on continents and islands, those beneath oceans, lakes, rivers, glaciers and other water bodies, as well as those on the terrestrial planets and moons of our Solar System. Contemporary geomorphologic investigations are most commonly conducted within a scientific framework (see Rhoads and Thorn 1996) although academic, applied or engineering interests may motivate them. A broad range of alternative research methodologies have been employed by geomorphologists, and past attempts to impose a systematic structure on the discipline have yielded stifling tendencies and overt resistance. Geomorphologists frequently profess to innate aesthetic appreciation for the complex diversity of Earth-surface forms, and, in this regard, a fitting definition of geomorphology is simply 'the science of scenery' (Fairbridge 1968). Past and present concerns have focused on the description and classification of landforms (including their geometric shape, topologic attributes and internal structure), on the dynamical processes characterizing their evolution and existence, and on their relationship to and association with other forms and processes (geomorphic, hydro-climatic, tectonic, biotic, anthropogenic, extraterrestrial, or otherwise). Geomorphology is an empirical science that attempts to formulate answers to the following fundamental questions. What makes one landform distinct from another? How are different landforms associated? How did a particular landform or complex landscape evolve? How might it evolve in the future? What are the ramifications for humans and human society?

Modern geomorphology is currently subdivided and practised along the lines of specialized domains. *Fluvial geomorphology*, for example, is concerned with flowing water (primarily in the form of rivers, streams and channels) and the work it accomplishes during its journey through the terrestrial phase of the hydrologic cycle. A very broad spectrum of interests are subsumed within fluvial geomorphology, ranging from the influence of turbulence on the entrainment, transport and deposition of sediment particles at the finest scale, to the mechanics of MEANDERING, POINT BAR formation and FLOODPLAIN development at middle scales, to the nature and character of DRAINAGE BASIN evolution at the coarsest scales. Within the other substantive areas of geomorphology are: *hillslope geomorphologists*, who boast expertise on the geotechnical properties of soil and rock, the mechanics of LANDSLIDEs, and the movement of water within the ground; *tectonic geomorphologists*, who study neotectonic (see NEOTECTONICS) stress fields, continental-scale sedimentary basins and active/passive margin landscapes; *glacial and periglacial geomorphologists*, who are interested in alpine and continental glaciers, PERMAFROST and other cold-climate forms or processes that involve ice, snow and frost; *karst geomorphologists*, who deal with soluble rocks (e.g. limestone) and chemical processes of DISSOLUTION that lead to landforms such as gorges, caverns and underground streams; *coastal geomorphologists*, who study nearshore, lacustrine and marine systems where oscillatory, rather than unidirectional, flow processes dominate; and *aeolian geomorphologists*, who study the transport of sand and dust by wind, mostly in desert or semi-arid environments, but also along beaches, over agricultural fields and on the moon and Mars. Other subspecialities include: soils geomorphology, biogeomorphology (zoogeomorphology), climatic geomorphology, tropical geomorphology, desert geomorphology, mountain geomorphology, extraterrestrial (planetary) geomorphology, remote-sensing geomorphology, experimental geomorphology, environmental geomorphology, forest geomorphology, applied geomorphology, engineering geomorphology and anthropogeomorphology (BAUER, 2004).

1.2 LEARNING OBJECTIVES

The present section aims to introduce the following topics –

- Plate tectonics as a unified theory of global tectonics
- Tectonic geomorphology: Influence of tectonics in landscape evolution
- Concepts in Geomorphology: spatial scale, temporal scale, systems, feedback, equilibrium and threshold
- Catchment process and fluvial processes; Factors regulating entrainment, transportation and deposition
- Adjustment of channel forms and patterns to morphodynamic variables
- Coastal morphodynamic variables and their influence in evolution of landforms
- Periglacial processes and landforms
- Elements of slope and different approaches to study slope development
- Concept of basin hydrology and run off cycle; Unit hydrograph, rating curve and their applications
- Storm water and flood management: storm water management, design of drainage system, flood routing through channels and reservoir, flood control and reservoir operation
- Drought management**:** drought assessment and classification, drought analysis techniques, drought mitigation planning
- Methods of water conservation: rainwater harvesting and watershed management

1.3 ASSESSMENT OF PRIOR KNOWLEDGE

Discussion about nature, scope and relevance of geomorphology is necessary. Discussion about recent trends in geomorphological research is necessary.

1.4 LEARNING ACTIVITIES

Preparation of short notes and essays on different topics discussed in class

1.5 FEEDBACK OF LEARNING ACTIVITIES

Debate and discussion on various topics discussed in the class may be conducted. Class seminar on various topics may be arranged.

UNIT-1: PLATE TECTONICS AS A UNIFIED THEORY OF GLOBAL TECTONICS

Over the past century, the theoretical framework of *continental drift* has been refined into a well-established theory called *plate tectonics,* which has been tested by collecting a great deal of evidence from the lithosphere. The theory of plate tectonics has revolutionized the Earth sciences and our understanding of Earth's history. Long ago, some scientists believed that Earth's landscapes were created by great cataclysms. They might have believed, for example, that the Grand Canyon split open one violent day and has remained that way ever since, or that the Rocky Mountains appeared overnight. This theory, called **catastrophism,** has been rejected. For almost two centuries, physical geographers, geologists, and other Earth scientists have accept instead the theory of **uniformitarianism,** which is the idea that internal and external Earth processes operate today in the same manner as they have for millions of years. Uniformitarianism, however, does not mean that processes have always operated at the same rate or with equal strength everywhere on Earth. In fact, our planet's surface features are the result of variations in the intensity of internal and external processes, influenced by their geographical location. These processes have varied in intensity and location throughout Earth's history. Furthermore, regular or episodic changes in the Earth system that may seem relatively small to us can dramatically alter a landscape after progressing, even on an irregular basis, for millions of years.

CONTINENTAL DRIFT

Most of us have probably noted on a world map that the Atlantic coasts of South America and Africa look as if they could fit together. In fact, if a globe were made into a spherical jigsaw puzzle, several widely separated landmasses could fit alongside each other without large gaps or overlaps. Is there a scientific explanation for this phenomenon? In the early 1900s, Alfred Wegener, a German climatologist, proposed the theory of **continental drift,** the idea that continents and other landmasses have shifted their positions during Earth history. Wegener's evidence for continental drift included the close fit of continental coastlines on opposite sides of oceans and the trends of mountain ranges on land areas that also match across oceans. He cited comparable geographical patterns of fossils and rock types found on different continents that he felt could not result from chance and did not reflect current climatic conditions. To explain the spatial distributions of these features, he reasoned that the continents must have been previously joined. Wegener also noted evidence of great climate change, such as ancient evidence of glaciation where the Sahara Desert is today and tropical fossils found in Antarctica, that could be explained best by large landmasses moving from one climate zone to another. Wegener hypothesized that all the continents had once been part of a single supercontinent, which he called **Pangaea,** that later divided into two large landmasses, one in the Southern Hemisphere (Gondwana), and one in the Northern Hemisphere (Laurasia). Later, these two supercontinents also broke apart into sections (the present continents) and drifted to their current positions. Laurasia in the Northern Hemisphere consisted of North America, Europe, and Asia. Gondwana in the Southern Hemisphere was made up of South America, Africa, Australia, Antarctica, and India. Continued continental movement created the geographical configuration of the landmasses that exist on Earth today.

SUPPORTING EVIDENCE FOR CONTINENTAL DRIFT

About a half century later, in the late 1950s and 1960s, Earth scientists began giving serious consideration to Wegener's notion of moving continents. New information appeared from research in oceanography, geophysics, and other Earth sciences, aided by sonar, radioactive dating of rocks, and improvements in equipment for measuring Earth's magnetism. These scientific efforts discovered much new evidence that indicated the movement of portions of the lithosphere, including the continents. As one example, scientists were originally unable to explain the varied orientations of magnetic fields found in basaltic rocks that had cooled millions of years ago. They knew that iron-bearing minerals in rocks display the magnetic field of Earth as it existed when the rocks solidified, which is a phenomenon known as **paleomagnetism.** Scientists at that time also knew that the exact position of the magnetic poles wandered through time, but they could not account for the confusing range of magnetic field orientations indicated by the basaltic rocks they studied. Magnetic field orientations of rocks of the same age did not point toward a single spot on Earth, and the indicated positions for the magnetic north pole ranged widely, including some that pointed toward the present south magnetic pole. The observed variations were more than could be accounted for by the known magnetic polar wandering. Scientists eventually used the paleomagnetic data to model where the sampled rocks would have to have been relative to a common magnetic north pole. Successful alignment was only possible if the continents had been in different positions than they are today. Using rocks of different age, they reconstructed locations of the continents during past periods in geologic history. Paleomagnetic data revealed that the continents were grouped together about 200 million years ago, just as Wegener's hypothesized two supercontinents began to split apart to form the beginnings of the modern Atlantic Ocean. Paleomagnetic data also revealed that the polarity of Earth's magnetic field had reversed many times in the past. A record of these polarity reversals was imprinted within the iron-rich basaltic rocks of the seafloor. Supporting evidence for crustal movement came from a variety of other sources in the mid-20th century. The widely separated patterns of similar fossil reptiles and plants found in Australia, India, South Africa, South America, and Antarctica, previously noted by Wegener, were mapped in detail. The fossils represented organisms that in each instance were so similar and specialized that they could not have developed without their now-distant locations being either connected or at least much closer together than they are today. When the positions of the continents were reassembled on a paleomap derived from paleomagnetic data and representing the time when the organisms were living, the fossil locations came together spatially. Other types of ancient environmental evidence, such as left by glaciations, could also be fit together in logical geographical patterns on reconstructed paleomaps of the continents or the world. How well Earth's landmasses match up when they are brought together on a paleomap was found to be even better when using the true continental edges—the continental slopes— which lie a few hundred meters below sea level. In this case, as also had been noted by Wegener, mountain ranges on opposite sides of oceans line up and rock ages and types match where the continents join. Knowledge of the geographical distribution of Earth's environments relative to latitude and climate zones provided additional insight. Evidence that ancient glaciation occurred simultaneously in India and South Africa while tropical forest climates (represented by coal deposits) existed in the northeastern United States and in Great Britain could only be explained by the latitudinal movement of landmasses, and their locations came together well on *paleogeographic* reconstructions.

Plate tectonics, the modern theory to explain the movement of continents, suggests that the rigid and brittle outer shell of Earth, that is, the lithosphere (crust and uppermost mantle), is broken into several separate segments called **lithospheric plates** that rest on, and are carried along with, the flowing plastic asthenosphere. Tectonics involves large-scale forces originating within Earth that cause parts of the lithosphere to move around. In plate tectonics, the lithospheric plates move as distinct and discrete units. In some places they pull away from each other (diverge), in other places they push together (converge), and elsewhere they slide alongside each other (move laterally). Seven major plates have proportions as large as or larger than continents or ocean basins. Five other plates are of minor size, although they have maintained their own identity and direction of movement for some time. Several additional plates are even smaller and exist in active zones at the boundaries between major plates. All major plates consist of both continental and oceanic crust although the largest, the Pacific plate, is primarily oceanic. To understand how plate tectonics operates and why plates move, we must consider the scientific evidence that was gathered to test this theory. We should also evaluate how well this theory holds up under rigorous examination. The supporting evidence, however, is overwhelming.

SEAFLOOR SPREADING AND CONVECTION CURRENTS

In the 1960s, several keys to plate tectonics theory were found while studying and mapping the ocean floors. First, detailed undersea mapping was conducted on a system of midoceanic ridges (also called oceanic ridges or rises) that revealed configurations remarkably similar to the continental coastlines. Second, it was discovered in the Atlantic and Pacific Oceans that basaltic seafloor displayed parallel bands of matching patterns of magnetic properties in rocks of the same age but on opposite sides of midoceanic ridges. Third, scientists made the surprising discovery that although some continental rocks are 3.6 billion years old, rocks on the ocean floor are all geologically young, having been in existence less than 250 million years. Fourth, the oldest rocks of the seafloor lie beneath the deepest ocean waters or close to the continents, and rocks become progressively younger toward the midoceanic ridges where the youngest basaltic rocks are found. Finally, temperatures of rocks on the ocean floor vary significantly, being hottest near the ridges and becoming progressively cooler farther away. Only one logical explanation emerged to fit all of this evidence. It became apparent that new oceanic crust is being formed at the midoceanic ridges while older oceanic crust is being destroyed along other margins of ocean basins. The emergence of new oceanic crust is associated with the movement of great sections or plates of the lithosphere away from the midoceanic ridges. This phenomenon, which represents a major advance in our understanding of how continents move, is called **seafloor spreading**. The rigid lithospheric plates diverge along the oceanic ridges and separate at an average rate of 2 to 5 centimeters (1–2 in.) per year as they are carried along with the flowing plastic asthenosphere in the mantle. The young age of oceanic crust results from the creation of new basaltic rock at undersea ridges and the movement of the seafloor with lithospheric plates toward ocean basin margins where the older rock is remelted and destroyed. As molten basalt cooled and crystallized in the seafloor, the iron minerals that they contain became magnetized in a manner that replicated the orientation of Earth’s magnetic field at that time. The iron-rich basalts of the seafloor have preserved a historical record of Earth’s magnetic field, including **polarity reversals** (times when the north pole became south, and vice versa). Plate tectonics includes a plausible explanation of the mechanism for continental movement, which had eluded Wegener. The mechanism is **convection.** Hot mantle material travels upward toward Earth’s surface and cooler material moves downward as part of huge subcrustal convection cells. Mantle material rises to the asthenosphere where it spreads laterally and flows in opposite directions, dragging the lithospheric plates with it. Pulling apart the brittle lithosphere breaks open a midoceanic ridge. Molten basalt wells up into the fractures, cooling and sealing them to form new seafloor. In this process, the ocean becomes wider by the width of the now-sealed fracture. The convective motion continues as solidified crustal material moves away from the ridges. In a time frame of up to 250 million years, older oceanic crust is consumed in the deep trenches near plate boundaries where sections of the lithosphere meet and are recycled into Earth’s interior.

TECTONIC PLATE MOVEMENT

The shifting of tectonic plates relative to one another provides an explanation for many of Earth’s surface features. Plate tectonics theory enables physical geographers to better understand not only our planet’s ancient geography but also the modern global distributions and spatial relationships among such diverse, but often related, phenomena as earthquakes, volcanic activity, zones of crustal movement, and major landform features. Let’s briefly examine the three ways in which lithospheric plates relate to one another along their boundaries as a result of tectonic movement: by pulling apart, pushing together, or sliding alongside each other.

Plate Divergence The pulling apart of plates, tectonic **plate divergence,** is directly related to seafloor spreading. Tectonic forces that act to pull objects apart cause the crust to thin and weaken. Shallow earthquakes are often associated with this crustal stretching, and asthenospheric magma wells up between crustal fractures. This creates new crustal ridges and new ocean floor as the plates move away from each other. The formation of new crust in these spreading centers gives the label *constructive plate margins* to these zones. Occasional "oceanic" volcanoes, like those of Iceland, the Azores, and Tristan da Cunha, mark such boundaries. Most plate divergence occurs along oceanic ridges, but this process can also break apart continental crust, eventually reducing the size of the landmasses involved. The Atlantic Ocean floor formed as the continent that included South America and Africa broke up and moved apart 2 to 4 centimeters (1–2 in.) per year over millions of years. The Atlantic Ocean continues to grow today at about the same rate. The best modern example of divergence on a continent is the rift valley system of East Africa, stretching from the Red Sea south to Lake Malawi. Crustal blocks that have moved downward with respect to the land on either side, with lakes occupying many of the depressions, characterize the entire system, including the Sinai Peninsula and the Dead Sea. Measurable widening of the Red Sea suggests that it may be the beginning of a future ocean that is forming between Africa and the Arabian Peninsula, similar to the young Atlantic between Africa and South America about 200 million years ago.

PLATE CONVERGENCE

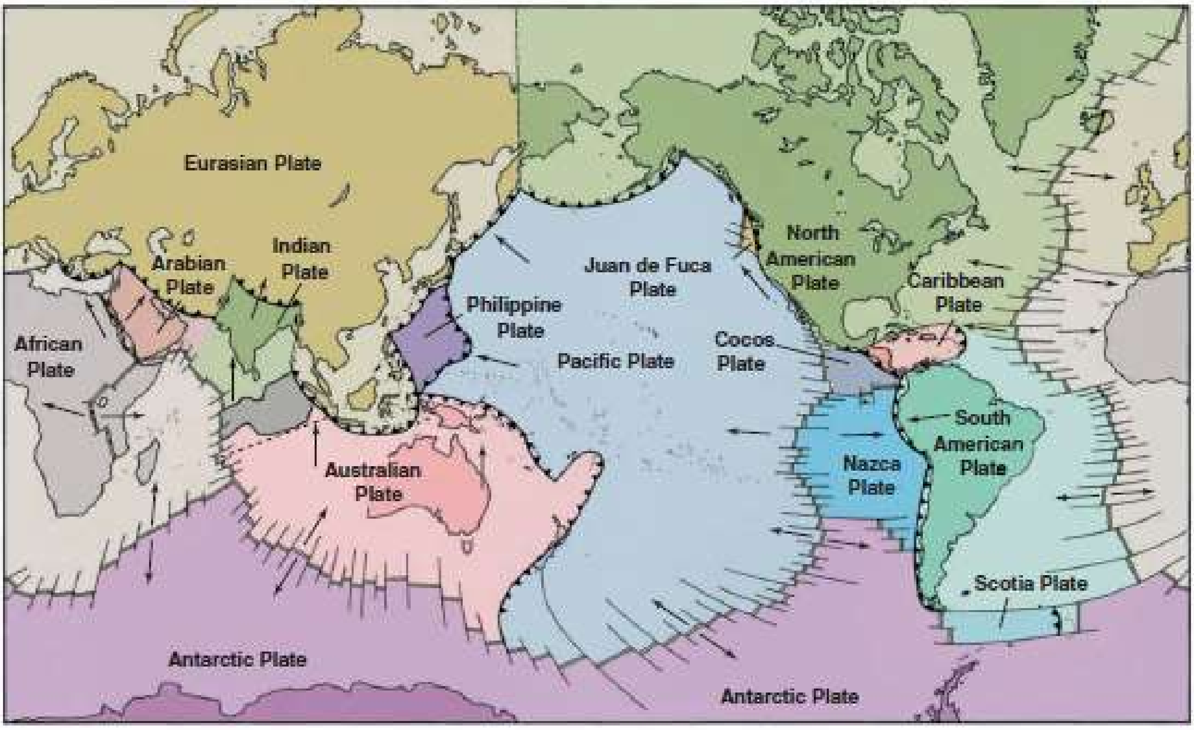
A vide variety of crustal activity occurs at areas of tectonic **plate convergence.** Despite the relatively slow rates of plate movement (in terms of human perception), the incredible energy involved in convergence causes the crust to crumple as one plate overrides another. The denser plate is forced deep below the surface in a process called **subduction.** Subduction is most common where dense oceanic crust collides with and descends beneath less dense continental crust. This is the situation along South America's Pacific coast, where the Nazca plate subducts beneath the South American plate, and in Japan, where the Pacific plate dips under the Eurasian plate. As oceanic crust, and the lithospheric plate of which it forms a part, is subducted, it descends into the asthenosphere to be melted and recycled into Earth's interior. Deep ocean trenches, such as the Peru–Chile trench and the Japanese trench, occur where the crust is dragged downward into the mantle. Frequently, hundreds of meters of sediments that are deposited on the seafloor or along continental margins are carried down into these trenches. At such convergent boundaries, rocks can be squeezed and contorted between colliding plates, becoming uplifted and greatly deformed or metamorphosed. These processes have produced many great mountain ranges, such as the Andes, at convergent plate margins. A subducting plate is heated as it plunges downward into the mantle. Its rocks are melted, and the resulting magma migrates upward into the overriding plate. Where molten rock reaches the surface, it forms a series of volcanic peaks, as in the Cascade Range of the northwestern United States. Where two oceanic plates meet, the older, denser one will subduct below the younger, less dense oceanic plate, and volcanoes may develop, creating major **island arcs** on the overriding plate between the continents and the ocean trenches. The Aleutians, the Kuriles, and the Marianas are all examples of island arcs near oceanic trenches that border the Pacific plate. As the subducting plate grinds downward, enormous friction is produced, which explains the occurrence of major earthquakes in these regions. Subduction zones are sometimes referred to as Benioff zones, after the seismologist Hugo Benioff, who first plotted the position of earthquakes extending downward at a steep angle on the leading edge of a subducting plate. **Continental collision** causes two continents or major landmasses to fuse or join together, creating a new larger landmass. This process, which closes an ocean basin that once separated the colliding landmasses, has been called *continental suturing.* Where two continental masses collide, the result is massive folding and crustal block movement rather than volcanic activity. This crustal thickening generally produces major mountain ranges at sites of continental collision. The Himalayas, the Tibetan Plateau, and other high Eurasian ranges formed in this way as the plate containing the

Indian subcontinent collided with Eurasia some 40 million years ago. India is still pushing into Asia today to produce the highest mountains in the world. In a similar fashion, the Alps were formed as the African plate was thrust against the Eurasian plate. Zones where plates are converging mark locations of major, and some of the tectonically more active, landforms on our planet: huge mountain ranges, chains of volcanoes, and deep ocean trenches. The distinctive spatial arrangement of these features worldwide can best be understood within the framework of plate tectonics.

Transform Movement Lateral sliding along plate boundaries, called **transform movement,** occurs where plates neither pull apart nor converge but instead slide past each other as they move in opposite directions. Such a boundary exists along the San Andreas Fault zone in California. Mexico's Baja peninsula and Southern California are west of the fault on the Pacific plate. San Francisco and other parts of California east of the fault zone are on the North American plate. In the fault zone, the Pacific plate is moving laterally northwestward in relation to the North American plate at a rate of about 8 centimeters (3 in.) a year (80 km or about 50 mi per million years). If movement continues at this rate, Los Angeles will lie alongside San Francisco (450 mi northwest) in about 10 million years and eventually pass that city on its way to finally colliding with the Aleutian Islands at a subduction zone. Another type of lateral plate movement occurs on ocean floors in areas of plate divergence. As plates pull apart, they usually do so along a series of fracture zones that tend to form at right angles to the major zone of plate contact. These crosshatched plate boundaries along which lateral movement takes place are called *transform faults.* Transform faults, or fracture zones, are common along midoceanic ridges, but examples can also be seen elsewhere, as on the seafloor offshore from the Pacific Northwest coast between the Pacific and Juan de Fuca plates. Transform faults are caused as adjacent plates travel at variable rates, causing lateral movement of one plate relative to the other. The most rapid plate motion is on the East Pacific rise where the rate of movement is more than 17 centimeters (5 in.) per year.

HOT SPOTS IN THE MANTLE

The Hawaiian Islands, like many major landform features, owe their existence to processes associated with plate tectonics. As the Pacific plate moves toward the northwest near these islands, it passes over a mass of molten rock in the mantle that does not move with the lithospheric plate. Called **hot spots,** these almost stationary molten masses occur in a few other places in both continental and oceanic locations. Melting of the upper mantle and oceanic crust causes undersea eruptions and the outpouring of basaltic lava on the seafloor, eventually constructing a volcanic island. This process is responsible for building the Hawaiian Islands, as well as the chain of islands and undersea volcanoes that extend for thousands of miles northwest of Hawaii. Today the hot spot causes active volcanic eruptions on the island of Hawaii. The other islands in the Hawaiian chain came from a similar origin, having formed over the hot spot as well, but these volcanoes have now drifted along with the Pacific plate away from their magmatic source. Evidence of the plate motion is indicated by the fact that the youngest islands of the Hawaiian chain, Hawaii and Maui, are to the southeast, and the older islands, such as Kauai and Oahu, are located to the northwest. A newly forming undersea volcano, named Loihi, is now developing southeast of the island of Hawaii and will someday be the next member of the Hawaiian chain.



This map shows Earth's major tectonic plates and their general directions of movement. Most tectonic and volcanic activity occurs along plate boundaries where the large segments separate, collide, or slide past each other. Barbs indicate boundaries where one plate is overriding another, with the barbs on the side of the overriding plate (Gabler, Petersen, Trapasso, & Sack).

UNIT-2: TECTONIC GEOMORPHOLOGY: INFLUENCE OF TECTONICS IN LANDSCAPE EVOLUTION

THE FLUVIAL SYSTEM AND TECTONIC GEOMORPHOLOGY

This section deals with tectonic effects on rivers and their deposits; these are landforms that are very susceptible to change (Schumm, 1977). Tectonic effects can be easily recognized in consolidated rocks where the drainage system is a function of the differential resistance the rocks. The simplest examples are found in various types of drainage networks: rectangular and trellis networks, for example (Schumm *et al*., 2000). Geomorphologists use various parameters to describe river shapes and processes and the balance between them in a fluvial system tells us that rivers are very sensitive to any type of change (Keller and Pinter, 1996). Quaternary climate changes and their effects on landforms are very obvious (Bull, 1991; Gutierrez, 2005). Anthropic activity has considerably affected fluvial aggradation and degradation processes (Goudie, 1981b; Wilkinson, 2005). In addition, as knowledge of tectonic geomorphology increases, geomorphologists can analyze the role of tectonic activity in fluvial system processes and landforms (Adams, 1980; Ouchi, 1985). Drainage systems adapt to changes of the surface slope; therefore, they hold a valuable record of information on the evolution of faults and folds (Ollier, 1981). Earthquake induced changes to fluvial systems that result directly from deformation that occurs during an earthquake can threaten people who live along the riverbanks. Slow deformation of the valley bottom can affect channel stability and pose risks to man-made structures (Schumm, 1986). Changes that occur gradually over thousands of years are more common than simultaneous deformation. These changes highlight areas of active deformation (Schumm, 1986). Earthquakes and fault movements may cause large rivers that flow across floodplains to undergo **avulsion**, the abandonment of one channel for another. A classic example is the Indus River of Pakistan where in 1819 a large earthquake caused an uplift 6 meters high, 16 km wide, and 80 km long. The earthquake totally blocked the eastern margin of the Indus River and flow was not re-established in the channel until 1828 (Holmes, 1968). A river may reach an **equilibrium stage** in which active factors cause equilibrium to move in a direction that tends to absorb the effect of change and to re-establish equilibrium (MacKin, 1948; Leopold and Maddock, 1953). This implies that the effects of deformation are propagated upstream and downstream. Numerous rivers have adjusted to tectonic activity (Melton, 1959). A geomorphologist must know the magnitude of deformation to determine whether it is rapid enough to produce recent adjustment of the channel (Schumm *et al*., 2000). Another type of deformation is non seismic, which can be described as progressive uplift and subsidence that operates over a specific period of time (United States Department of Commerce, 1972; in Schumm *et al*., 2000).

UPLIFT, SUBSIDENCE, AND EARTHQUAKES IN FLUVIAL SYSTEMS

Numerous cases have been studied, mainly in the United States, in which the fluvial system provides independent evidence to show that geodetic measurements are valid and deformation is real (Adams, 1980). The reconstruction of longitudinal profiles of rivers can be used to show repeated uplift over a long time period. These reconstructions involve erosional and depositional terraces. The primary geomorphic indications of tectonic activity may be landforms created by long-term continuous deformation, which modifies the drainage network and deforms the set of terraces. These characteristics contrast with more subtle river changes (Schumm *et al*., 2000). Tilting can produce changes in an entire drainage network. In addition, continuous tilting can cause lateral erosion and the development of an asymmetric valley. The formation of large lakes in sedimentary basins may be partially due to tectonic activity. Lakes Titicaca and Uyuni, at elevations of about 4000 m in the Altiplano of Bolivia, are associated with plate convergence and uplift (Schumm *et al*., 2000).

Riverbeds along the Gulf Coast of the United States show that big rivers adjust to deformation more rapidly than small rivers, including uplift that affects flooding. Floods are much more frequent on the upstream side of uplifts than on their downstream side. The New Madrid earthquakes affected the Mississippi River in significant way and the effects persist 200 years later. Profiles of the maximum channel level (*bankfull stage*) in 1880 and 1915 show a transverse uplift (Lake County Uplift) that is shown by various convexities at the time of measurement. This convexity can be seen over about 40 miles (Russ, 1982). Profiles constructed downstream in 1880, 1976, and 1988 show that the 1880 profile is the lowest of all of them, which demonstrates that the river has eroded. The New Madrid convexity observed in the 1880 profile is not visible in the 1976 and 1988 profiles. The steepest part of the 1976 and 1988 profiles is located downstream of New Madrid. It shows a drop of 1.2 m over 3.2 km. This may reflect the exhumation of a fault scarp in the channel of the Mississippi River. The drop is located 8 km upstream of a fault where a waterfall formed after the earthquake of February 7, 1812 (Russ, 1982).

DEFORMATION IN COLLUVIAL AND FLUVIAL DEPOSITS

In tectonically active zones, slope profiles can cut faults, producing a step. This gradient change causes colluvium to fill the step, a process that continues until the scarp disappears. The **wedge conceptual model** is used in the study of colluvial deposits to interpret the history of a fault (Wallace, 1977). Japanese researchers (Okada *et al*., 1989) have given the name "D structure" to a fault-and-colluvium unit because it resembles the letter D. After the faulting, detritus fills the step, creating a wedge shape. The photograph shows ceramic fragments at the top of the wedge deposits. Because the ceramics postdate the middle Bronze and predate the Iberian epoch, they must have been produced between 1200 and 500 BC. We can conclude that significant tectonic activity occurred after the period of 1200 to 500 BC (Burillo *et al*., 1985). If tectonic activity is followed by a prolonged period of stability, the fault scarp disappears, the slope reaches equilibrium, and soil may develop on the scarp (Gutierrez *et al*., 1983, Gutierrez and Pena, 1994a). The soil may in turn be disturbed by a new stage of tectonic activity. A new colluvial wedge may form. Multiple faulting events produce a series of colluvial wedges that stack up on the down-dropped block of the fault. These wedges, which may be separated by a soil layer, represent stages of sedimentation that occur after fault movement (McCalpin, 1996b). This model has been applied to numerous faults in the western United States to identify up to four paleoseismic events (Forman *et al*., 1991). A **facies model** has been proposed for sedimentation at the base of a normal fault scarp (Nelson, 1992). This model establishes **detrital facies** made up of clasts and soil fragments, mostly from weathering processes on the slope above the fault plane. As the scarp fills, *slopewash* and *creep* occur on the slope and colluvium that is deposited is finer grained, has better stratification and sorting, and is richer in organic material. These deposits are called **wash facies**. In areas of normal faults, angular unconformities are present and associated with tilting due to drag on the fault plane. In tilted areas, the dip of the strata may increase with increasing age and at increasing depth (McCalpin, 1996a). Along the Concud normal fault, near the village of Concud and the city of Teruel (Spain), Quaternary sediments outcropping on the down-dropped block show evidence of several events of seismic movement simultaneous to faulting (Gutierrez *et al*., 2005). The oldest event is indicated by faults that displace the QU2 unit but not the overlying QU3 unit. Deformation simultaneous with faulting occurred between 62 and 71.6 Ka. It took several more recent events (events, 2, 3, and 4) to fill three fissures that extend into unit QU3. These fissure fillings are bounded by well-defined surfaces and show very clear differences of texture and color. The two oldest contain large blocks that fell from terrace deposits as rock falls. These events must have occurred after 62 Ka. If four large seismic events occurred over 71.7 Ka, the recurrence interval on this fault is less than 18 Ka. Obviously the meager chronological information on the paleoseismic history of the Concud Fault is not sufficient for us to draw conclusions on the precise frequency of large earthquakes. Supposing that the fault that ruptured during the earthquake is about 23 km long, we can use the empirical ratio of Wells and Coppersmith (1994) to anticipate earthquakes with magnitudes around 6.6 along this structure. In the eastern area of the Jiloca half graben, a set of faults indicates recent activity (Capote *et al*., 1981, Gutierrez *et al*., 1983a and b; Gutierrez *et al*., 2005). In the Rubielos de Cerida region numerous carbonate outcrops are compartmentalized by N-S valley fill. These elongate depressions have been interpreted as tectonic valleys (Gutierrez *et al*., 1983a) generated by Quaternary activity along a dominant system of normal synthetic faults and accessory normal antithetical faults (Capote *et al*., 1981). The latter are oblique to the principal NW-SE faults (Palomera Fault,), which define the eastern margin of the Jiloca half graben. These narrow grabens are less than 4 km long and 100 to 800 m wide. At the margin of tectonic valleys, the profile of the rocky slope generally shows a small scarp at the contact with the valley fill, usually not more than a meter high. A quarry located about 2 km northwest of Rubielos de Cerida (Spain) was excavated in 1980, exposing the location of a fault plane with carbonate-cemented fault breccia. Slickensides, striations, and grooves with relatively long wavelengths are present at the surface. The fault has displaced 50 cm of calcareous regolith present on the slope of Jurassic material and on top of the colluvium. It is not a fresh scarp but the slope changes its gradient from about 22o to 30o. The colluvial sequence is cemented and dips 10o practically parallel to the topographic surface. This unit of stratified detrital material is interpreted as *sheetwash* from the Jurassic slope. There are no wedges or other features that would indicate episodes of movement simultaneous with seismic activity. At this site, fragments of charcoal were collected from 2.1 m and 7.9 m below the surface. 14C dating has shown that the most recent sample is 43,070 ± 1200 BP and the oldest is > 48,500 BP, beyond the scope of the method (Gutierrez *et al*., 2005). Using the age of the first sample, we obtain a vertical displacement velocity of 0.05 to 0.07 mm/year. This number suggests slow movement along the fault and leads us to believe that earthquakes of magnitudes greater than 5.6 are not to be expected (Wells and Coppersmith, 1994; Anderson *et al*., 1996; Villamayor and Berryman, 1999). Until now we have analyzed vertical movement but in a river significant lateral movements of a channel can give rise to **asymmetrical valleys**. This asymmetry may be related to tectonic influences. Active tilting and movement toward the fault can affect riverbeds in half grabens (Leeder and Gawthorpe, 1987). Leeder and Alexander (1987) state that the floodplain of the Madison River contains oxbow remains that have a concave orientation toward the present position of the river. According to dating that has been done, the total amount of river migration since 7000 years BP is 1600 m, indicating an average displacement of 23 cm/year. According to these authors, the drift is due to tilting; this agrees with the concavity toward the northeast in the abandoned meanders. The lower reach of the Indus River in Pakistan is made up of numerous channel systems and wide floodplains over which ancient riverbeds flowed. Rivers change from braided to anastomosing or meandering when they cross various active structures. The river has avulsed numerous times across the floodplain in recent times. The zones where avulsion is most frequent are near significant tectonic landforms (Jorgensen *et al*., 1993). **Fluvial terraces** are the morphologies most often used to detect and measure tectonic activity during the Pleistocene. These terraces are flat and thus they easily preserve fault scarps. A fault that cuts a terrace is obviously more recent than the terrace; on the contrary, a fault that does not cut the terrace postdates the terrace. When doming occurs, sedimentation occurs in low-lying areas and rivers downcut in elevated areas. Sedimentation and entrenching are processes that keep the profile in equilibrium. The morphology of deformed terraces shows the nature of the deformation. For this reason, longitudinal profiles of terraces have been used to detect deformation (Bullard and Lettis, 1993). Deformed lake terraces can also indicate isostatic and tectonic activity. The most basic example is the deformation of the Lake Bonneville terraces in Utah (Crittenden, 1963). Mountain fronts in arid and semiarid zones show the development of **alluvial fans**. High values of *Smf* and *Vf* for the mountain front uplift rate indicate relatively low uplift rates for mountain front activity (Keller and Pinter, 1996). The situation is more complex when the alluvial fan is tilted. If the fan tilts farther than the threshold, the fan is confined and a new fan segment forms farther from the mountain front. A study of alluvial fans in Death Valley shows that tilting produces segmented fans (Hooke, 1972).

DEFORMATION IN LITTORAL AND LACUSTRINE ENVIRONMENTS

Littoral landforms are used to study tectonic deformation because sea level is a reference level that can be used to determine the age of these landforms (Keller and Pinter, 1996; Bull, 2007; Nelson. 2007). On a global scale, sea level has remained essentially constant for about 6000 years; this makes it possible calculate displacement with respect to sea level. Along coasts of where immersion is limited by **coral reefs**, sea level rises and reefs are eroded by waves (Chappell, 1974). The shape of the resulting platforms depends on the reef geometry, its orientation to the waves, and the time between successive uplift events. The amount of uplift during successive seismic events can be easily obtained from the height of benches and notches compared to analogous features formed along the present day high tide line. The presence of corals colonizing the benches and notches provides a rapid method of dating (Burbank and Anderson, 2001). Erosional coasts have an **abrasion platform** that has a flat surface and a cliff on the landward side. Coastal uplift creates a hanging abrasion platform or plateau (called rasa in Spanish) (Hernandez-Pacheco, 1950). The platform surface may have detrital sediments and fossils that can be used for age dating. Successive pulses of uplift create a stairway of platforms, such as the well-developed stairway on the Spanish Cantabrian coast, Almeria, the Huon Peninsula (New Guinea), and in southern Peru. Faults frequently cut **marine terraces**. When this happens, the fault movement postdates terrace formation. This morphology can be correlated over long distances. Variations in terrace height indicate modifications of the uplift rate. Along the Atlantic coast of the United States, one single level terrace ranged in height from less than 50 m to about 100 m, reflecting postglacial rebound that has migrated northward over the last 18,000 years (Winker and Howard, 1977). Along the Estrecho coast of Gibraltar from Cadiz to Malaga in Spain, the altimetric distribution of marine terraces of isotopic stages 5e and 5c shows symmetrical doming of the coastline with a maximum of +19 m in Tarifa and minimum of +1 m in the Cadiz and Malaga areas (Zazo *et al*., 1999b). In addition, these same terraces have different elevations at the intersection of important faults and the coastline, with up to 6 m of vertical displacement, indicating that they have been active during the last 95,000 years. A preliminary study was recently conducted in preparation for the construction of the underwater Gibraltar Tunnel to unite Europe and Africa (Silva *et al*., 2006). The criterion for evaluating fault activity has been seismicity, surface faults, geomorphology of fault traces, identification of paleoseismic fractures, and evaluation of uplift rates on coastal segments. The study included detailed cartography that shows that the NE-SW Cabo de Gracia strike-slip fault has been active during at least the last 128 Ka and that it probably caused moderate events affecting the Roman city of Baelo Claudia. Deformation studies also analyze the position of lake coastlines after isostatic rebound, such as studies done at Lake Bonneville (Crittenden, 1963) and the Lahontan lakes (Hanks and Wallace, 1985). In tectonically active zones such as the Dead Sea, Holocene fan-delta deposits appear to be highly deformed (Galit *et al*., 1995) and synthetic and antithetic fault systems can be observed (Guttierez, 2013).

UNIT-3: CONCEPTS IN GEOMORPHOLOGY: SPATIAL SCALE, TEMPORAL SCALE, SYSTEMS, FEEDBACK, EQUILIBRIUM AND THRESHOLD

A system is a collection of related objects and the processes that link those objects together. Within fluvial systems, objects such as hillslopes, the channel network and floodplains are linked together by the processes that move water and sediment between them. In common with other systems, the fluvial system is hierarchical, in that there are integrated sub-systems operating within it. In this chapter you will learn about:

- How energy and materials move through the system.
- Fluvial system variables (e.g. channel slope, discharge, bedload transport rate).
- The way in which some variables control others and how these relationships depend on the space and time scales considered.
- How feedbacks between variables can counteract or enhance system response to change.
- The role of thresholds in system behaviour.
- How equilibrium can be defined at different time scales.

INPUTS, OUTPUTS AND STORES

The basic unit of the fluvial system is the drainage basin. Fluvial systems are open systems, which means that energy and materials are exchanged with the surrounding environment. In closed systems, only energy is exchanged with the surrounding environment.

Inputs

The main inputs to the system are water and sediment derived from the breakdown of the underlying rocks. Additional inputs include biological material and solutes derived from atmospheric inputs, rock weathering and the breakdown of organic material. Most of the energy required to drive the system is provided by the atmospheric processes that lift and condense the water that falls as precipitation over the drainage basin. The pull of gravity then moves this water downslope, creating a flow of energy through the system. This energy is expended in moving water and sediment to river channels and through the channel network.

Outputs

Water and sediment move through the system to the drainage basin outlet, where material is discharged to the ocean. Not all rivers reach the ocean; some flow into inland lakes and seas, while others, such as the Okavango River in Botswana, dry up before reaching the ocean. This reflects another important output from fluvial systems: the loss of water by evaporation to the atmosphere. Most of the available energy is used in overcoming the considerable frictional forces involved in moving water and sediment from hillslopes into channels and through the channel network. Much of this energy is 'lost' to the atmosphere in the form of heat.

Stores

A certain amount of material is stored along the way. For example, water is stored for varying lengths of time in lakes and reservoirs, and below the ground in the soil and aquifers. Sediment is stored when it is deposited in channels, lake basins, deltas, alluvial fans and on floodplains. This material may be released from storage at a later stage, perhaps when a channel migrates across its floodplain, eroding into formerly deposited sediments which are then carried downstream. Ferguson (1981) describes the channel as 'a jerky conveyor belt', since sediment is transferred intermittently seawards.

TYPES OF SYSTEMS:

Three types of system can be identified in fluvial geomorphology. These are morphological systems, cascading systems and process–response systems.

Morphological (form) systems

Landforms such as channels, hillslopes and floodplains form a morphological system, also referred to as a **form system**. The form of each component of a morphological system is related to the form of the other components in the system. For example, if the streams in the headwaters of a drainage basin are closely spaced, the hillslopes dividing them are steeper than they would be if the streams were further apart from each other. Relationships such as this can be quantified statistically.

Cascading (process) systems

The components of the morphological system are linked by a cascading system, which refers to the flow of water and sediment through the morphological system. Cascading systems are also called **process systems** or **flow systems**. These flows follow interconnected pathways from hillslopes to channels and through the channel network.

Process–response systems

The two systems interact as a process–response system. This describes the adjustments between the processes of the cascading system and the forms of the morphological system. There is a two-way feedback between process and form. In other words, processes shape forms and forms influence the way in which processes operate (rates and intensity). This can be seen where a steep section of channel causes high flow velocities and increased rates of erosion. Over time erosion is focused at this steep section and the channel slope is reduced. Velocity decreases as a result, reducing rates of erosion. In order to examine the components of the fluvial system in more detail, it can be divided into sub-systems, each operating as a system within the integrated whole. One way of doing this is to consider the system in terms of three zones, each of which is a process–response system with its own inputs and outputs. Within each zone certain processes dominate. The sediment **production zone** in the headwater regions is where most of the sediment originates, being supplied to the channel network from the bordering hillslopes by processes of erosion and the mass movement of weathered rock material. This sediment is then moved through the channel network in the sediment **transfer zone**, where the links between the channel and bordering hillslopes, and hence sediment production, are not so strong. As the river approaches the ocean, its gradient declines and the energy available for sediment transport is greatly reduced in the sediment **deposition zone**. It is primarily the finest sediment that reaches the ocean, as coarser sediment tends to be deposited further upstream. In fact, only a certain proportion of all the sediment that is produced within a drainage basin actually reaches the basin outlet.

FLUVIAL SYSTEMS VARIABLE:

Variables are quantities whose values change through time. They include such things as drainage density, hillslope angle, soil type, flow discharge, sediment yield, channel pattern and channel depth.

Internal and external variables

An important distinction exists between internal and external variables. All the examples given above are **internal variables**, which operate within the fluvial system. Internal variables are influenced by other internal variables, and also by variables that originate from outside the system. These **external variables**, such as climate, control or regulate the way in which the system operates. Unlike the internal variables, external variables operate independently, in that they are not influenced by what is going on inside the fluvial system. At the basin scale the external variables are climate, base level, tectonics and human activity. If you are considering a sub-system, such as a reach of channel in the transfer zone, the external variables would include the supply of flow and sediment to the channel. This is because these variables originate from outside the channel sub-system, even though they are internal variables at the basin scale. To avoid confusion, the 'ultimate' external variables – climate, base level, tectonics and human activity – will be referred to as **external basin controls**.

The external basin controls

The variables defined in this section act as regulators of the whole system. Any change in one of these variables will lead to a complex sequence of changes and adjustments within the fluvial system.

● ***Climate*** describes the fluctuations in average weather. Although the weather is always changing, longer-term characteristics such as seasonal and inter-annual variations can be defined. Other characteristics include how often storms of a given size can be expected to occur and the frequency and duration of droughts. Where no long-term changes are occurring in the climate, the combination of such attributes defines an envelope of 'normal' behaviour. **Climate change** occurs when this envelope shifts and a new range of climatic conditions arises.

● ***Tectonics*** refers to the internal forces that deform the Earth's crust. These forces can lead to large scale uplift, localised subsidence, warping, tilting, fracturing and faulting. Where uplift has occurred inputs of water have to be lifted to a greater elevation, increasing energy availability; some of the highest rates of sediment production in the world are associated with areas of tectonic uplift. Valley gradients are altered by faulting and localised uplift, which may in turn affect channel pattern. Lateral (sideways) tilting can cause channel migration and affect patterns of valley sedimentation.

● ***Base level*** is the level below which a channel cannot erode. In most cases this is sea level. If there is a fall in sea level relative to the land surface, more energy is available to drive flow and sediment movement. Conversely a relative rise in base level means that less energy is available, resulting in net deposition in the lower reaches of the channel. Over time these effects may be propagated upstream through a complex sequence of internal adjustments and feedbacks.

● ***Human activity*** has had an increasing influence on fluvial systems over the last 5,000 years, especially during recent times. Activities within the drainage basin such as deforestation, agriculture and mining operations all affect the flow of water and production of sediment. These are referred to as indirect or diffuse activities. River channels are also modified directly when channel engineering is carried out. Advances in technology over the last century have meant that dam construction, channel enlargement for navigation and flood control, channel realignment, the building of flood embankments and other engineering works can now be carried out at an unprecedented scale. Today there are very few rivers that have not been affected in some way by the direct and indirect effects of human activity. It can be argued that, under some circumstances, human activity can be considered to be *both* an internal and an external variable. Many of the direct modifications described above are in response to some local human perception of the system. For example, channels are dredged because they are not deep enough for navigation, or flood defence works carried out because floods occur too frequently. Urban (2002) suggests that direct human intervention can often be classified as an internal variable, although it is more appropriate to consider *indirect* human activities as external. Some internal variables have a greater degree of independence in that they are only affected in a limited way by the fluvial system. These variables are geology, soils and vegetation and topography (which includes relief, altitude and drainage basin size). All are internal variables because they are controlled to some extent by the external basin controls, however their main influence on the operation of the fluvial system is a controlling one.

Adjustable (dependent) and controlling (independent) variables

From the discussion above it can be seen that some variables control the adjustment of other variables. For example channel pattern is, among other things, affected by the supply of sediment to the channel. In this case, channel pattern is the **adjustable** or **dependent variable** while sediment supply is the controlling or independent variable. Things can get a little confusing because controlling variables may in turn be adjusted by other variables. Extending the previous example, sediment supply is itself controlled by hillslope vegetation cover. In this case, sediment supply is the adjustable variable and vegetation cover the controlling variable. All internal variables are adjustable because their operation is ultimately regulated by the external basin controls. They are also influenced to a greater or lesser extent by other internal variables. Because the relationships between variables are so complicated, it can be very difficult to isolate the effect of one variable on another. The hierarchical nature of the fluvial system means that variables operating at larger scales tend to affect the operation of variables at smaller scales. For example, climate affects vegetation cover and hillslope erosion, which in turn determine sediment supply, which influences channel pattern, which affects the small-scale flow dynamics in the channel, which governs the movement of individual grains. This is not a one-way process, however. Over long periods of time, the cumulative effect of small-scale processes, such as the erosion and deposition of individual grains, can lead to larger scale changes. These include changes in channel pattern and, over time periods of tens to hundreds of thousands of years, can adjust the slope of the whole river valley. Time itself is an important controlling variable. Every drainage basin has a historical legacy resulting from past changes that have taken place in the basin. This includes the cumulative effect of processes such as erosion, transport and deposition over long periods of time. It also includes the far-reaching effects of changes in the external basin controls, such as the variations in climatic conditions since the Last Glacial Maximum 18,000 years ago, which have greatly affected fluvial systems worldwide. In the temperate zone, many rivers underwent a transition from a braided to a meandering form as climate conditions ameliorated, vegetation became established and sediment loads decreased. However, vast quantities of sediment still remain in formerly glaciated drainage basins, where many fluvial systems are still adjusting to this glacial legacy.

Feedbacks

A **feedback** occurs when a change in one variable leads to a change in one or more other variables, which acts to either counteract or reinforce the effects of the original change. Two types of feedback are observed: negative feedback and positive feedback. Both are initiated by a change in one of the system variables, which in turn leads to a sequence of adjustments that eventually counteract the effect of the original change (**negative feedback**) or enhance it (**positive feedback**). When there is a change in one of the external controls, negative feedbacks allow the system to recover, damping out the effect of the change. An everyday example of a negative feedback loop is a central heating system controlled by a thermostat, which switches the source of heat on and off as the room cools and warms. An equilibrium is maintained as the temperature fluctuates around an average value. A commonly cited example of negative feedback within the fluvial system occurs when a section of channel is suddenly steepened by tectonic faulting. This leads to a local increase in the flow velocity and rate of bed erosion. Over time this acts to reduce the channel slope, counteracting the effects of the original change. It should be noted that the actual sequence of events is usually rather more complex. This is because change in one part of the system can lead to complex changes, both locally and throughout the rest of the system. The nature of complex response will be discussed later in this chapter. Positive feedbacks have a very different effect. Soil erosion is a natural process and an equilibrium exists if rates of soil removal over a given period are balanced by rates of soil formation over that period. However, an external change, such as the deforestation of steep slopes, can lead to a dramatic increase in soil erosion. The upper soil layers contain the most organic matter, which is important in binding the soil together. It also increases soil permeability, allowing rainfall to soak into the soil rather than running over and eroding the soil surface. If the topsoil is removed, the lower permeability of the underlying soil layers means that more water runs over the surface, increasing erosion and removing still more soil layers. In this way several centimetres of soil can be removed by a single rain storm (Woodward and Foster, 1997). This greatly exceeds the rate of soil formation. Referred to colloquially as 'vicious circles' or 'the snowball effect', positive feedbacks involve a move away from an equilibrium state. They usually involve the crossing of a **threshold** as the system moves towards a new equilibrium. A small-scale example of positive feedback is the build up of sediment during the formation of a channel bar. Bar formation is initiated when bedload sediment is deposited at a particular location on the channel bed. This affects local flow dynamics, causing the flow to diverge over and around the initial deposit. As the flow diverges, it becomes less concentrated and therefore less able to transport the coarser sediment. Localised deposition occurs, further disrupting the flow and promoting further deposition and bar growth. Several feedbacks, both positive and negative, exist between channel form, water flow and sediment transport. The form of a channel has an important influence on the way that water and sediment move through it. For example, flow is concentrated where the channel narrows, increasing erosion potential. The character of the channel bed is also significant, since the size and arrangement of sediment determines bed roughness and resistance to flow. Where resistance is high, the average velocity of flow in the channel is reduced. This influences hydraulic conditions near the bed of the channel, which are significant for processes of erosion and deposition. Considerable differences are seen across the channel bed, giving rise to spatial variations in erosion and deposition. These processes themselves modify the form of the channel, feeding back to influence flow.

Thresholds

Thresholds are another important concept in systems theory and you will come across many examples in fluvial geomorphology. For example, a threshold is crossed when a sand grain on the bed of the channel is entrained (set in motion). Movement is resisted by the submerged weight of the grain and friction between it and the neighbouring grains. If the driving force exerted on the grain by the flow is less than these resisting forces, no movement will occur. It is only when the driving force of the flow exceeds the submerged weight of the grain that entrainment will take place. In this example, channel flow is an external variable. When a threshold is crossed there is a sudden change in the system, for example when loose material on a slope becomes unstable and starts to move down the slope as a landslide. The gradual processes by which rock is broken down and loose material builds up on a hillslope take place over time scales of tens to hundreds of years. Why, then, does a landslide occur at a particular point in time? Such a transition can come about when a change in one of the controlling variables leads to instability within the system – as a direct result of an earthquake (tectonics) for example. Thresholds that are crossed as a result of external change are called **external thresholds**. Instabilities may also develop over time without any external change having occurred. For this reason it is possible for a major landslide to be triggered by a relatively minor rainfall event that falls well within the expected climatic norm, because instability has gradually developed over time and the system is ready 'primed'. This is an example of an **internal threshold**. Another commonly cited example is the threshold that is crossed when a meander loop is cut off to form an oxbow lake. Again this is something that can take place without there having been any change in the external variables. Whether or not either kind of threshold is crossed depends on how 'sensitive' the system is, in other words how close to a threshold it is. To illustrate this point, consider a pan filled with water that is heated by 10°C at normal atmospheric pressure. If the water had an initial temperature of 25°C you would not expect to see much change in its appearance. However, if the initial temperature was 90°C, the same increase in temperature would lead to a dramatic change as a threshold was crossed and the water started to boil. In the second example, the sensitivity of the system is much greater because the water temperature is closer to boiling point. Once a threshold has been crossed, the system reaches a new equilibrium. Relating this back to the sand grain example, two scenarios could be considered. In the first, the particle is resting on a flat bed and is fully exposed to the flow. In the second, it is buried beneath a layer of gravel. The threshold for movement will be much lower for the exposed grain than for the buried grain, which cannot move unless the overlying gravel is removed.

Complex response

The response of the fluvial system to change is often complex because of the many interrelationships that exist between the different components of the system. An example is the complex response of a tributary to a lowering in base level elevation at its outlet (Schumm, 1977). Here the main river, into which the tributary flows, has degraded, or lowered its channel elevation by erosion. This leads to a complex sequence of episodic erosion and deposition in the tributary as the system searches for a new equilibrium.

● *Stage 1.* A fall in base level results in a local steepening of the channel gradient at the downstream end. This might exist as a sharp break in channel slope or as a steeper section of channel. Erosion is focussed at this steeper part, leading to a localised lowering in channel elevation. As a result of this lowering, the section of channel immediately upstream becomes steeper. The steep zone therefore migrates upstream over time. The initial response of the channel to the fall in base level is therefore incision (downcutting) at the downstream end. At this stage the upstream part of the channel is unaffected, so there is no change in the supply of sediment coming from upstream.

● *Stage 2*. Over time, incision is propagated upstream. As a result there is an increase in the amount of sediment coming from further upstream.

● *Stage 3*. This increase in sediment supply from upstream leads in turn to a build up of sediment in the channel at the downstream end. This process is called **aggradation** and results in a reduction in channel gradient. Meanwhile, the zone of incision continues to migrate further upstream.

● *Stage 4*. Aggradation is propagated upstream. Since there is now a net deposition of sediment in the upstream part of the channel, there is a decrease in the volume of sediment supplied from upstream.

● *Stage 5*. At the downstream end, less sediment means that there is additional energy for erosion and a new cycle of incision begins. Aggradation continues to occur at the upstream end.

● *Stage 6*. The wave of incision extends upstream, once again increasing the supply of sediment from upstream.

● *Stage 7*. The increased sediment supply leads to aggradation at the downstream end of the channel. This example illustrates the way in which downstream changes can be propagated upstream, while the resultant upstream changes in turn control what happens downstream. In this way there is a complex cycle as periods of **episodic erosion** are interspersed with periods of deposition.

SCALE IN FLUVIAL GEOMORPHOLOGY

Scale is an important consideration in fluvial geomorphology, with process–form interactions occurring over a huge range of space and time scales. At one end of this range is the long-term evolution of the landscape. At the other are small-scale processes, such as the setting in motion of an individual grain of sand resting on the bed of a channel. Space scales therefore encompass anything from a few millimetres to hundreds of kilometres. Relevant time scales stretch from a few seconds to hundreds of thousands of years or more. In order to understand how the fluvial system operates we can examine the relationships between processes and form in more detail at finer scales. This can be done by examining individual sub-systems, or sub-systems within sub-systems. When focusing in like this it is important to remember that these sub-systems are all part of an integrated whole and therefore cannot be considered in isolation from the rest of the system.

Space scales (spatial scales)

In studying the fluvial system, the scale of relevance varies according to the type of investigation. At the largest, drainage basin scale, it is possible to see the form and characteristics of the drainage network and drainage basin topography. These reflect the cumulative effect of processes operating over long time scales, as well as past changes imposed by the external basin controls. At a smaller scale, the form of a reach of meandering channel can be examined in the context of drainage basin history and the influence of controlling variables at the channel scale, such as the supply of water and sediment from upstream. The way in which the form and position of the channel has changed over time scales extending to thousands of years may be preserved as floodplain deposits, which can be used in reconstructing drainage basin history. Moving in to look at an individual meander bend, process–form interactions can be observed at a smaller scale. These include flow hydraulics within the bend and associated sediment dynamics. Investigations of rates of bend migration or bank erosion processes are also carried out at this scale. Depositional channel units such as the point bar are of interest to sedimentologists, providing evidence about the flows that formed them. At a finer scale still are individual ripples on the bar surface formed by the most recent high flow and, moving even closer, the internal arrangement of grains. At the finest scale are individual grains of sediment.

Time scales (temporal scales) and equilibrium

At smaller spatial scales, process–form interactions generally result in more rapid adjustments. At the largest scale, the long term evolution of channel networks occurs over time scales of hundreds of thousands of years or more, while the migration of individual meander bends can be observed over periods of years or decades, and small-scale flow-sediment interactions within minutes. The perspective of the historically oriented geomorphologist concerned with the large-scale, long-term evolution of landforms is therefore very different to that of the process geomorphologist or engineer who is interested in the operation of channel processes at much shorter time scales (Schumm, 1988). Historical studies show that the fluvial system follows an evolutionary sequence of development that is interrupted by major changes induced by the external basin controls. However, over the much shorter time periods involved in the field measurement of processes, there may be little or no significant change in fluvial landforms. This might not matter too much if flow–sediment interactions at very small scales are of interest, although basin history certainly does have an influence at the reach scale, since channel form has been shaped by past changes in flow and sediment supply. The precise definition of equilibrium is also time dependent. Equilibrium refers to a state of balance within a system, or sub-system. Negative feedback mechanisms help to maintain the system in an equilibrium state, buffering the effect of changes in the external variables. However, different types of equilibrium may exist at different time scales. These were defined by Schumm (1977) with reference to changes in the elevation of the bed of a river channel above sea level. If you were to observe a short section of river channel over a period of a few hours you would not see any change in its form (unless there happened to be a flood), although you might see some sediment transport. Over this short time period the channel is said to be in a state of **static equilibrium**. The same river, observed over a longer time scale of a decade, would show some changes. During this time, floods of various sizes pass through the channel, scouring the bed. In the intervening periods, deposition builds up the channel bed again. As a result of these cycles of scour and fill, the elevation of the channel bed fluctuates around a constant average value and **steady state equilibrium** exists. Over longer time scales, from thousands to hundreds of thousands of years or more, erosion gradually lowers the landscape. At these time scales, the channel elevation fluctuates around a changing average condition, the underlying trend being a reduction in channel elevation. As you know, the influence of the external basin controls cannot be ignored. Changes in any of these variables can lead to positive feedbacks within the system and a shift to a new equilibrium state. For example, in tectonically active regions, the section of channel might be elevated by localised uplift. Such episodes of change occur over much shorter time scales than the gradual evolution of the landscape, resulting in abrupt transitions. This type of equilibrium delights in the term **dynamic metastable equilibrium** (Charlton, 2008)**.**

UNIT-4: CATCHMENT PROCESS AND FLUVIAL PROCESSES: FACTORS REGULATING ENTRAINMENT, TRANSPORTATION AND DEPOSITION

SEDIMENT ENTRAINMENT AND TRANSPORT

The process of particle entrainment

Whether or not a given particle is set in motion depends on the balance between the forces driving and resisting its movement. The driving force is provided by the combined effect of two **fluid forces** exerted on the particle by the flow: a drag force and a lift force. The fluid **drag force** acts in the same direction as the flow and can be thought of as the 'force of the flow' that is felt when you wade out into the current of a stream. It comes about because the pressure exerted on an object by the flow is greater on its upstream side than on its more sheltered downstream side. The second fluid force, the **lift force**, acts vertically upwards and is caused by a pressure difference above and below the particle. Water flowing over the particle has to move faster. According to the Bernouilli principle, an increase in velocity results in a decrease in pressure above the particle, while the pressure below it stays the same. This difference in pressure generates lift. In theory, if this force exceeds the gravitational force, the particle will be lifted from the bed. In practice, the presence of other particles complicates matters considerably. Sediment transported as bedload is generally gravelsize and larger, although coarse sands may also form part, or all, of the bedload component. Finer bedload, which is too coarse to be transported in suspension, is moved along the bed in a series of short jumps by **saltation**. Saltating grains are lifted from the bed at a relatively steep angle by the combined forces of lift and drag. As a grain moves upwards into the flow, the lift force decreases and it starts to fall back towards the bed. The falling grain is carried downstream by the drag force, following a shallow trajectory towards the bed. Larger particles, which cannot be lifted, are rolled or dragged along the bed. This movement is usually sporadic because of variations in bed shear stress. In addition, particles tend to become lodged behind other particles or obstacles on the bed. The weight of smaller particles carried in suspension is supported by turbulence. Descending saltating grains may also be temporality lifted upwards by turbulent movements. This is called incipient suspension.

Size-selective theories of sediment transport

A considerable amount of research has focused on deriving critical flow or **entrainment thresholds** from easily measured flow parameters. There are various practical reasons why we might want to know the flow conditions that will move particles of a certain size. These include the planning of reservoir releases to flush out fine sediment from fish spawning grounds (without removing the gravel), or determining when structures such as bridge piers are at risk of being undermined by erosion. The threshold conditions for the entrainment of particles of a given size can be defined according to a critical mean flow velocity (i.e. cross-sectional average) or a critical bed shear stress. Using the mean flow velocity is an indirect method, since it is actually the hydraulic conditions near the bed of the channel that are significant. However, both relationships show similar basic trends. An explanation will first be given in terms of a critical mean flow velocity since this relationship is conceptually easier to understand. The critical mean flow velocity curves were derived from a large amount of experimental data accumulated by Filip Hjulstrøm in the 1930s. They show the entrainment and fall (or settling) velocities for particles of different sizes, from fine clay to coarse gravel and small boulders. Note that a logarithmic scale is used on both axes to cover the wide range of particle sizes and the corresponding range of flow velocities. The upper curve on the graph shows the **entrainment velocity** required to set different particle sizes in motion. Sand grains, with a diameter of between 0.2 mm and 0.7 mm, are the easiest to entrain. In the case of larger particles, which have a greater immersed weight, the entrainment velocity increases with particle size as might be expected. However, the relationship is rather different for particles smaller than 0.2 mm, since the entrainment velocity actually *increases* as the particle size decreases from fine sand to silt and clay. Reasons for this include the fact that these small particles tend to be partly or wholly enclosed within the laminar sublayer during most flows. Drag forces are lower within this layer, and particles are not exposed to turbulent lift forces. In addition, the cohesive forces between clay particles further increases the force required to set them in motion. An alternative approach, which is more relevant to modern sediment transport theory, was devised by the American engineer Albert Shields in 1936. This defines the critical bed shear stress necessary to set particles of a given size in motion. The critical bed shear stress is actually defined in a dimensionless form. The dimensionless critical bed shear stress is often referred to as the **Shields parameter**. It appears in a number of sediment transport equations and is represented by the Greek letter theta (θc – the subscript is short for 'critical'). Critical bed shear stress increases with particle size but also depends on bed roughness. Shields related the dimensionless bed shear stress to the boundary Reynolds number. The boundary Reynolds number is proportional to the ratio between grain size and laminar sublayer thickness. For hydraulically rough surfaces, the critical bed shear stress is independent of the boundary Reynolds number and the critical bed shear stress reaches a constant value of 0.06 (Richards, 1982). The lowest critical bed shear stress is associated with sand grains in the size range 0.2 mm to 0.7 mm (Knighton, 1998). It is important to note that the Hjulstrom and Shields experiments were carried out using well sorted bed sediment of a single size. This is not representative of the conditions on the bed of many channels, where there is a mixture of grain sizes. The arrangement of grains on the bed and the mixture of grain sizes is very significant, affecting both the entrainment of individual grains and overall transport rates.

Sediment transport in mixed beds

The mobility of individual particles is greatly affected by the size and arrangement of the particles surrounding them. In most natural channels the mixture of sediment sizes, and an irregular bed surface, makes the situation rather more complicated. This can be defined in terms of a **friction angle**, which is greatest where small particles overlie larger ones, meaning that a greater force is required to pivot smaller particles away from the bed. The degree of **sorting** reflects the range of particle sizes in a particular sample of bed material. Well-sorted sediments have a narrow range of particle sizes, whereas poorly sorted material shows a much wider range.

Armour layers

In gravel-bed rivers the development of an **armour layer** has a very significant impact on rates of bedload transport. As a result, fine sediment is removed from the bed leaving a layer of coarse sediment, usually about one particle diameter in thickness. This armour layer protects the finer material beneath from subsequent high flows. Once a bed is armoured, a much higher critical threshold is required to break it up. Bathurst (1987a) defined **'twophase' flow** for armoured channels. During phase 1 flow an armour layer is present and rates of bedload transport are low (although finer sediment can still be supplied from further upstream). Once the armour layer breaks up phase 2 transport takes place, with a dramatic increase in transport rates as the finer sediment becomes available. This can lead to complex variations in bedload transport through time. For example, where two high magnitude flow events occur in close succession, the initial rate of transport is often much higher for the second event than for the first, which breaks up the armour layer. Research has shown that ephemeral channels do not tend to develop an armour layer because, in the absence of low flows, there is no mechanism for removing fine sediment to create the armour layer. This could mean that rates of bedload transport are greater for ephemeral channels than for channels in humid settings (Nanson *et al*., 2002).

The theory of equal mobility transport

On a level bed with a uniform sediment size, all the particles might be expected to begin moving under approximately the same flow conditions (Reid *et al*., 1997). However, on mixed beds, the relative size of a given sediment particle determines its degree of exposure to the flow. As a result, larger particles shelter smaller particles, which then require a higher shear stress for entrainment than would otherwise be the case. In contrast, coarser grains are more easily entrained when surrounded by fine grains. This is because they are relatively more exposed to the forces of entrainment (Andrews, 1983). Particles of an intermediate size are relatively unaffected by the sheltering/hiding effects. Empirically, this 'reference size' has been shown to approximate the median size (*D*50) (Bathurst, 1987b). On the basis of field data, Parker *et al*. (1982) introduced a theory of **equal mobility** for channel beds composed of a mixture of sediment sizes. This states that the threshold condition for each size fraction is not dependent on the grain size. In other words, the movement of particles of different sizes can be initiated under similar critical flow conditions. The theory of equal mobility transport therefore challenges the size-selective transport theory of Shields (1936). However, any deviation away from an equal mobility condition represents some degree of size-selective transport. Under conditions of equal mobility, the bedload transport rate could be calculated from a single representative grain diameter such as the median size, *D*50 (Parker *et al*., 1982). Equal mobility transport is the subject of some debate, however. Field investigations into the occurrence of equal-mobility transport have mainly been carried out in gravel-bed channels, where the largest grains are cobblesize or smaller (for example Andrews, 1983; Ashworth and Ferguson, 1989). Measurements made in rivers during steady uniform flows have shown that the transport of mixed sediment is only weakly size selective at low shear stresses. At higher shear stresses, sediment transport approaches equal mobility (Parker *et al*., 1982; Andrews, 1983; Marion and Weirach, 2003). However, observations made over a wider range of flows (e.g. Ashworth and Ferguson, 1989; Wilcock, 1992) have emphasised the size-selective nature of gravel transport. Only during the highest flows does sediment transport approach equal mobility. For example, Wilcock (1992) observed a progressive shift away from unequal to equal mobility transport with increasing shear stress, although equal mobility was not observed until the shear stress was over twice the critical stress required to initiate motion. One of the biggest problems associated with these investigations is obtaining sufficient field data to include a representative range of flow conditions (Reid *et al*., 1997).

BEDLOAD TRANSPORT

Bedload transport does not necessarily take place all the time, and rates may approach zero during low flows. Even when transport is occurring, it is likely that only part of the bed will be mobile at any one time. Part of the reason for this is the uneven distribution of bed shear stresses, which is directly controlled by variations in turbulent fluctuations. The ejection of low momentum fluid away from the bed also allows finer sediment to be lifted up away from the bed and into the turbulent profile, maintaining it in suspension. The availability of bed sediment has an important influence on overall rates of bedload transport in a given reach of channel, and many bedload-dominated channels are transport limited. This means that transport rates might be lower than expected at a particular flow because of a lack of available sediment. This 'lack' does not necessarily refer to the total volume of bed sediment in a reach, more relevant is the availability of sediment of a certain size or calibre. Thus a flow that is competent to transport only fine gravels will not be able to entrain the larger material in a boulder-bed stream, no matter how abundant this is. The supply of bedload can be especially limited in bedrock channels and the flow capacity often exceeds that required to transport the available load.

Bedforms

Bedforms in sand-bed channels In sand-bed channels the sand grains can be transported at both high and low flows because of their low entrainment threshold. As a result the bed is easily shaped by flows to form periodic features known as bedforms. For the purpose of explanation, the starting point is assumed to be a plane bed, something that is rare in natural channels because even the smallest flows start to shape the bed. When water starts to flow over a flat bed the sand grains start to move, individually at first and then in patches, until periodic **ripples** develop, with crests perpendicular to the direction of flow. Field and laboratory research suggests that the wavelength, or spacing, of ripples is mainly dependent on particle size and is typically between 150 mm and 450 mm. As the flow intensity increases, ripples start to give way to **dunes**, larger features with rounded crests. Dunes are common in alluvial channels and are continuous along the bed for hundreds of kilometres in large rivers like the Mississippi and Niger. Dunes vary greatly in size, being scaled with the depth of flow and ranging from a few centimetres to a few metres in height. Dune wavelengths also vary, from tens of centimetres to more than a hundred metres in the largest rivers. Ripples and dunes migrate downstream over time, as the flow moves sand grains up the more gentle upstream slope towards the crest, from where sediment falls down the steeper downstream slope. A critical mechanism in this process is the deposition of coarse grains at the crest, where flow separation occurs. At higher flows, dunes become unstable and are 'washed out' because the flow velocity is too great to sustain deposition at the dune crest. Dunes then give way to a **plane bed**, but one that is rather different from the initial flat bed. Above the bed is a clearly defined zone of suspended sediment within which 'dust storm conditions' prevail (Leopold *et al*., 1964). This marks the transition to the **upper flow regime**, where the Froude number (ratio of inertial and gravitational forces) is greater than 1 and flow becomes supercritical. Upper flow regime bedforms include **standing wave antidunes**, where the sediment is moving but the waves themselves are stationary. This is because rates of deposition on the upstream side are matched by erosion on the downstream side. The position of standing waves is marked by waves at the water surface, the sand and water waves being in phase with each other. At higher flows sediment is thrown up from the downstream side of the bedforms at a faster rate than it can be replenished, which results in **antidunes**. These migrate *up*stream, while the sediment continues to move downstream. At very high flows, a series of **chutes and pools** develop. Chutes have a near-plane bed and shooting flow, which enters downstream pools: deeper sections that are marked by hydraulic jumps. Bedforms in gravel and mixed sand–gravel channels Bed structures also form in gravel-bed channels and have been a focus of research over recent decades. **Pebble clusters** are commonly found in this type of channel and form when a single large particle acts as an obstacle, protruding into the flow and encouraging the accumulation of coarse material on its upstream side. This upstream material may have an imbricated structure, increasing stability and requiring larger lift and drag forces to entrain the constituent particles. Finer particles are found on the downstream side of the obstacle, where shelter is provided from lift and drag forces. **Transverse ribs** are another type of gravel bedform and consist of regularly spaced ridges of coarser pebbles, cobbles or boulders that lie transverse to the flow. Like sand bedforms, these features affect flow resistance as well as rates of bedload transport. Where the bed is composed of a mixture of sand and gravel, the different mobility of the constituent particles can lead to some interesting effects. For example, longitudinal ribbons of sand have been observed to travel downstream, snaking from side to side over immobile gravel beds. Bedload can also move in thin sheets as an elongated procession of sediment with a thickness of one to two grain diameters. The coarsest sediment accumulates at the leading edge and there is a progressive fining of sediment behind it. Bedload sheets appear to be fairly common in mixed channels and are related to rates of sediment supply, becoming less frequent and reduced in extent as supply rates are reduced (Dietrich *et al*., 1989).

Assessing rates of bedload transport

From the preceding discussion you will have some idea of just how complex and variable bedload transport is. There is a general paucity of data on rates of transport because the available techniques can be expensive and time-consuming to employ. These include the collection of bedload over a period of time using portable samplers or traps excavated in the bed. Another approach is to track the movement of individual particles.

SUSPENDED LOAD TRANSPORT

Processes Particles carried in suspension are kept aloft by turbulent eddies and will remain in suspension as long as their weight is supported by the upward component of turbulent eddies. In a fluid at rest, a suspended particle will fall through the fluid column. The rate of fall, or **fall velocity**, is a function of the density, size and shape of the particle. It is also determined by the viscosity and density of the transporting fluid. Since the falling particle displaces fluid, its movement is resisted by an equal and opposite fluid drag force. If sufficient depth is available, the falling particle will accelerate until it reaches a terminal velocity. In channels, the fall velocity is further affected by flow turbulence and the interactions of surrounding particles (Chanson, 1999). Considerable variation is seen between particles of different sizes. The fall velocity for the finest wash load component is very low, meaning that this sediment can be transported over considerable distances. For example the terminal fall velocity of a silt grain (0.001 mm) is approximately 0.004 cm s-1, but increases to 34 cm s-1 for a 10 mm gravel particle (Chanson, 1999). Suspended sediment is transported by processes of **advection** and **turbulent diffusion**. Advection is the transport of sediment within the flow, where the sediment moves with the flow itself. Turbulent diffusion refers to the mixing of sediment through the depth profile by turbulent eddies. Within the depth profile, the greatest concentration of suspended sediment is found towards the bed of the channel. Although there is continuous movement of individual suspended grains, the overall concentration and average grain size generally decrease rapidly away from the bed. This is due to interaction between the fall velocity and the vertical component of flow associated with turbulent eddying (Knighton, 1998). The upward migration of sediment to zones of lower concentration is both an advective and a diffusion process. A related process, which is called **convection**, involves the entrainment of sediment by large-scale vortices. For example, sediment is suspended in vortices generated as a result of flow separation in the toughs of ripples and dunes (Bridge, 2003). Large-scale vortices also occur where there are sudden drops in bed elevation, at hydraulic jumps and during overbank flows.

Sediment supply and transport rates

The main sources of suspended sediment include material washed in from hillslope erosion and the release of fine material and aggregates from bank erosion. The supply of fine sediment is a major control on rates of suspended sediment transport. Most suspended transport, particularly the wash load, is supply limited. This means that the supply of fine sediment often has a greater influence on the sediment concentration than flow conditions in the channel. The rate of supply varies during individual events, between events, seasonally and annually. These variations are controlled by a number of variables, including antecedent conditions, rainfall intensity, hydrograph shape and vegetation growth. High discharges tend to be associated with greater concentrations of suspended sediment. This is because the supply is increased by storm-induced erosion of hillslopes and channel banks, and the release of fine sediment from storage. With all this in mind, it is hardly surprising that no simple relationship exists between suspended sediment concentration and flow discharge for a given crosssection.

DEPOSITION

Sediment particles are deposited when there is a reduction in the competence and capacity of the flow. The process itself takes place at a very small scale and involves individual grains, although depositional forms can be observed over a wide range of spatial scales, from the smallest bedforms to vast floodplains and deltas. The construction and development of depositional forms might be likened to the building of an anthill. The process of building the anthill involves individual ants carrying soil one crumb at a time to the site of the ant hill. Although this process takes place at a small scale, the resulting feature is much bigger than the individual ants and crumbs of soil that created it. Thresholds for deposition are associated with the fall (or settling) velocity defined earlier. The deposition of suspended sediment takes place when the fall velocity dominates over turbulent diffusion. Since the fall velocity is closely related to particle size, the coarsest sediment tends to be deposited first. This leads to sediment sorting, a vertical and horizontal gradation of sediment, from coarse to fine. It should be noted that the fall velocity is also affected by the viscosity and density of the fluid. These are both influenced by changes in suspended sediment concentration. In addition, finer material can be transported as agglomerations of sediment called flocs. These have a greater fall velocity than the individual particles forming them. In the case of bedload transport, the near-bed flow conditions are significant. Bedload deposition occurs where the bed shear stress drops below the critical shear stress (Shields's parameter) required to transport particles of a given size. Local patterns of sediment sorting are well known, for example a downstream reduction in bed particle size is commonly observed along channel bars (e.g. Bluck, 1982; Smith, 1974).

Where sediment is deposited

There are a number of different circumstances that lead to deposition. These include:

● ***Reductions in flow discharge*** which are seen as flows recede, or along dryland rivers, where downstream losses are caused by high rates of evaporation and percolation.

● ***Decreases in slope*** which can be localised, or involve a gradual reduction over a longer length of channel and cause a reduction in average flow velocity and stream power.

● ***Increases in cross-sectional area*** cause the flow to diverge and become less concentrated. Flow resistance increases because there is more contact between the flow and channel boundary. There is a large increase in cross-sectional area when overbank flows occur.

● ***Increases in boundary resistance*** are associated with vegetation and coarse bed sediment. When overbank flows occur, velocity is reduced by the increased roughness of the floodplain surface, leading to the deposition of suspended sediment.

● ***Flow separation*** which causes sediment to become decoupled from the flow.

● ***Obstructions to flow.*** Sediment often accumulates behind obstructions. These include boulders, outcrops or islands of bedrock, woody debris and man made structures such as bridge piers, dams and flow control structures. Changes in the supply of sediment are also important. For example, sediment tends to accumulate immediately downstream from scour zones caused by flow convergence, when the material scoured from the channel bed is deposited immediately downstream. At a larger scale, increases in the supply of sediment to a channel reach are caused by changes within the upstream drainage area.

Depositional environments

Although deposition does occur in the production and transfer zones of the fluvial system it dominates in the deposition zone, where there is a decline in gradient and energy availability. Large-scale deposition leads to the development of characteristic landforms, including floodplains, alluvial fans and deltas. Within channels, **bars** represent smaller-scale depositional features. They are commonly found on the inside of meander bends, along the edges of channels, and where tributaries join the main channel. Braided channels are characterised by numerous midchannel bars. **Floodplains** border the channels of alluvial rivers and are formed from a mixture of in-channel and overbank deposits. Their development and evolution, is governed by a number of factors, including the supply of sediment (volume and calibre), the energy environment of the channel, and the valley setting. Sediment is laid down by rivers as they migrate across the floodplain, being deposited on the inside of meander bends or when braid bars are abandoned. These channel deposits are relatively coarse in comparison with the much finer sediment that is laid down by overbank flows. Processes of erosion can also be significant in reworking sediment or in removing part, or all, of the floodplain surface. **Alluvial fans** are typically found in situations where an upland drainage basin flows out onto a wide plain. The sudden change from confined to unconfined conditions leads to flow divergence, while mean flow velocity is decreased by the reduction in slope. The resultant deposition leads to the formation of a conical feature with a convex cross-profile. Most fans have a radius of less than 8 km, but can be more than 100 km wide in some cases. Where a number of individual fans develop along a mountain front, they may grow laterally and coalesce to form a sloping apron of sediment called a **bajada**. Fans are commonly found in dry mountain regions, where an abundant sediment supply is associated with extreme discharges and frequent mass movements. Frequent shifts are often seen in the position of the braided channels that cross the fan surface, although only part of the fan surface may be active during a major flood event. In long profile, the slope is steepest at the fan head, progressively decreasing along the length of the fan. There is also a down-slope reduction in sediment size, although deposits are coarse and poorly sorted. Incision and fan head trenching is associated with decreases in sediment supply, or increases in slope. Such changes can be caused by tectonics, climatic variations, a fall in regional or local base level, or human activity. In the absence of external change, the progressive lowering of the landscape will also result in a decline in sediment yield over time. Arid fans are generally smaller and steeper than those found in humid regions, a large-scale humid example being the Kosi Fan on the southern Himalayan mountain front. This covers an area of 15,000 km2 and formed where the Kosi River descends onto the wide alluvial plain of the Indus. It has a very low gradient, only averaging 1 m km-1 at its head, with further decreases downstream (Summerfield, 1991). **Deltas** are found where sediment-charged flowing water enters a body of still water. They extend outwards from shorelines where rivers enter lakes, inland seas and oceans. In coastal areas deltas form where the supply of sediment is greater than the rate of marine erosion, although sediment is redistributed by coastal processes. The influence of fluvial processes tends to dominate in the case of lake deltas (Charlton, 2008).

UNIT-5: ADJUSTMENT OF CHANNEL FORMS AND PATTERNS TO MORPHODYNAMIC VARIABLES

The form of a channel is largely a function of the water and sediment supplied to it. Adjustments to channel form occur as a result of process feedbacks that exist between channel form, flow and sediment transport. At the reach scale, the type of adjustment that can take place is constrained by the valley setting, the nature of bed and bank materials, and bank vegetation. This gives rise to a wide diversity of different channel forms. In this chapter you will learn about:

- How the supply of water and sediment drive channel form adjustments
- Feedbacks between flow, channel form and sediment transport
- Types of channel adjustment that can be made within the boundaries imposed by local conditions
- Space and time scales over which adjustments are made

CONTROLS ON CHANNEL ADJUSTMENT AND FORM

Flow and sediment supply both fluctuate through time, meaning that continuous adjustment takes place through the erosion, reworking and deposition of sediment. The flow and sediment regimes are called **driving variables** because they drive these processes. Along a given reach, channel adjustment is constrained within certain boundaries that are imposed by local conditions. For example, a sand-bed river flowing across a wide floodplain is able to adjust its form much more readily than a bedrock channel confined within a narrow gorge. Energy availability is also important, and channel adjustments are often limited for rivers that flow over low gradients, especially where cohesive banks are protected by vegetation. These constraints are called **boundary conditions** and include valley confinement, channel substrate, valley slope and riparian vegetation. Not all channels are in regime, and there are many examples of non-regime, or disequilibrium, channels. This may be because the channel is evolving in response to long term changes in the flow or sediment regime, caused by a change in one of the external basin controls. Examples include incising or aggrading channels and those that are undergoing a change in channel pattern. Alternatively, some bedrock and dryland channels may exist in a permanent state of disequilibrium because it is only during flood flows that adjustments take place. In such cases, low flows have little or no influence on the overall channel form. Many empirical relationships have been developed to relate 'regime dimensions' (e.g. channel width or depth), to control variables (e.g. bankfull discharge). It is important to realise that these regime dimensions represent an average and are not applicable to all channel types or flow regimes. The available stream power along a given reach is determined by the discharge and valley slope. At the sub-reach scale there are spatial variations in energy expenditure, which result from variations in channel shape and resistance to flow. These in turn influence patterns of erosion and deposition. For example, energy and erosion potential are concentrated where the channel narrows. Conversely, flow resistance is increased by obstructions to flow such as boulders, bedforms, bars or woody debris, which can lead to localised deposition. There is therefore two-way feedback between channel form and flow hydraulics – form influences flow and flow influences form. This point is well illustrated by the work of Ashworth and Ferguson (1986) on a glacially fed braided river in Arctic Norway. An intensive monitoring programme was carried out to make detailed measurements of channel morphology, velocity and shear stress, bedload size and transport rate, and the size of bed material. Starting at the top left of this diagram is the discharge of the river, which is unsteady (varies over time). The irregular form of the channel creates non-uniform flow conditions over the rough channel bed. As a result, complex spatial variations are seen in velocity, which also changes over time. Rates of bedload transport are determined by bed shear stress as well as the size and amount of bed material that is available for transport. As with velocity and shear stress distributions, rates of bedload transport are spatially variable, and also change with time. Bedload transport may maintain the existing channel shape, size and pattern. Alternatively, channel form can be modified as a result of scour, fill and possible lateral migration. The nature of such changes is spatially variable, and in turn feeds back to influence the velocity distribution within the channel. The character of the bed material determines the roughness of the channel, in turn affecting the velocity distribution in the channel.

Driving variables

Flow regime

The flow in natural river channels is unsteady, fluctuating through time in response to inputs of precipitation to the drainage basin. Characteristics of the flow regime include seasonal variations, flood frequency–magnitude relationships and the frequency and duration of low flows. Since discharge influences stream power, velocity and bed shear stress, the characteristics of the flow regime have an important influence on channel form. Of morphological significance is the bankfull discharge. The bankfull discharge marks a morphological discontinuity between within-bank and out-of-bank flows. Since the flow in natural channels is unsteady, the bankfull discharge provides a representative flow. The geomorphological effectiveness of a given flood depends not only on its size, but also on the frequency with which it occurs. Large floods can carry out a considerable amount of geomorphological work. However, their comparative rarity means that the cumulative effect of smaller, more frequent flows may be more significant in shaping the channel.

Sediment regime

The supply of sediment varies through time. It is not only the volume of sediment that is important but also its size distribution. Fluctuations in the volume and size of sediment are brought about by variations in sediment supply from the drainage basin and processes of sediment transfer through the channel network. As with the flow regime, it is the processes in the drainage basin, upstream from a given reach, that influences sediment supply.

The balance between stream power and sediment supply

There is an important balance between the supply of bedload at the upstream end of a channel reach and the stream power available to transport it. This is known as the Lane balance, having first been described as a qualitative equation by Lane in 1955. The left hand side of the scales represents the volume and size of sediment supplied to a channel reach over a given period of time. Balanced against this is the stream power available to transport it. This is determined both by the volume of water that enters the reach (over the same time period), and by the slope over which it flows. If the stream power is exactly sufficient to transport the sediment load, both sides of the scales are in balance and there is no net erosion or deposition along the reach. This is not to say that there is no erosion or deposition whatsoever, because these processes do occur at a localised scale in response to local variations in hydraulic conditions. Rather it means that, on balance, neither erosion nor deposition will predominate. An imbalance will occur if there is an increase in the volume or calibre of the sediment load in relation to the available stream power (sediment calibre is important because it determines the flow competence required to transport it). This means that there is insufficient stream power to transport all the sediment, with the result that the excess is deposited along the reach. In this case, the balance tips towards **aggradation**, with net deposition occurring along the reach. Aggradation can be triggered in several ways, for example where the sediment supply is increased by upstream channel erosion, mass movement, or human activities such as mining. Aggrading channels are characterised by numerous channel bars in a wide, shallow channel. Deposition within the channel may lead to the channel bed becoming elevated above the surface of the floodplain. This, together with reduced channel capacity, increases the incidence of flooding and also promotes channel migration. A different situation arises when the stream power exceeds what is needed to transport the sediment load through the reach. This excess energy has to be expended somehow, so it is used to entrain sediment from the bed and erode the channel boundary. In this case **degradation** predominates. Degradation can be caused by an increase in discharge, perhaps caused by an increase in flood frequency, or by a decrease in sediment supply. The Lane balance is simplistic because much depends on the calibre of bed sediment within the reach. For example, no degradation can occur in a boulder-bed stream if the bed sediment is too coarse to be moved by the available stream power. This can be true even if the stream power exceeds the sediment supplied at the upstream end of the reach. Even when degradation does occur, another limitation of the equation is that it does not tell us *where* within the reach erosion will occur (Simon and Castro, 2003). This means that the equation cannot be used to predict the actual nature of channel change. For example, if the channel bed is more resistant to erosion than the banks, bank erosion is likely to be an initial adjustment. However, in a sand-bed channel with cohesive banks it is more likely that an initial adjustment would be scouring of the bed (Simon and Castro, 2003). Resistance to erosion can be highly variable within a given reach, as can the specific stream power along that reach. This gives rise to spatially complex adjustments along the reach, even if there is net aggradation or net degradation along the reach as a whole.

Boundary conditions

Valley slope

This refers to the downstream slope of the valley floor (as opposed to the slope of the channel itself) and determines the overall rate at which potential energy is expended along a given reach. The valley slope imposed on a given reach of channel is determined by a combination of factors including tectonics, geology, the location of the reach within the drainage basin and the long-term history of erosion and sedimentation along the valley. Although the overall energy available along a given reach is largely determined by the valley slope, it is possible for adjustments to occur that increase flow resistance at different scales (channel resistance, form resistance and boundary resistance). Different types of channel and floodplain morphology are associated with low, medium and high-energy environments.

Valley confinement

A channel may be defined as confined, partly confined, or unconfined, depending on how close the valley sides are. The degree of valley confinement is important for several reasons. In **confined** settings channel adjustments are restricted by the valley walls, which also increase flow resistance. In addition, valley width influences the degree of slope–channel coupling that exists. Inputs of sediment from mass movements and other slope processes may exceed transport capacity, in turn influencing channel form. The episodic nature of mass movements means that these contributions can vary considerably over time. In **partly confined** settings, some degree of lateral migration and floodplain development is possible. However, where the river comes against the valley wall or hillslope it is prevented from migrating further, which can lead to the development of over-deepened sections of channel. Stream power is also concentrated within the narrow valley and sections of the floodplain surface may be stripped during major floods. Where the hillslopes are a long way from the channel and have relatively little influence in contributing to the channel load, the channel is described as

unconfined. Typically these settings are found in the lower reaches of rivers where there is very little interaction between channel and hillslopes.

Channel substrate

Considerable variations are seen in the form and behaviour of channels developed in different substrates. The substrate determines how resistant the channel is to the erosive force of the flow. It also influences boundary roughness, and therefore flow resistance. Alluvial channels formed in sand and gravel are generally more easily adjusted than those with cohesive silt and clay substrates. This is because the individual particles can be entrained at relatively low velocities, so non-cohesive substrates tend to be associated with wider, shallower cross-sections and faster rates of channel migration. Bedrock and mixed bedrock-alluvial channels are influenced over a range of scales by various geological controls.

Riparian vegetation

Vegetation on the banks and bed of river channels controls channel form in various ways. It often acts to protect and strengthen the banks, and research has shown that a dense network of roots can increase erosion resistance by more than a factor of ten. As a result, channels with vegetated banks are often narrower than those with non-vegetated banks under similar formative flows. This effect is most marked for densely vegetated banks (Hey and Thorne, 1986). Flow resistance can also be increased by vegetation growing on the bed and banks, as well as by woody debris (fallen trees and branches) that enters the channel from the banks. An interesting example of the influence of riparian vegetation on channel form is provided by the Slesse Creek, British Columbia, Canada, and is reported by Millar (2000) and MacVicar (1999). The Slesse Creek drains an area of 170 km2 within the Fraser River basin, flowing southwards from the United States into British Columbia.

Downstream changes

Downstream changes in slope, discharge, valley confinement, sediment supply and particle size give rise to different balances between erosion and deposition along different parts of the profile. This leads to downstream changes in channel and floodplain morphology. In general terms, the cumulative supply of sediment increases downstream but the available energy decreases. The discharge in most river channels increases in a downstream direction, as a progressively larger area is drained. In order to accommodate the growing volume of flow, channel dimensions (width and depth) typically increase downstream, and are often accompanied by a slight rise in velocity. The way in which these parameters change with increasing discharge can be described by the hydraulic geometry of the channel. Observations show that there is a general decline in sediment size along the channel. The main causes of this downstream reduction are widely recognised as being abrasion and selective transport (Rice and Church, 1998). Abrasion refers to the reduction in size of individual particles by chipping, grinding and splitting. Physical and chemical weathering processes are also significant in the pre-weakening of individual particles. Selective transport refers to the longer travel distances associated with smaller grains, which are more mobile. The rate of reduction in sediment size varies considerably and downstream *increases* are often observed at several locations. The downstream decrease in sediment size is often disrupted by inputs of coarser material. These include material from bank erosion, inputs from tributaries, and colluvial material. Material entering the main stream from tributaries is typically coarser than that in the main channel (Knighton, 1998). This causes a sudden increase in sediment size followed by a progressive fining further downstream. Complex patterns of downstream size reduction are seen where slope channel coupling is strong and non-alluvial supplies are dominant. These include contributions from hillslopes (e.g. mass movements), the erosion of bedrock outcrops and glacial material (Rice and Church, 1998). The channel slope is typically steepest in the headwaters, becoming gentler in the lower reaches. The resulting long profile of many rivers is concave in shape, although the degree of concavity varies. Downstream increases in discharge, together with a decrease in bed material size, mean that the load can be transported over progressively shallow slopes. Exceptions to this are seen in arid and semi-arid regions, where downstream conveyance losses and high rates of evaporation lead to a downstream reduction in discharge. In this case a straight or convex profile may develop, since increasingly steep slopes are needed to compensate for the downstream reduction in flow. Irregularities are often seen in the long profile, for example flatter sections are associated with lakes and reservoirs, and steeper sections at the site of waterfalls. In addition, there is often a change in the channel slope where tributaries join the main channel, because of the sudden increase in discharge. In tectonically active areas, where rates of uplift may be similar to erosion rates, rivers are in a state of dynamic equilibrium constantly trying to 'catch up' with tectonically driven changes. It takes time for a concave profile to develop, so the *overall* shape of long profiles in tectonically active regions tends to be straight rather than convex.

CHANNEL ADJUSTMENT

Time scales of adjustment

Different components of a channel's morphology (e.g. bedforms, cross-sectional shape, slope) change over different time scales. This is because some components are more readily adjusted than others. For example, bedforms in a sand-bed channel are rapidly modified by a wide range of flows. Adjustments to channel width and depth take place over months to years, planform adjustments occur over tens to hundreds of years, while changes in the long profile may take thousands of years. Morphological adjustments therefore tend to lag behind the changes that cause them. This means that it can be difficult to link processes of flow and sediment transport with channel dimensions and form. Channel form is directly controlled by flow regime and sediment supply. These all act as controls on the flow and sediment regimes and, through a complex sequence of adjustments, lead to long-term changes in channel form and behaviour.

How adjustments are made

Channel form and behaviour reflect the driving variables and boundary conditions influencing a given channel reach. These controls also influence the ways in which channel adjustments are made. There are potentially four **degrees of freedom** or variables that can be modified: channel cross-section, slope, planform and bed roughness. Modifications to the cross-sectional size and shape are associated with changes in width and depth of the channel by processes such as bank erosion, incision of the bed, or aggradation. Channel slope can be adjusted in different ways. Negative feedback reduces the slope of steeper sections by erosion, and the slope of flatter sections is increased by deposition. Increases or decreases in channel length also affect channel slope. There are several different types of channel planform adjustments. These include lateral migration, meander bend development, reworking of bars, and even wholesale shifts of the channel to a new course. Finally, changes in bed roughness are brought about when the channel rearranges bed material, for example, in sand-bed channels, where bedforms are modified in response to changes in flow conditions. Mutual interrelationships exist between these variables, with adjustments made to one affecting one or more of the others. For instance, the formation of a meander cut-off alters the channel planform as well as increasing channel slope. The influence of the driving variables and boundary conditions often reduces the degrees of freedom that a particular channel has. In the case of the mixed bedrock-alluvial channel, depth increases are greatly restricted by the rock bed of the channel. On the other hand, the alluvial banks allow the channel to be widened much more easily. However, reductions in cross-sectional size by deposition may be limited if the channel has degradational tendencies (Charlton, 2008).

UNIT-6: COASTAL MORPHODYNAMIC VARIABLES AND THEIR INFLUENCE IN EVOLUTION OF LANDFORMS

The Work of Waves and Tides

This chapter brings together two different agents of erosion: wind and waves. Wind is the movement of air, a fluid of low density; waves involve the movement of water, a fluid of much higher density. This difference in fluid density means that the power of moving air to create landforms is generally much weaker than the power of moving water as waves. That said, vast desert landforms on our planet have been shaped by wind, whereas waves can act only on coastlines of oceans and large lakes. Yet the two agents are related because waves are set in motion by wind blowing across a water surface. We will first examine how waves create landforms.

WAVES

Waves are the most influential of the agents that shape coastal landforms. When winds blow over broad expanses of water, they generate waves. Both friction between moving air and the water surface and direct wind pressure on the waves transfer energy from the atmosphere to the water. Waves are seen and felt as a regular rising and falling of the water surface that causes a floating object to move up and down, forward and back. Waves have *crests* and *troughs* . *Wave height* is the vertical distance between trough and crest. *Wave length* is the horizontal distance from trough to trough or from crest to crest. The wave typically travels in a forward direction as parallel fronts of crests and troughs. The *wave period* is the time in seconds between successive crests or successive troughs that pass a fixed point. Wind-generated ocean waves are an example of *oscillatory waves* . In this type of wave, a tiny particle, such as a drop of water or a small floating object, completes one vertical circle, or *wave orbit* , with the passage of each wave length. The circle of motion grows rapidly smaller with depth. For lake and ocean waves pushed by the wind, particles trace an orbit in which the forward speed of each particle at the crest is slightly greater than the backward speed at the trough. As a result, there is net forward motion of water associated with most lake and ocean wave trains. The amount of motion depends on the size and steepness of the waves, and strong waves are capable of pushing large amounts of water toward the beach. This causes local sea level to rise and generates local currents of water heading back out to sea. The height of waves is determined by wind speed, wind duration, and *fetch* —the distance that the wind blows over the water. Given a sustained wind, average and peak wave heights increase with fetch and duration, until waves are fully developed. Wave height increases rapidly with wind speed. Waves retain most of their energy as they travel across the deep ocean, but when a wave reaches the shore, it begins to expend its energy. As it reaches shallow water, the drag of the bottom slows and steepens the wave. At the same time, the wave top maintains its forward velocity and eventually falls down onto the face of the wave, creating a *breaker.* Tons of foamy turbulent water surge forward in a sheet, riding up the beach slope. This powerful *swash* moves sand and gravel on the beach landward. Once the force of the swash has been spent against the slope of the beach, the return flow, or *backwash* , pours down the beach. This undercurrent can sweep unwary bathers off their feet and carry them seaward beneath the next oncoming breaker. The backwash carries sand and gravel seaward, completing the wave cycle.

LITTORAL DRIFT

Waves breaking along the shore generate tremendous energy to move sediment along the shoreline in a process called **littoral drift** . (*Littoral* means "pertaining to a coast or shore.") Littoral drift includes two transport processes. *Beach drift* is the transport of sediment along the beach. This occurs when waves reach the shore at an angle, however slight, and the swash from the breaking waves carries sand up the beach at a similar angle. The backwash, then, is drawn downward toward the breaker zone by gravity in a straighter path. The result is that sand is carried down the beach in a series of steps, setting in motion a net movement of sand away from the angle of wave attack. Just offshore, the net forward motion of breaking waves pushes water into the breaker zone. The water is propelled along the shoreline, away from the direction produces a longshore current, which moves sand along the bottom as *longshore drift* , the second component of littoral drift. If the wave direction remains more or less the same for long periods, littoral drift can transport sediment for long distances along the coast.

WAVE REFRACTION

The process by which waves erode sediment along a shore depends on two factors: the amount of energy the waves have and the resistance of the shore materials. Where waves strike a rocky coastline, areas of softer rock are eroded more quickly, carving out bays and coves and leaving behind jutting landforms of resistant rock, called *headlands* . When the coastline has prominent headlands that project seaward, and deep bays in between, approaching wave fronts slow when the water becomes shallow in front of the headlands. This slowing effect causes the wave front to wrap around the headland, in a process called *wave refraction*. This concentrates wave energy on the headland, enhancing erosion. The angle of the waves against the side of the headland also initiates a longshore current that moves sediment from the headland into the surrounding bays, creating crescent-shaped *pocket beaches* in the bays. Over time, the action of waves against a coast tends to have the effect of straightening the coast as it erodes sediment from headlands and deposits it into bays.

TIDES

Most marine coastlines are also influenced by the ocean tide, the rhythmic rise and fall of sea level under the influence of changing attractive forces of the Moon and Sun on the rotating Earth. In the tidal system, the Earth and Moon are coupled together by their mutual gravitational attraction, which is balanced by an inertial force generated by their revolution around a common center of mass. While the inertial force is constant at all points on the globe, the gravitational attraction of the Moon is greater on the near side of the Earth; seawater responds to this attraction, causing the ocean to "bulge" toward the Moon. On the far side of the globe, the gravitational force of the Moon is weaker; but because the inertial force is the same, the ocean water is pushed away from it, and this causes a second bulge to emerge on the far side. The bulges remain essentially stationary while Earth rotates through them, creating two high tides and two low tides per day. The Sun affects tides in a similar way, but its tide-producing force is only about half as strong as the Moon's. When the Sun and Moon are aligned on opposite sides of the Earth, their gravitational forces combine to produce tides with a higher range, called *spring tides* . In contrast, when the Moon and Sun are positioned at a right angle to the Earth, tides have a lower range and are called *neap tides* . Some marine organisms time their egg laying on the beach to coincide with high spring tides, so that the eggs can incubate without subsequent exposure to wave action. Hatchlings emerge at the next spring tide, about 14 days later, and are washed out to sea. The survival of these species is dependent on close synchronization with these extreme tides. Each day, coastal regions experience two high tides, called *flood tides* , and two low tides, called *ebb tides* ; collectively, they are referred to as *semidiurnal tides* , meaning occurring approximately every half day. There are, however, two exceptions, due to unique orbital geometry: the Gulf of Mexico and South China Sea have only one high tide and one low tide each day, called *diurnal tides* . The difference between the heights of successive high and low waters is known as the *tidal range.* If the tidal range is wide, the changing water level can play an important role in shaping coastal landforms. Tides contribute to the erosion, transport, and deposition of sediment by ocean waters. As tides ebb and flood, ocean waters flow into and out of bays. Tidal currents pouring through narrow inlets that connect bays with the ocean are very swift and pick up sediment as they travel, eroding the inlets. This keeps the inlet open, despite littoral drift that would otherwise block the inlet with sand. Tidal currents also carry large amounts of fine silt and clay in suspension. Some of this sediment is deposited in bays.

TSUNAMIS

Although most waves are generated by wind, waves can also be set in motion by any sudden displacement of large amounts of water, such as by an earthquake or an undersea landslide. These waves are called **tsunamis** , or seismic sea waves. (Seismic sea waves are also sometimes called tidal waves, but they are unrelated to tides.) A tsunami originates when the sudden movement of the sea floor—resulting from an earthquake, for example—generates a train of water waves. These waves travel over the ocean at 700 to 800 km/hr (435 to 500 mi/hr), moving outward in all directions from their source. In deep ocean waters, the tsunami 's motion is normally too gentle to be noticed, making it hard to detect in the open ocean. But as the wave approaches land, it is subject to the same physics as wind-driven waves; that is, as the wave drags along the bottom, and slows, it shortens and steepens to a height of 15 m (50 ft) or higher. Ocean waters rush landward and surge far inland, destroying coastal structures and killing inhabitants as they pass, at speeds of up to 15 m/s (34 mi/hr) for several minutes. Considering that each cubic meter of water weighs about 1000 kilograms (about 2200 lbs), it is easy to understand the tremendous power of such a surge of water. Even as the waters retreat, they continue to devastate the land, pulling people and debris back out to sea with them. Tsunamis typically consist of several surging waves, one after the other—and the largest wave is not necessarily the first to arrive. With their tremendous force, tsunamis can dramatically reshape coastal areas and inflict damage in regions on both sides of the ocean. The deadliest tsunami recorded so far struck the Indian Ocean region in December 2004, following a massive undersea earthquake—measuring 9.0 on the Richter scale—in the Java Trench, west of Sumatra. About 1000 km (600 mi) of fault ruptured, rocketing the nearby sea floor upward about 5 m (16 ft). This rapid explosion of a vast area of ocean bottom launched a giant tsunami that devastated populations in Southeast Asia, India, and Africa, reshaping coastal landscapes, wiping out cities, and killing more than 265,000 people. Where early-warning systems have been installed, tsunamis can usually be predicted far enough in advance to evacuate local populations. By monitoring seismic activity on the sea floor, it is possible to trigger tsunami warnings along coastlines that might be at risk. However, if an earthquake occurs very near the shore, or if local warning systems are not effective, the toll taken by tsunamis can be incalculable. One such event took place on March 11, 2011, in Japan. On that date, the Pacific Tsunami Warning Center in Hawaii detected an undersea earthquake of magnitude 9.0, located 70 km (43 mi) off the east coast of Tohoku, the north eastern portion of Honshu, Japan 's largest island. The Great Tohoku earthquake, the largest ever recorded in that country, occurred along a 300-km (186-mi) subduction zone at a depth of 32 km (20 mi) under the ocean floor, with its epicenter 70 km (43 mi) east of the coastline. And in spite of Japan 's technologically advanced warning systems and regularly scheduled emergency training program, nothing could stave off entirely the onslaught of destruction set in motion by the earthquake and the huge tsunami it triggered— these early-warning systems did save tens of thousands of lives minutes before the 9.2 m (30 ft) wave came ashore. Unfortunately, worse was yet to come when nuclear reactors built along the coastline were flooded, and eventually experienced partial core meltdowns that released harmful radioactivity. The final death toll of the Great Tohoku earthquake and tsunami exceeded 15,000.

Coastal Landforms

As waves break upon the shore, they expend tremendous amounts of energy. This energy, along with the coastal currents it produces, shapes coastlines by eroding steep cliffs, carving out bays, and building beaches. Two terms are important to distinguish in describing coastal landforms and processes: **Shoreline** refers to the dynamic zone of contact between water and land. **Coastline** (or **coast**) refers to the zone of shallow water and nearby land that fringes the shoreline; it is the zone in which coastal processes shape landforms. The Earth's coastlines are widely varied. For example, along most of the East Coast of the United States, we fi nd a coastal plain gently sloping toward the sea, with shallow lagoons and barrier islands at the coastline. On the West Coast, in contrast, we often find coastlines of rocky shores, with dramatic sea cliffs, headlands, and pocket beaches. To explain these differences we look to plate tectonics. The eastern coastline is a passive continental margin, without tectonic activity. Abundant sediment from rivers on the continent, accumulating over millions of years, provided the material to build the coastal plain, beaches, and islands. Here, the characteristic coastal landforms are depositional—built by sediment moved by waves and currents. In contrast, the western coastline is the site of great tectonic activity, with rocks rising from the sea along subduction boundaries and transform faults. Here, the characteristic landforms are erosional, with waves and weathering slowly eroding the strong rock exposed at the shoreline. Coastlines around the world are likewise varied, but for other reasons. During each glaciation of the recent Ice Age, sea level fell by as much as 125 m (410 ft). Rivers cut canyons to depths well below present sea level to reach the ocean. Waves broke against coastlines that are now well underwater. Now that sea level has risen, many coastal landscapes are coastlines of submergence, where fluvial landforms are now under attack by wave action. Still others are dominated by coral reefs, where ocean waters are warm enough for coral to grow.

EROSIONAL COASTAL LANDFORMS

The breaking of waves against a shoreline yields a variety of distinctive features. If the coastline is made up of weak or soft materials—various kinds of regolith, such as alluvium—the force of the forward-moving water alone easily cuts into the coastline. Here, erosion is rapid, and the shoreline may recede rapidly. Under these conditions, a steep bank, or *marine scarp*, will form and steadily erode as it is attacked by storm waves. Where resistant rocks meet the waves, **sea cliffs** often occur. At the base of a sea cliff is a *notch* , carved largely by physical weathering. Constant splashing by waves followed by evaporation causes salt crystals to grow in tiny crevices and fissures of the rock, breaking it apart, grain by grain. Hydraulic pressure of waves, and abrasion by rock fragments thrust against the cliff, also chisel the notch. Undercut by the notch, blocks fall from the cliff face into the surf zone. As the cliff erodes, the shoreline gradually retreats shoreward. Sea cliff erosion results in a variety of erosional landforms, including *sea caves*, *sea arches*, and *sea stacks.* The retreat of the cliffs leaves behind a broad, gently sloping plane, called a *shore platform*, at the base of the cliffs. If tectonic activity elevates the coastline, or if sea level falls, the platform is abruptly lifted above the level of wave attack. What was a shore platform is raised up and becomes a **marine terrace**. Repeated uplifts produce a series of marine terraces in a step like arrangement. Marine terraces are common along the continental and island coasts of the Pacific Ocean, where tectonic processes are active along the mountain and island arcs.

DEPOSITIONAL COASTAL LANDFORMS

Most of the sediment we find along a coastline is provided by rivers that reach the ocean. Waves then transport and deposit this sediment to take shape as shoreline features such as beaches, bars, and spits. These depositional landforms are relatively transitory, however, appearing, disappearing, or migrating as a result of seasonal changes, storms, and human engineering. **Beaches** A **beach** is a wedge-shaped sedimentary deposit, usually of sand, built and worked by wave action. The face of a beach varies over time as waves either deposit or erode more sand. During short periods of storm activity, waves cut back the beach, giving it a long, flat, sloping profile. The sand moves just offshore and along the shore via longshore drift. Gentler waves return the sand to the beach, building a steeper beach face and a bench of sand at the top of it. When sand leaves a section of beach more rapidly than it is brought in, the beach is narrowed and the shoreline moves landward. Conversely, when sand arrives at a particular section of the beach more rapidly than it is carried away, the beach is widened and builds oceanward. During the winter, many midlatitude beaches experience more active waves and are eroded to narrow strips. This change is called *retrogradation* (cutting back). In summer, the gentler wave climate allows sediment to accumulate, and so the beach is replenished. This change is called *progradation* (building out). Retrogradation and progradation can also happen on longer time cycles, related to changes in sediment input, climate, or human activity in the coastal zone. Although most beaches are made from particles of fi ne to coarse quartz sand, others are built from rounded pebbles or cobbles. Still others are composed of fragments of volcanic rock, or even shells. **Coastal Dunes** Where ample sand is available, a narrow belt of dunes, called *foredunes* , often occurs in the region landward of beaches. These dunes are usually held in place by a cover of beach grass. Although these dunes are typically built up by wind, they play an important role in maintaining a stable coastline by trapping sand blown landward from the adjacent beach. As sand from the beach collects along the foredunes, the dune ridge grows upward, becoming a barrier several meters above high-tide level. This forms a protective barrier for tidal lands on the landward side of a beach ridge or barrier island. In a severe storm, the swash of storm waves chisels away the upper part of the beach. Although the foredune barrier may then be eroded by wave action and partly cut away, it will not usually give way. Between storms, the beach is rebuilt, and in due time, if a vegetative cover is maintained, wind action restores the dune ridge. If, however, the plant cover of the dune ridge is undermined by human foot and vehicular traffic, inlets may open up and allow ocean water to wash in during storms or high tides. Many coastal communities now protect their dunes by erecting raised pathways over them, to restrict foot traffic, or by planting protect **Spits And Bars** Where littoral drift moves the sand along the beach toward a bay, the sand is carried out into the open water, extending like a long finger, called a *spit* . As the spit grows, it forms a barrier, called a *baymouth bar* , across the mouth of the bay. Once the bay is isolated from the ocean, it is transformed into a lagoon. And where a spit grows to connect the mainland to a near-shore island, it forms a *tombolo*. **Barrier Islands** Much of the length of the Atlantic and Gulf coasts of North America is flanked by *barrier islands* , low ridges.

COASTLINES OF SUBMERGENCE

Coastlines of *submergence* are formed when rising sea level partially drowns a coast or when part of the coast sinks. At the end of the last glaciation, sea level rose by about 120 m (394 ft), submerging coastal landscapes. Climate scientists expect sea level to rise even higher over the next century as a result of global climate change. Today, many of the major river delta areas are experiencing submergence resulting from land use practices and sea-level rise. A *ria coast* is formed when a rise of sea level or a crustal sinking (or both) brings the shoreline to rest against the sides of river valleys previously carved by streams. The shoreline rises up the sides of the stream-carved valleys, creating narrow bays. Streams that fl owed through the valleys add freshwater to the bays, making them estuaries of mixed fresh- and saltwaters, thereby producing a unique habitat for many plants and animals. A *fiord coast* is similar to a ria coast except that the bays are formed in valleys that have been scoured by glaciers. Glaciers eroded their walls and scraped away loose sediment and rock, leaving behind broad, steep-sided valleys that are now flooded with seawater. Fiords, the Scandinavian word for bays, are common features along the northern coastal countries; they can be found where glaciers occupied coastlines during the Ice Age. These deep, glacially scoured bays can be hundreds of meters deeper than the adjacent seas, and they often meander tens of kilometers inland.

CORAL REEFS

So far we have been talking about coastal landforms built by sediment. **Coral reefs** are different—unique, in fact—because their origins are biological; they are made by living organisms. Growing together, corals and algae secrete rocklike deposits of carbonate minerals. As old coral colonies die, new ones grow upon them, accumulating as composite layers of limestone. When coral fragments are torn free by wave attack, the pulverized fragments accumulate on land as coral sand beaches. For dense coral reefs to grow, water temperatures must be above 20°C (68°F); that is why coral reef coasts are usually found in warm tropical and equatorial waters between 30° N and 25° S. Furthermore, the seawater must be free of suspended sediment, and be well aerated, for vigorous coral growth to take place. For this reason, corals live near the water surface and thrive in locations that are exposed to waves from the open sea. They are not found near the mouths of muddy streams because muddy water prevents coral growth. Sediment runoff from poor land-use practices and development has been identified as a primary factor in most coastal coral reef die-offs. There are three distinctive types of coral reefs—fringing reefs, barrier reefs, and atolls. These reefs are often found around a hotspot volcano at different stages of submergence. *Fringing reefs* build up as platforms attached to shore. They are widest in front of headlands where the wave attack is strongest. *Barrier reefs* lie out from shore and are separated from the mainland by a lagoon. At intervals along barrier reefs, there are narrow gaps through which excess water from breaking waves is returned from the lagoon to the open sea. *Atolls* are more-or-less circular coral reefs enclosing a lagoon; no land is inside. Most atolls are rings of coral growing on top of old, sunken hotspot volcanoes. They begin as fringing reefs surrounding a volcanic island. Then, as the volcano sinks, the reef continues to grow, until eventually only the reef remains. Coral reefs are highly productive ecosystems that support a diversity of marine life-forms. They also perform an important role in recycling nutrients in shallow coastal environments. They serve, too, as physical barriers that dissipate the force of waves, protecting ports, lagoons, and beaches behind them. Finally, they are an important aesthetic and economic resource. Unfortunately, because of their very specific environmental requirements, coral reefs are highly susceptible to damage from human activities, as well as from natural causes such as tropical storms. Some 50 percent of coral reefs worldwide are currently under threat of human activities. Earlier in the chapter, we noted the effects of global warming on corals, causing bleaching and, ultimately, the death of corals from elevated water temperatures. But geographers are also worried about the impact on corals of rising carbon dioxide levels in the atmosphere. As more CO 2 dissolves in seawater it becomes more acidic, making it harder for corals to develop their calcium carbonate skeleton structure (Strahler, 2013).

UNIT-7: PERIGLACIAL PROCESSES AND LANDFORMS

The periglacial environment

Introduction

The Polish geomorphologist W. von Lozinski introduced the term 'periglacial' in 1909 to describe climatic conditions and geomorphic features of areas peripheral to Pleistocene icesheets in the Carpathians. The term was later extended to include processes and landforms characteristic of cold climates regardless of their proximity to glaciers or glaciated areas. Periglacial activity predominates in vast zones that are unrelated to glacial areas such as eastern Siberia (Jahn, 1975). The term periglacial has been used broadly because of its imprecise definition. The word 'geocryology,' which is commonly used by Russian researchers, has been suggested as an alternative term (Washburn, 1979). Geocryology is commonly defined as the study of terrestrial materials that have temperatures below 0°C, in other words permanently or seasonally frozen ground (Tricart, 1967; Washburn, 1979; French, 1996, 2007). Periglacial environments are characterized by freezethaw cycles and permafrost or permanently frozen ground. Either or both occur within the periglacial environment although some periglacial landforms are not associated with the presence of permafrost. The periglacial environment is present in polar areas and in middle and low latitude alpine areas of cordilleras throughout the world (Harris, 1986). Today the periglacial domain covers about one-fifth of the global land surface; periglacial conditions covered an additional 20 percent of the Earth's surface during the Pleistocene cold periods (French and Karte, 1988). Paleoclimatologists recognize the importance of periglacial phenomena in paleoclimatic reconstructions and of permafrost (Isarin, 1997). Ecologists and environmental scientists have pointed out the vulnerability of present-day periglacial regions as a consequence of increased human use and the implications of expected climate change (Vandenberghe and Thorn, 2002). Periglacial environments have a wide range of climatic conditions with mean annual temperature close to or far below the freezing point and the range between maximum and minimum yearly temperatures is generally large. Total annual precipitation varies considerably from one environment to another with values between 130 and 1400 mm according to Peltier (1950) and between 50 and 1250 mm according to Wilson (1969). Tricart (1967) established a climatic classification based on temperature, precipitation, wind conditions, and their seasonal distributions. He distinguished three main types of periglacial climates: A) **Dry climates with cold winters.** These conditions occur in subpolar areas of the northern hemisphere. These areas have low-temperature winters, short summers, low precipitation, and strong winds. This climate favors the presence of pergelisol (permafrost). The morphogenetic system of this climate type is characterized by strong frost action, limited influence of runoff, and significant wind activity. B) **Cold wet climates with extreme winters.** This group includes arctic and mountain types. The **arctic type** is controlled by the ocean and thus has wide climate variations. Temperatures are similar to climate A but with a lower annual range. Precipitation is higher than 300 mm, violent winds are common, and permafrost is present. As a result, frost action is less intense or persistent than in climate A, snow cover reduces eolian processes, and runoff is relatively important. The **mountain type** occurs in alpine meadow areas in temperate zones. Temperatures are similar to the arctic type but with higher annual averages and smaller ranges. Precipitation is more significant than in the arctic climate. Slope and orientation play important roles in these areas. Overall frost action is significant, permafrost is generally absent, wind influence is weak, and runoff-related processes are widespread. C) **Climates with small annual temperature ranges.** The mean temperature is close to 0°C and the annual temperature range is around 10°C. There are two types of climates with small annual temperature ranges. The climate of **high latitude islands** has significant weather instability, a small temperature range, and snow precipitation above 400 mm that minimizes the effects of wind. These climatic conditions lead to multiple freeze-thaw cycles that have limited penetration into the ground. **Low latitude mountain** climates lack seasonal temperature variations; the daily range is high and larger than the annual range. Precipitation is high except in arid mountains. These climatic conditions prevent permafrost formation and increase numerous freeze-thaw cycles and the weak penetration of frost action; there is insignificant wind activity except in arid mountains. The periglacial environment has two principal types of vegetation: the subarctic or northern forest and the arctic tundra. There is thus a distinction between forest periglacial environments and woodland-free environments (French, 1996).

Permafrost characteristics

Permafrost, also called pergelisol or permanently frozen ground, is defined as a soil or a surficial deposit or rock mass that remains below 0°C continuously for more than two years; the definition does not consider other properties such as moisture content and lithology (Muller, 1945). Glaciers and ice sheets are not considered to be permafrost. Permafrost covers approximately 22% of the land surface of the Northern Hemisphere and extensive areas in the Andes and the Antarctic (Harry, 1988). Globally, permafrost underlies 14% of the Earth's surface (Washburn, 1979). Half of Canada and 80% of Alaska are underlain by permafrost (Harry, 1988). Permafrost also occurs at the bottom of shallow seas in high latitudes (**submarine permafrost**). Two types of permafrost occur on continents: **polar permafrost** and **alpine permafrost.** Alpine permafrost occurs in mountain environments and is a function of latitude and elevation. The lower altitude limit of permafrost increases from north to south (Harris, 1988). **Continuous permafrost** wedges out toward low latitudes and breaks up into patches, giving way to **discontinuous** and **sporadic permafrost** (Harris, 1986). This classification is based on the percentage of underlying permafrost at a particular location. The proportion of frozen ground in continuous permafrost is higher than 80 percent, is between 30 and 80 percent in discontinuous permafrost, and is less than 30 percent in sporadic permafrost (Brown, 1970) is a north-south cross section of Canada that shows permafrost increment with increased latitude and increased fragmentation of blocks separated by unfrozen ground, called **talik**, with decreased latitude The maximum documented thickness of permafrost occurs in the River Markha in Siberia where it reaches a depth of 1450 m (Washburn, 1979). During summer the upper part of the permafrost thaws; this zone, which is subject to continuous cycles of freezing and thawing, is called the **active layer** or **mollisol** although some authors restrict this term to seasonally frozen ground, that is ground that remains below 0°C only in the winter. The thickness of the active layer varies from year to year. The **transition layer** is a near-surface layer that contains abundant ice, indicating a transition toward the upper zone of permafrost (Shur *et al*., 2005). Overall, the active layer thickens at lower latitudes but does not exceed 3–4 m in depth. Water that freezes within permafrost comes from the atmosphere, runoff, or underground flow. The frozen water cements mineral and organic particles together; ice may occur as individual crystals, particle coatings, or pore fillings. Ice that formed in pre-existing sedimentary sequences is called epigenetic; ice that forms during sedimentation is called syngenetic (Harry, 1988); an example of the latter is the advance of permafrost in deltaic and fluvial sequences. Ice bodies in permafrost have variable geometries (Washburn, 1979). Ice can be found in pores, holes, and cavities and in both unconsolidated sediments and rocks (French, 2007). Remains of old **glacial ice** that are found buried by sediments in permafrost can be recognized by their deformation. **Icings** or **aufeis** are tabular or planar-convex bodies of fresh-water ice that form in the winter where groundwater emerges as springs or where floodwaters freeze in river valleys. They may be incorporated into permafrost upon burial. **Ice lenses** are horizontal lenticular layers of clear ice of variable thickness; they form in situ as ice segregates during the freezing process. **Ice veins** and **wedges** are vertical cracks in the ground that fill with ice. Meteoric water supplies the veins, which form wedges as they grow. The V-shaped veins taper downward and the ice crystals have a vertical arrangement (Lachenbruch, 1962); as they expand, they deform the adjacent sediment. Ice wedges formed by epigenetic freezing are commonly 2–3 m wide and penetrate about 5–8 m. In Siberia, syngenetic ice wedges that form in alluvium can be 5 m wide and 40–50 m deep (Harry, 1988); Bray *et al*., 2006). They frequently form part of polygonal systems. If ice wedges disappear during warming, the existing void may fill with detritus. Fossil or relict wedges are indicators of the presence of permafrost during previous epochs (Eissmann, 1981; Murton, 2007). Finally, the ice may become the nucleus of a mound, called **pingo ice**. Pingos are massive fresh-water ice bodies with a flat convex morphology; they answer to particular hydrological and geothermic local conditions such as the freezing of meteoric water, water-rich sediments, or artesian water intruded through the permafrost. Permafrost depth is determined by the balance between increasing heat at depth due to the geothermal gradient (1°C every 30–60 m) and heat loss at the surface. The basal limit of permafrost is the depth where the temperature is 0°C as a result of the geothermal gradient (Brown, 1970). The active layer forms near the surface; its base is at a depth where the maximum annual temperature is 0°C. Temperature fluctuations that occur near the surface due to seasonal fluctuations attenuate at depth to zero annual thermal oscillation; this occurs between 6 and 16 m (French, 1996). Permafrost is subject to aggradation or increased extent of permafrost or degradation over time due to changes in soil temperature (Williams, and Smith, 1989). Subtle climatic variations may affect permafrost at temperatures near 0°C. In addition climatic changes that occur on the scale of decades or centuries may significantly modify permafrost area and thickness. Most permafrost is relict; it was formed in past Quaternary periods and is not in equilibrium with present climatic conditions. Abundant evidence corroborates this assertion but perhaps the most convincing is the discovery of tissues of woolly mammoths (*Mamuthus primigenius*) and other Pleistocene animals preserved in Siberian permafrost (in Washburn, 1979; French, 2007). The presence of tissue shows that permafrost already existed when the animals died; otherwise they would have decomposed.

Periglacial processes

Most morphogenetic processes operating in the periglacial domain are also active in other morphoclimatic zones. However, processes related to freeze-thaw cycles reach their maximum intensity in the periglacial environment. In addition, thawing on slopes promotes mass movements. These processes predominate in cold regions along with surface runoff and eolian activity in cold desert areas (King, 1976).

Frost action

Frost action is the term used for processes that result from freezing and thawing. In many situations, multiple processes linked to frost action act jointly but in some cases they may be associated with processes of a different origin. The study of freezing processes can be relatively complex because it requires knowledge of the magnitude and duration of temperatures below 0°C. Furthermore, frost action may affect materials that have different moisture contents and variable heat conductivities (Williams and Smith, 1989). The understanding of frost-related processes requires knowledge of freeze-thaw cycles at the ground surface throughout the year. Daily variations in these zones are due to insolation changes. In some parts of the Andes more than 300 annual freeze-thaw cycles have been recorded (Troll, 1944). An annual average of 80 freeze-thaw cycles has been recorded at the Ransol meteorological station at an altitude of 1640 m in the Pyrenees in Andorra (Raso and Garcia Loureiro, 1998). At the University Hostel (2510 m) in the Sierra Nevada in the Betic Cordillera of Spain, 117 have been recorded in a year (Gomez Ortiz and Salvador Franch, 1997). Multiple processes are associated with frost action; some of their most significant effects are analyzed below. Gelifraction was discussed in Chapter 5 on Weathering.

Frost heaving

Frost-induced expansion exerts pressures in all directions. The vertical and horizontal components of the resulting displacements are called **heaving** and **translation** (Eakin, 1916). Heaving is more significant than translation. By installing sticks and pins to different depths in the ground, researchers have made field measurements of the magnitude of heaving. Maximum heaving occurs in areas with high moisture content during the spring when the ground has a relatively higher proportion of water from ice melting. Sticks driven into the ground may show different degrees of heaving or may even be expelled. Heaving values of 1–5 cm/year are common although much larger values have been measured (French, 1996). Heaving increases with depth to about 30 cm. It affects mainly the upper part of the active layer; the lower part is passive with respect to frost action. Studies done in the Sierra Nevada of Spain at an altitude of 3200 m indicate that the lower boundary of the horizon affected by frost action is at a depth of 25–30 cm (Gomez Ortiz and Salvador Franch, 1997). The heaving of rounded clasts around finer particles is common in periglacial areas; experiments indicate that larger clasts move upward more rapidly than small ones (Corte, 1969; 1971). Planar particles tend to rotate and rest on edge with the larger faces perpendicular to the cooling surface (Schunke, 1974), as shown in laboratory studies (Kaplar, 1965). Clast lifting causes vertical sorting of soil particles; this is a very common characteristic of periglacial accumulations. Researchers have explained frost heaving in two ways. One hypothesis proposes that both stones and fine particles rise as the ground expands. During thawing the fine grained material settles but ice remains beneath the clasts. Later the ice lens melts and the hole resulting from the melting fills with fine particles (*frost-pull hypothesis*) (French, 2007). This mechanism seems to be the most common. The second hypothesis applies to clasts located near the surface. The water that flows around the clasts collects below them; as it freezes it pushes the clasts upward. The clasts have a greater thermal conductivity than surrounding fine particles so the base of the clast reaches the freezing point first. During thawing, the clast does not return to its initial position because the frost translation component compresses the hole originally occupied by the clast; fine material slides toward the position of the basal ice (*frost-push hypothesis*) (Bowley and Burghardt, 1971). Another frost heaving mechanism is related to **needle ice** or **pipkrake**, which are ice crystals that form near the surface and perpendicular to it. These crystals, which range in size from a few millimeters to 40 cm, are common in alpine environments. As they grow, they lift clasts that have collapsed during the thaw and in so doing sort the soil (French, 2007). Some features of patterned ground in the French Alps are thought to be related to thawing and collapse of needle ice (Pissart, 1977). **Nubbins** are oval or elongate ground bulges that measure centimeters in size. They are thought to result from needle ice. **Voids around clasts** result from frost heaving and later settling during thawing (Washburn, 1979).

Mass displacement

Mass displacement is movement of unconsolidated material as a result of frost action. Particle movement has a predominantly vertical component although it may also be affected by horizontal displacement. Scientists have suggested numerous mechanisms to explain these internal movements but **cryostatic pressure** seems to be the main cause (Washburn, 1956). Cryostatic pressure increases as freezing-induced pressures propagate to unfrozen parts of the ground located between the freezing front and the top of the permafrost. The ground surface starts to freeze in autumn and the freezing front advances downward during the winter. Because different parts of the ground have variable moisture contents, the ground freezes irregularly, causing differential volumetric expansion. The pressures transmitted to the unfrozen ground may liquefy it or generate dome-shaped bulges at the ground surface. Laboratory experiments have demonstrated the high magnitude of cryostatic pressure (Corte, 1969; Pissart, 1970). Researchers believe that cryostatic pressures are the primary genetic cause of **cryoturbation** or **periglacial involution** (Sharp, 1942b; Vanderberghe, 2007) although other causes have also been suggested. These chaotic structures are reflected in cross section as disharmonic folds, diapiric injections, and faults that form during freezing and thawing of the soil in seasonally frozen zones. Involutions are also observed in other morphoclimatic zones, primarily arid, but the mechanisms that cause differential soil expansion are obviously different in the periglacial environment (French, 1996).

Ice cracking

As the temperature of frozen soils decreases, thermal contraction causes cracking. The soils behave as rigid solids; fissures and cracks form in the ice as the temperature falls. The cracks commonly occur as systems of polygons with four, five, or six sides (Lachenbruch, 1962). Researchers have classified frost cracks into two main types. Cracks form in winter; in the spring, they fill with snow, percolating water, groundwater, and water vapor. The upper part melts during the seasonal thaw whereas cracks in permanently frozen snow are preserved. The cracking is repeated every year; as cracks fill and freeze, they form an **ice wedge** (Lachenbruch, 1962; Yershov, 1998). Ice wedges have vertical banding; the number of bands shows the age of the wedge. They are generally thousands of years old (Yershov, 1998). Ice wedges are typical of humid environments; they require specific temperature conditions for preservation and growth. In Alaska, the southern boundary of active ice wedges corresponds to an average air temperature of 6°C to −8°C (Pewe, 1966). When they thaw, they fill with detritus; preserved ice wedge casts are valuable paleoclimatic indicators (Murton, 2007). In dry environments with less than 100 mm of precipitation, the lack of moisture makes it impossible for ice to fill the wedges; under these circumstances they may fill with eolian sand from winter winds or other detritus as in central Yakutia (Yershov, 1998). This is how **sand wedges** form (Pewe, 1959; Jahn, 1975). Their paleoclimatic significance is similar to ice wedges because they are genetically equivalent. In addition to ice cracks, other mechanisms are known to cause fissures in periglacial areas. Desiccation cracking is a common process that forms small polygons. Evaporation causes desiccation; in addition, experimental studies show that moisture loss in areas of ice formation also causes desiccation (Washburn, 1973, 1979). Another mechanism is dilatation cracking caused by differential expansion of surface materials. Heaving causes radial cracks and subsequent collapse after thawing creates systems of concentric cracks such as those observed in pingo formation (Benedict, 1970; Jahn, 1975).

Ice sorting

Ice sorting is a complex mechanism of soil displacement and segregation by particle size. The previously explained processes of clast heaving, needle ice action, and mass displacement all contribute to particle sorting. Laboratory experiments show three types of sorting (Corte, 1966, 1969; 1971). The first is heaving, which causes larger particles to move upwards, creating horizontal levels of different particle sizes. This is called **vertical sorting**. As the freezing front advances, the finest particles move in parabolic trajectories away from the front, leaving coarser particles behind. This is called **lateral sorting**. Finally, when domeshaped bulges form at the surface, coarser particles move by gravity toward the margins of the dome, leaving the finer material at the center. This **mechanical sorting** can also be observed in polygons where the clasts located in the center of the bulged cell move toward cracks and fill them. Experimental studies show that sorting increases with increased moisture and when freezing is slow.

Chemical weathering

In periglacial environments, frost-shattered clasts commonly accumulate on surfaces of variable slope gradients; this shows that gelifraction is the primary weathering process. Although this process generates primarily large clasts, it can also produce clay-size particles as has been demonstrated in experiments conducted in climate chambers (Guillien and Lautridou, 1970). Because water is in the solid state except during thawing stages, chemical weathering is generally limited (Hall *et al*., 2002). In addition, chemical reaction rates slow down at low temperatures. However, rounded blocks of diabase and sandstones that show exfoliation due to hydration and oxidation processes have been observed in some areas (Czeppe, 1964). In some cases authors interpret chemical weathering features as indicators of past non-periglacial conditions. In Mediterranean mountain environments, as in the Sierra Nevada of Spain, chemical weathering of mica-schist during the warm season is thought to be significant in silt production (A. Gomez Ortiz, pers. comm.). In coastal environments, wind-transported seawater spray often leaves a sodium-rich saline efflorescence on rocks. Snowmelt helps the saline water percolate into the rock; later precipitation of dissolved ions may fracture the rocks and cause them to disintegrate. This salt weathering process (haloclasty) is attributed to the formation of honeycomb and tafoni weathering along with wind action that helps transport the particles generated by the breakup of the rock (Washburn, 1969; Selby, 1972; French and Guglielmin, 2002). Even so, the activity and intensity of chemical weathering in these periglacial areas is not well understood because not much research has been done in this field. Surprising results were obtained from a nine-year study of an area of mica-schist and amphibolite in northern Sweden (Rapp, 1986). This experiment showed that 48% of the weathering products were removed by dissolution in runoff. Other research shows very low values for this type of transport. Carbonate rock solubility increases with increased carbon dioxide content in water; it reaches maximum solubility in water at 0°C. For this reason, periglacial environmental conditions favor the dissolution of carbonate rock. However, solution rates are lower than in warm areas due to sparse biotic activity, which is generally the main source of CO2 (Birot, 1960). Karkebagge conducted field experiments in which he installed limestone and granite disks in the soil; at the end of five years the limestone disks had weathered much more than the granite (Thorn *et al*., 2002). Ice and snow supply water during the thaw; carbonate dissolution creates karren and doline fields and widens discontinuity planes. Water penetrates into the bottom of the massif and forms shafts and cavities, especially in alpine areas. In areas of continuous permafrost, meltwater cannot penetrate the permafrost so it flows on the surface. Small cavities can form in discontinuous permafrost; they do not exceed 300 m long and stalactites and stalagmites may also form in them (Sweeting, 1972).

Mass movement

Mass movement is defined as the movement of detrital material downslope under the influence of gravity. These processes are significant in periglacial environments because the high water content of the active layer reduces friction and soil cohesion. Permafrost is impermeable and the top of the permafrost behaves like a water-lubricated plane to facilitate the movement of the overlying active layer (French, 1996). Different mass movement processes have different intensities due to climatic, topographic, and lithologic characteristics. In steep areas of glacial origin, marine cliffs, and fluvial scarps, water penetrates into cracks in the rocks; as it freezes it expands and can cause sporadic **rockfalls**. These occur primarily during the thaw and are exacerbated by basal undermining (Harris, 2007). The result is a gelifract-covered slope at the foot of a rocky cliff. Scarp retreat rates calculated by different methods show values of 0.3 to 0.6 mm/ year (Andre, 1993), well below those recorded in temperate and semiarid areas. Solifluction is one of the most common processes in periglacial areas. This type of movement can occur under a variety of climatic conditions; it is called **gelifluction** in periglacial regions (Baulig, 1957). **Frost creep** is a related process that results from soil displacement as it dilates perpendicularly into the ground during freezing and settles during thawing (Washburn, 1967).

Nival processes

Nival activity refers to rapid snow movements called avalanches and to slower movements related to static snow deposits; these processes are called nivation. **Avalanches** are rapid downslope movements of snow that commonly occur in mountains within temperate zones but rarely in polar regions. They occur on slopes with gradients from 15 to 50°; they do not commonly occur on steeper slopes because the higher gradient prevents a thick accumulation of snow cover. They can travel at velocities greater than 100 km/hr. Several criteria are used to classify avalanches (Lopez Martinez, 1988). An avalanche can be composed of dry or wet snow, may transport snow or snow and detritus, can slide over the ground, or can move through the air. Some avalanches are contained within a channel while others are not confined at all. Several factors cause the snow cover to be unstable (Embleton, 1979b). Avalanches are especially common after heavy snow falls that overload previously deposited snow. The presence of frozen ice layers covered by new precipitation creates structurally weak planes that can trigger avalanches. In addition, the percolation of snow meltwater can generate avalanches. Two fundamental types of avalanches can be distinguished (Furdada, 1996). **Powder avalanches** are an unformed mass with a very limited geomorphic effect although they can destroy trees and houses in their trajectories. **Slab avalanches** are rigid bodies of cohesive snow that move on a sliding surface controlled by an old snow layer (surficial avalanche) or by the snow-ground interface (fulldepth avalanche). Full-depth slab avalanches generally incorporate detrital material; they can have significant erosional effects. Avalanches are especially common during the thaw period and in winter (Owens, 2004). Movements of water-saturated snow masses or slush avalanches that flow down fluvial valleys in Arctic areas during the spring should also considered avalanches. They can transport blocks that weight up to 100 tons down slopes that have a 5° gradient (Hestners, 1985). Avalanches are highly variable in space and time. They are rare in some zones but occur regularly in others. Sometimes avalanches re-use previous avalanche routes; these are valuable indicators to use in avalanche risk cartography. As previously stated, whereas the geomorphic activity of powder avalanches is limited, full-depth avalanches can do significant erosion. Many talus slopes and detritus cones are partially produced by avalanches. Avalanches remove and transport loose particles; they erode as they move and deepen the channels they travel (Keylock, 1997). **Nivation**, first defined by Matthes (1900), designates a group of processes including gelifraction, gelifluction, frost creep, and runoff from snowmelt (Thorn and Hall, 2002). Water from thawing can freeze again where permafrost is present beneath the snow. In unfrozen soils, a snow thickness of more than a meter has a temperature near the freezing point and the snow layer protects the ground from freezethaw cycles (Embleton and King, 1975b). Thin snow cover promotes gelifraction. Along the eastern slopes of Penalara (Sierra of Guadarrama, Spain) nivation is effective only on outcrops of moraines or the weathering layer. The western slope shows no signs of nivation activity except where catastrophic mass movements have altered the slope profiles (Palacios *et al*., 2003). Nivation is more active in subarctic and alpine environments; the most notable geomorphic feature is **nivation niches** (Thorn, 1988a).

Fluvial activity

Fluvial activity in periglacial environments has not been studied extensively. Fluvial activity was traditionally considered irrelevant because soils are frozen for a significant part of the year. Periglacial area rivers do not have an allogenic supply; therefore they have flowing water only during the thaw period. For this reason researchers thought that runoff played a minor role in landform development

(French, 1996). Rivers in periglacial areas are supplied by snow and melted ice, rain, groundwater, a small quantity of frozen soil, and extraperiglacial runoff. Most rivers are strongly affected by snowmelt; their seasonal flow pattern has been described as a nival regime that consists of a high flow period sustained by meltwater with prominent diurnal runoff (Vandenberghe and Woo, 2002). Snow that falls in winter (Birot, 1960) thaws over a period of two to three weeks in the spring. Snowmelt causes floods whose hydrographs show a steep rising limb and an acute flood peak followed by a prolonged falling limb that lasts two or three months (Arnborg *et al*., 1967). During the summer, most recharge comes from fields of permanent snow or glaciers. In arid zones, rainstorms may occur during the summer. Although annual precipitation is small, runoff is concentrated over a few days; it is estimated that between 30 and 90% of the annual runoff occurs within two or three weeks of snowmelt flooding (Clark, 1988). Runoff finally ends at the onset of winter when a new ice stage begins. In periglacial environments, sediment transport is generally low; it may consist partially of glacial sediments. Bedload, which usually predominates, may account for 90 percent of the sediment load, dissolved load is negligible (<1%), and suspended load ranges from 5 to 25 percent (Church, 1972). Various methods have been used to estimate transported volumes in different drainage basins; values of surface lowering are between 0.03 and 0.5 mm/year (Arnborg *et al*., 1967). Beyrich *et al*., (2003) reach similar conclusions. Fluvial channels are predominantly *braided* because of their high bedload; channels are unstable and contain highly mobile bars that may be inundated during high-stage periods. River water causes thermal erosion; it helps thaw frozen riverbanks and undercuts the banks, leading to collapse. This continuous thermal erosion widens the bed and creates U-shaped channels (French, 1996). In addition, large rivers can create a talik under their bed and incise into it. Large rivers that cross periglacial regions are the MacKenzie in Canada and the Obi, Yenisei, and Lena in Siberia. Costard and Gautier (2007) summarize the hydrology and morphology of the Lena River. The Lena is 4,300 km long and has more than 500 tributaries with a total length of 60,000 km. The surface area of its drainage basin is 2.49 × 1016 km2. It discharges 525 km3 of water into the Arctic Ocean every year; it carries between 12 and 17 million tons of suspended load and 49 million tons of dissolved load. At its mouth a large delta has formed; it has a surface area of 30,000 km2 and supports abundant fauna and flora. Costard and Gautier (2007) collected this data from various Russian authors. The Lena River flows almost north-south and consequently crosses various latitudinal zones inhabited by a wide variety of flora and fauna. Boreal forest (taiga) covers 83% of the basin and 20° of the latitude. Near the Arctic Circle the landscape is arctic tundra. The Lena River crosses Yakutia where the Earth's lowest temperatures (−72°C) and maximum permafrost thickness (1500 m) have been recorded. Most of central Yakutia is not covered by snow because precipitation is sparse; the lack of snow has helped form and preserve continuous and discontinuous permafrost. The permafrost along the floodplain of the Lena River contains a large quantity of massive ice and ice wedges. Fluvial thermokarst is present as collapsed pingos and alases (Soloviev, 1973). Alases form as flat bottom landforms and depressions with steep slopes as massive ice layers melt. They may coalescence to produce mature alas valleys. The bottom of an alas has a chaotic topography of collapsed masses and temporary sag ponds. Alases range between 100 m and 15 km in length but are normally about 1 km long. Thermokarst is located primarily on the Pleistocene plains of central Yakutia (Czudek and Demek, 1970b). As a result, alases cover 40–50% of the plains of central Yakutia. At the mouth of the Lena River the subaerial delta covers about an area of about 28,500 km2 that remains frozen about six months of the year. Its tundra has various types of patterned ground and thermokarst alases. The fluvial system is composed of anastomosing rivers with multiple channels separated by permanent alluvial islands (Knighton and Nanson, 1993). As a result, the sediment load, which is primarily sandy, does not move large distances; instead, it accumulates in wide bars, long islands, and anastomosing branches (Gautier and Costard, 2000). The headlands of the islands undergo significant erosion, with average values of 12 m/year and maximum values reaching 20 m.

The Lena River is of interest because of its spectacular floods, which are controlled by periglacial dynamics. Yang *et al*. (2002) show that maximum floods are recorded in May in association with positive thermal anomalies. As flooding begins, the river transports large quantities of ice and floating trunks that create local dams. The dams cause the water level to rise, which rapidly floods the floodplain and sometimes even the first flood terrace. In May 2001, the region experienced its worst flood in 100 years, which affected hundreds of thousands of people. In many places, it was necessary to use explosives to destroy the dams. This event was caused by a combination of a severe winter and an abnormally warm spring (Yang *et al*., 2002). According to Yang *et al*. (2002), the average temperature in the Lena River basin has increased 1.3°C since 1930. This temperature rise increases the thickness of the active permafrost layer and the extent and subsidence of thermokarst depressions (Pavlov, 1994; Serreze *et al*., 2002). Climate change is undoubtedly affecting the behavior of the river in a substantial way. In a study of six large Eurasian rivers near the Arctic, Peterson *et al*. (2002) calculated an annual increase in water volume of 2 ± 0.7 km3, an increase of approximately 7%.

Wind action

In Arctic and Antarctic regions where extreme cold and arid conditions prevent vegetation growth, wind can be an important erosional agent because there are few obstacles to impede it (Yershov, 1998). During the Pleistocene, significant transport and deposition of eolian particles occurred over broad areas of Eurasia and the central United States (Embleton and King, 1975a, b). Silt constitutes an extensive deposit in central Europe but it is patchy in Great Britain. Three phases of eolian sand deposition have been documented in the period between the Last Glacial Maximum and the beginning of the Holocene but they do not strictly coincide with stadial or glacial conditions. Aridity, sparse vegetation cover, and delayed responses of eolian activity to climate change determine the timing of eolian phases (Kasse, 2002). In addition, the wind redistributes snow during the winter. In the summer, snow melts more rapidly on sun-facing slopes. These changes in snow cover are important in nivation processes and the creation of landforms such as asymmetric valleys and oriented lakes, which will be discussed below. In cold polar regions wind action predominates in summer, causing the deflation of fine particles from the land surface and creating periglacial pavements on alluvial plains and till piles. Eolian abrasion creates ventifacts (Pissart, 1966b; French, 2007). Overexcavation in some places causes deflation blowouts (Seppala, 2004). During transport, successive impacts create ventifacts or faceted clasts, yardangs in rocky outcrops (elongate mushroom forms), microyardangs in unconsolidated sediments (knife-shaped and aligned features), and tafoni near the base (French, 1996). High latitude dunes supply glacial and fluvioglacial accumulations. They are common in North America, Europe, Asia, Tibet, Argentina, and in areas of Antarctica that lack ice. Several small dune fields have been found in unfrozen areas of Antarctica. In their study of Victoria Valley (Antarctica), Bourke *et al*. (2009) calculated dune field movements at 1.5 m/year. The combined action of eolian transport and deposition of snow and mineral particles creates nivo-eolian deposits (Pissart, 1966b; Koster and Dijkmans, 1988). As in ice masses, differential melting produces topographies of small mounds and depressions that disappear from one year the next. The most outstanding deposit formed by particle accumulation in sand dunes and particularly as silt layers is called **loess** (Pewe, 1955 and 1959). Loess is transported by dust clouds that deposit sediments far from the source. It is a massive very well sorted deposit; 50 to 60 percent of the particles range from 0.01 to 0.05 mm (silt) in size and the rest are sand and clay size (Muhs, 2007). Loess is gray in color; as it weathers it becomes brownish-gray. It contains ice wedge structures, solifluction, and cryoturbation casts. During the Pleistocene, extensive accumulations of loess were deposited on the southern margins of the North American and Eura- sian ice sheets; the study of these accumulations has been of great paleoclimatic importance (Muhs and Bettis, 2003; Muhs, 2007). However, not all loess is periglacial; much of it originates in tropical desert environments. The dominant wind and atmospheric conditions (isobars) of past periods can be inferred from paleocurrents obtained from internal structures of old eolian deposits (Poser, 1950).

Periglacial landforms

The previously described periglacial processes are responsible for a variety of periglacial landforms that commonly form in two principal environments (1) high latitudes and (2) medium and low latitudes. They are most extensive at high latitudes with seasonal climatic contrasts whereas climatic variations are primarily daily fluctuations in mountainous areas of medium and low latitudes such as the Andean cordilleras, the Himalaya, and mountains of equatorial Africa. Differences in climatic conditions play an important role in most processes that generate periglacial landforms (King, 1976). Variations in climatic belts during the Pleistocene changed the distribution of periglacial environments. It is important to differentiate between relict and active periglacial

Patterned ground

Patterned ground consists of small-scale landforms of circular, polygonal, and banded geometries; they attracted the attention of early explorers in periglacial areas to the detriment of other more widespread morphologies (French, 1996). Patterned ground, however, is not specific to periglacial areas; similar forms occur in other environments (Hallet, 1990; Ahnert, 1994), especially in hot deserts (Hunt and Washburn, 1966). Thus we have a problem of morphological convergence or equifinality; different processes generate the same morphologies in response to variations of soil moisture. This is important to keep in mind when making environmental interpretations of relict forms. In the semidesert environment of the Ebro Depression and near Zaragoza (about 200–300 m of elevation), several investigators (Johnson, 1960; Brosche, 1971, 1972) interpreted various structures in terrace deposits as ice wedges and cryoturbation although these forms in fact result from the dissolution of underlying gypsum and swelling of clay-rich formations among other processes (Zuidam, 1976). The only morphologies specific to the periglacial environment are those caused by thermal cracking of frozen ground because these forms provide evidence of permafrost or intense cold. Periglacial patterned ground forms in the active layer; Washburn's classical morphologic systematization (1956) classifies patterned ground into circles, polygons, nets, steps, and stripes. The classification describes the presence or absence of particle sorting for each type. Circles, polygons, and networks develop primarily on horizontal surfaces. These cells tend to occur on slopes with gradients between 2 to 7° and they elongate to form stripes (Sharpe, 1938; Budel, 1960). **Circles** range in diameter from 0.5 to 3 m. Circles that lack sorting bulge and their vegetation-free center is cracked. Circles usually have a high content of fine material; cross sections show structures that indicate uplift of the central areas. Sorted circles have a rim of stones that surrounds an inner part composed of fine-grained material. The planar stones tend to rest on edge with their large faces oriented vertically (Furrer, 1968). Both sorted and unsorted circles may occur in isolation or grouped. Holness (2003) claims that circle formation is due to differential frozen uplift associated with convection cells. The most important processes involved in circle formation are differential frost heaving, mass displacement, and frost sorting (Jahn, 1975). **Polygons** with diameters less than one meter and those larger than one meter are classified separately because they may have different origins (Washburn, 1979). Small unsorted polygons may be as small as 5 cm in diameter whereas the large ones can exceed 100 m. They form in soils with variable sized particles where vegetation tends to concentrate along the margins and emphasize the polygonal pattern. Some polygons have ice wedges along their margins. These are domed cells during the growth stage and depressions with ponds during the recession stage (Jahn, 1972). The margins exhibit the opposite behavior. Some polygons may host small pingos and smaller polygonal forms inside the principal cell. In sorted polygons the margins are formed of stones that surround a cell of finer material. The minimum size of the smaller features is 10 cm and the maximum size of the large polygons does not exceed 10 m. Clasts decrease in size with depth and/or narrow downward or they may gradually enlarge and merge with a continuous layer of clasts (Jahn, 1975). Cracking is fundamental to polygon formation. Large polygons seem to be produced by frost cracking whereas the smaller forms result from desiccation or dilation cracking (Washburn, 1979). **Nets** are intermediate forms between polygons and circles; they can be sorted or unsorted. **Hummocks** are included in this group. These are mound-like forms covered by vegetation; their maximum diameters are 1–2 m and they are about 50 cm high. Their internal structure shows cryoturbation. They occur above the tree line in areas of low slope; their origin appears to be related to mass displacement (Schunke and Zoltai, 1988) or cryoturbation (Ballantyne, 2007) although other types of origins have been proposed. **Steps** are another type of patterned ground; steps form benches on slopes that dip 3–20°. The margins of unsorted steps are vegetated whereas the borders of sorted steps are composed of imbricated clasts (Sharp, 1942a). Unsorted steps may be derived from hummocks and the sorted variety may evolve from circles and sorted polygons (Washburn, 1979). Their origin may be related to differential mass movement. **Bands** or **striated soils** are a set of strips oriented parallel to the steepest slope gradient. Unsorted stripes consist of parallel lines of vegetation-covered ground and intervening strips of relatively bare ground. Sorted stripes are parallel lines of stones and intervening strips of fine material with the latter often several times wider than the coarser lines. Stripes are generally straight and may reach 120 m in length (Washburn, 1969). The clasts are usually on edge with the major axis parallel to the stripes. Clast size generally decreases with depth and the sorting penetrates a meter at most. On slopes with gradients greater than 2° sorted polygons grade progressively into sorted stripes (Collard, 1973) because mass movements seal the sides perpendicular to the slope gradient (Furrer, 1968). Attempts have been made to explain the origin of patterned ground in periglacial environments but it is still not well understood. Patterned ground is undoubtedly polygenic; similar forms may result from different processes. In addition, certain processes may generate different landforms (Washburn, 1979).

Ice-cored mounds

Ice-cored mounds are small hills produced by frost action. Most of them are generated as new permafrost forms in nonfrozen floodplain areas and limnic basins (Yershov, 1998). Researchers have proposed so many terms to classify them that significant terminological confusion has resulted. These mounds can be perennial like palsas and pingos or seasonal, such as hydrolaccoliths or blisters. The latter, which are 1–8 m tall and 3–50 cm in diameter, have not been studied extensively because of their ephemeral nature and formation in winter under snow cover. They form as groundwater freezes on top of permafrost (Pollard, 1988).

Palsas

Palsas are small hills that form in swampy areas and discontinuous or sporadic permafrost areas (Zoltai, 1971). They are present today in areas with a mean annual temperature below 0°C such as Canada, Iceland, Sweden and alpine regions (Kershaw and Gill, 1979). Palsas occur in isolation or in groups. They are distinguished on the basis of their morphology: dome-shaped, ribbon-shaped, and platform. The dome shapes are between 0.5 and 7 m high, 10–30 m wide, and 15–150 m long. Large size palsas tend to be considerably less conical than smaller ones. Ribbon-shaped palsas are lower in height; they may be parallel or perpendicular to contour lines and can extend several hundred meters in length. Platform palsas rise 1 to 1.5 m above the surrounding bog and may cover several square kilometers. They are related to differential thawing. Finally, palsa complexes are composed of several types of palsas at different growth stages; they contain numerous closed depressions created by thawing (Seppala, 1988). Some authors use the term *string bogs* to designate a group of different types of palsas. The surface of a palsa is cracked due to dilation and desiccation (Embleton and King, 1975b). Palsas are classified into two types on the basis of their internal structure. Palsas with a peat nucleus have a surface layer of dry peat that thaws in summer. Between 80 to 90% of the interior is composed of small ice crystals and lenses of segregated ice 5–10 mm in thickness that thickens to 5–10 cm toward the bottom. The base is usually composed of silt-sized mineral particles. Palsas with a silt nucleus or **lithalsas** (Pissart, 2000, 2002) have a surface layer of peat and a nucleus of silt and clay with lenses and veins of ice arranged in a horizontal and vertical reticular fashion (Seppala, 1988). Palsas form as a result of different heat conductivities of dry and saturated frozen peat; the heat conductivity of saturated and frozen peat is about twenty times higher than dry peat (Washburn, 1979). As the peat gets wet in winter, its thermal conductivity increases and it penetrates deep into the ice, forming segregated ice and dome-shaped bulges as a result of cryostatic pressures. The outer peat layer dries up in summer and has low heat conductivity; it acts as an insulator that prevents the inner part of the palsa from thawing. Palsas may disintegrate gradually from their margins inward because of rising temperatures; they may turn into small shallow sag ponds where a new generation of peat can begin (Friedman *et al*., 1971). Seasonal palsas linked to ooze have been reported in the Paramera of Avila Province (Central Spain) at an elevation of 1200 m (Molina and Pellitero, 1982). Although most palsas in Scandinavian countries formed 1000–3000 years ago, absolute dating indicates that some of them formed recently (Seppala, 1988).

Pingos

Pingo is the Eskimo word for mound. These structures are also called hydrolaccoliths; in Siberia they are called bulgunniaks. Pingos are conical shaped mounds that have an ice nucleus and that generally stand up above extensive plains. Submarine pingos were recently discovered on continental shelf. Scientists do not know if they grew under the sea or if they formed on land and were later invaded by marine waters. The highest documented pingo on land and above sea level is the 48 m high Ibyuk pingo, located near the MacKenzie Delta (Canada). Most pingos range in diameter from 30 to 600 m; their diameter is inversely related to their height. Some are more than a kilometer long and have a maximum height of 9 m. Because they maintain a constant diameter as they grow, their slope gradient increases but never exceeds 45° (Pissart, 1988). Two types of cracks occur at the surface. Dilatation radial cracks converge at the top; they result from the upward push of the ice core as it grows (Washburn, 1979). The opening of the cracks may lead to partial thawing of the underlying ice, leading to a craterlike subsidence depression at the center. Concentric cracks, which are not as conspicuous as radial cracks, result from thawing of the ice core during the negative growth stage of the pingo (Muller, 1959). Unlike palsas, pingos have a massive ice core that penetrates several meters below the ground surface (Tricart, 1967; Lundquist, 1969). The cover material is usually loose sediment such as gravel, sand, and silt but pingos with sandstone and shale covers have also been observed. The cover layer may reach a thickness of 14 m. Their growth rate ranges from very low to 1.5 m/year (Mackay, 1973). All known pingos are younger than 10,000 years; some of them are only a few hundred years old. As temperatures rise, pingos can gradually disintegrate and become thaw lakes. Two origins of pingo formation have been proposed. The so-called **closed-system or hydrostatic** origin suggests that pingos of the MacKenzie Delta (Canada) and Central Yakutia (Siberia) developed in areas of thick continuous permafrost (Mackay, 1979). They form in relation to lacustrine depressions. As permafrost develops, entrapped water freezes in the lake and creates a massive ice core through downward percolation and aggradation. The volume increase generates cryostatic pressure that domes the lake sediments and may even cause water to be extruded towards the surface. If water reaches the surface, it may form an aufeis; if it contains gas it may cause an explosion. Water does not generally reach the surface; instead it freezes and forms a massive ice core (Mackay, 1979).

The cryostatic origin of the MacKenzie Delta pingos is supported by the fact that 98% of the 1380 mapped pingos are located at the edge of or inside contemporaneous or old lakes (Stager, 1956). A number of **open-system or hydraulic pingos** are located in areas of relatively thin permafrost in Alaska and Greenland (Mackay, 1979). They form as a result of a pressure gradient due to height differences (Holmes *et al*., 1968); for this reason, pingos often form in topographic lows such as valley bottoms or distal sectors of low angle slopes (Muller, 1959).

Slope morphology and evolution

Slopes in periglacial environments have a wide range of morphologies because they develop under varying conditions of temperature, moisture, lithology, and vegetation cover. Multiple processes may be involved in their origin such as congelifraction (frost shattering), mass movement, runoff, or nivation. These processes may operate together or separately with variable intensity. Many of the resulting forms are not specific to periglacial regions; they also form in other morphoclimatic zones (French, 2007). In addition, some periglacial slope morphologies formed in the past (relict forms) and may not be in equilibrium with current conditions.

Gelifluction slopes

Differential movement and downslope displacement of colluvial deposits due to frost creep and gelifluction create landforms that can be distinguished on the basis of their geometry (Washburn, 1979). **Gelifluction sheets** have a broad lateral extent and a festooned lower edge. They form on slopes with gradients greater than 1 to 3°. **Gelifluction benches** are characterized by a terrace shape; their long dimension tends to run parallel to topographic contours. **Gelifluction lobes** have a tongue-shaped morphology and widths of 30–50 m; in combination with benches they may have gradients of 20–25°. The center of the lobe moves faster than the margins (Harris, 1981; Matsuoka, 2008). Gelifluction deposits that elongate considerably in the direction of maximum slope are called **gelifluction streams**. In summer these masses may slide (as *debris flows*) over glacial ice at the surface, which is at the melting point; this promotes basal sliding. All these landforms develop easily on slopes with sparse vegetation and on sun-facing slopes where high solar radiation promotes thawing. Gelifluction deposits are usually heterometric; some have rough stratification. Clasts in these deposits are commonly angular and are oriented with their major axis parallel to the direction of movement (Cailleux and Tricart, 1956). These characteristics may also be present in deposits caused by gelifluction in other morphoclimatic regions but those that form in periglacial conditions are more angular and are not affected by chemical weathering.

Cryoplanation terraces and cryopediments

Cryoplanation terraces and cryopediments are low-angle planated surfaces that form in periglacial regions; they have been given multiple descriptive and genetic names in the geomorphic literature. They are intimately related to the periglacial cycle (Peltier, 1950; Birot, 1960). **Cryoplanation terraces** form in the middle and upper parts of hillslopes and mountains up to elevations of 3,000 m. They have a terraced profile and outcrops of bedrock and *tors* (rocky residual relief) (Czudek, 1964). These terraces form on slopes with gradients lower than 25° and at the foot of small scarps. Widening of the highest terrace may eventually form a planation surface at the summit. Terrace widths range from 5 m to more than a kilometer and their lengths range from 30 m to more than 10 km. The gradient ranges from 1 to 14° and the scarp height may reach 50 m. The contact between the scarp and terrace is sharp and usually has an accumulation of snow. Gelifluction deposits up to 3 m thick may cover the terraces (Priesnitz, 1988). Initially, terraces may be structurally controlled although their formation is linked to nivation processes involving material supplied by congelifraction and snowmelt water and mobilization by gelifluction and pellicular runoff. These nivation processes cause scarp retreat, widen the terrace, and increase snow retention capacity (Demek, 1969). Excellent examples of cryoplanation terraces have been documented in the Sierra Nevada (Betic Cordillera, Spain), in the Pyrenees (Gomez Ortiz, 1996), and in Hardanger (Norway). **Cryopediments** are slightly dipping erosional surfaces that form at the foot of slopes in valley margins. Generally, only one level is present although in some cases several stepped levels may form. They are larger than altiplanation or cryoplanation terraces; they can extend to several tens of kilometers. Cryopediment slopes range from 1 to 12° and have a straight or slightly concave longitudinal profile. A thin veneer (<2 m) of detrital material often covers them; bedrock may crop out locally. Gelifluction landforms are common in the proximal sectors of cryopediments and patterned ground commonly occurs in distal areas (Priesnitz, 1988). Cryopediments form as a result of frost action in the source area where gelifluction and runoff mobilize detritus; this also occurs on cryoplanation terraces. The latter process is more active in cryopediments and as a result, particles are transported longer distances. In distal sectors, runoff is diffuse and particles are transported in shallow braided fluvial channels. Continuous particle production and transport cause slopes to retreat slowly and cryopediments to enlarge (Czudek and Demek, 1970a). Cryopediments have an obvious morphological resemblance to hot desert pediments although their origin is different (Demek, 1969).

Talus slopes and debris cones

Talus slopes and cones are accumulations of angular clasts on slopes. Although talus is common in periglacial environments and especially in alpine areas, it also occurs in other morphoclimatic zones such as deserts. **Talus slopes or scree** are wedge-shaped heaps of fragmented rock that cover the middle and distal part of the slope; clasts come from higher cliff areas composed of resistant rocks (Luckman, 2007). The thickness of talus slopes ranges from very thin to 30 m (Brunner and Scheidegger, 1974). Talus slopes commonly have a concave longitudinal profile with a higher gradient at the top. The talus slope reflects the angle of repose of the coarser particles, which ranges from 25 to 40° (French, 1996). If the scree slope is long, clast sorting may occur; the smaller clasts are at the top and the larger ones below, as a result of their greater kinetic energy (Washburn, 1979; Church *et al*., 1979). The long axis of elongate clasts is usually parallel to the slope gradient. Multiple processes can fragment rock at the rock scarp near the top of the slope although congelifraction is the dominant process in periglacial regions. The continuous supply of frost-shattered clasts causes rock scarp retreat; this process is largely controlled by rock type and structure. Retreat rates of about one mm/year have been measured in Laponia and Spitzbergen (Rapp, 1957) and one to three mm/ year in Great Britain (Ballantyne and Kirkbride, 1987). Clasts move along the channel by rolling, creep, and occasionally by sliding. Movement is usually restricted to the upper half meter of the channel; velocity is highly variable, ranging from 1 to 500 cm/year (Washburn, 1979). When rock fragments from congelifraction concentrate in a channel, they incise bedrock as they move; these channels may become preferential pathways for rock and snow avalanches. Rock fragments deposited at the foot of these channels create **debris cones**, which are similar to talus slopes although they have convex profiles and usually an obvious basal concavity (Luckman, 2007). Adjacent cones may coalesce and develop morphologies similar to talus slopes.

Block fields, block slopes and block streams

These occur on top of or at the foot of hard rocks of various lithologies that have bedding or joint planes spaced decimeters apart. Weathering, which is primarily congelifraction in periglacial environments, acts on these rocky masses and generates block-size clasts that measure decimeters or meters in addition to smaller size particles. These block accumulations may be covered by vegetation; in this case interstitial fill is present within the block framework. When fine particles at the surface are removed by water, the deposit develops an open texture in which the blocks support each other. Depending on their topographic position, these accumulations may be called block fields or plains, block slopes, or block streams, which are located in valley bottoms or in previously existing incisions on slopes (Washburn, 1979). Most of the forms described in the literature are relict (Wilson, 2007). All these morphologies occur in elevated areas at high and alpine latitudes. **Block fields** are extensive accumulations of angular clasts located on plateaus or summits at high or medium latitudes (Rea, 2007) where more than half the surface is covered with blocks and smaller size particles. Block fields develop on horizontal or low gradient surfaces; an inclination of <10° is the limit between block fields and block slopes (Rea *et al*., 1996). They show evidence of intense gelifraction activity (Smith, 1953; Dahl, 1966). Numerous block fields are present in the northwestern Iberian Peninsula (Valcarcel and Perez Alberti, 2002; Garcia de Celis, 2002). **Block slopes** are relatively common accumulations; they occur at the foot of scarps of resistant rocks where frost weathering liberates gelifracts that move downslope. The long axes of clasts are generally oriented in the direction of maximum slope; the deposit is sorted with larger clasts at the surface and finer ones at depth (Gutierrez and Pena, 1977; Boelhouwers, 1999). A high percentage of fine particles can be seen in distal parts of the slope; clast size decreases progressively from top to bottom. Differential movement of the accumulation causes lobular and benchlike shapes in the middle and lower parts of the slope (Gutierrez and Pena, 1977). **Block streams**, which are made up of large gelifracts, occupy slope gullies or the bottom of principal valleys. The latter are much less common than the former; they are supplied by blocks from the slopes and they end up forming block streams. A spectacular example of a block stream is in the Paleozoic Tremedal massif in the central-eastern sector of the Iberian Cordillera (Gutierrez and Pena, 1977) where most main and slope valleys contain block streams. The maximum dimensions of block streams are 2.6 km long and 0.25 km wide. Their thickness is difficult to estimate because the bedrock is not visible although thicknesses of about 4 m have been measured. The longitudinal slopes of principal valleys vary widely even within the same valley. There may be transverse steps several meters tall, possibly of structural origin. On the block surface small closed depressions about a meter in diameter result from the illuviation of underlying fine material and subsequent settling. Beneath the open-texture surface blocks, a heterometric matrix fills in between the blocks. The absence of fines at the surface is attributed to runoff washing away this material, possibly as a result of snow melt (Andersson, 1906; Smith and Smith, 1945; Smith, 1953; Potter and Moss, 1968). This lack of fine material prevents plant growth. No fabric type has observed in the Tremedal block streams, unlike other areas where researchers have observed lobe shapes in the block arrangement and high dips of the long axes in the cross-slope direction (Potter and Moss, 1968; Caine, 1972). Movement in the Tremedal must have been minor because none of the block streams in the principal valleys has reached the lower part of the massif. If this had reached the lower part, they would have generated block pediments or fans of blocks. Today, water circulates under the block accumu- lations and flows out at the foot of the transverse steps. Lichens cover the blocks, which indicates lack of movement. Harris *et al*. (2001) calculated velocities of 0.6 to 2.1 m/year in field experiments in block streams in the Qinghai Province of China. The processes involved in the generation of these morphologies are congelifraction, ice creep, and gelifluction, sorting by ice and by movement of fine particles to the base of the deposit, and mechanical illuviation of the interstitial matrix (Washburn, 1979). Even so, the origin of the Tom Mays Canyon (Texas) block stream studied by Degenhardt *et al*. (2007) has been interpreted as a rock avalanche.

Rock glaciers

Rock glaciers are outstanding features of alpine periglacial areas. They are tongue- or lobe-shaped masses of angular clasts; they have ice in the interior and they move downslope (Barsch, 1996; Kaab, 2007). The ice cements the detritus or is an ice nucleus covered by rock fragments (Potter, 1972); it creates a surface depression as it melts. This feature has caused problems in classifying rock glaciers as glacial or periglacial forms (Tricart and Cailleux, 1962). Today, researchers agree that glaciers with an ice nucleus are part of the glacial system and those with interstitial ice are typical periglacial morphologies characteristic of permafrost environments. In addition, confusion in the distinction between block streams and slopes and rock glaciers arises from the transition that sometimes exists between these landforms (Corte, 1976). Rock glaciers generally occur at the foot of steep walls that may be the margins of cirque amphitheaters or abrupt slopes of trough valleys (Serrat, 1979; Gutierrez and Pena, 1981; Alonso and Trombotto, 2009). Rock glaciers are supplied by congelifraction clasts that move downslope creating talus slopes, block slopes and streams, and debris cones, and occasionally moraine deposits. They are classified morphologically (Wahrhaftig and Cox, 1959) as **tongue-shaped rock glaciers** if they have a length/width ratio greater than one, **lobate rock glaciers** if their length/ width ratio is less than one, and **spatulate rock glaciers** if they are similar to tongue-shaped rock glaciers but have broader distal parts (Barsch, 1996). Some authors add other types: **composite rock glaciers** formed by the superposition of several individual rock glaciers (Washburn, 1979). Rock glaciers range in width from 100 to 500 m; they can exceed a kilometer in length and range up to 50 m thick. Rock glaciers have scarped fronts due to their movement; the gradient at the front ranges from 35 to 45° (Barsch, 1988). The lateral margins are also abrupt. Their interior has a surface relief of arcuate ridges and grooves that is generally perpendicular to the flow direction. The volume transported by a rock glacier normally exceeds a million cubic meters of which 40–50% is composed of clasts and 50–60% is interstitial ice and ice lenses. Their movement, which results from the plastic deformation of ice, is controlled by slope, ice content, temperature, and the size of clasts in the interior. Their velocity, which is usually higher in summer than in winter, ranges from several centimeters (Barsch and Hell, 1975) to 5 m/year. The highest velocities have been measured at the rock glacier in Obergurl (Innsbruck, Austrian Alps) at abrupt changes in slope. The weight of the rock glacier and the push of slope deposits on the slope above the entrainment zone help move the glacier. Rock glaciers move by creep and ice deformation processes; maximum surface velocity occurs at their center line and internal velocities decrease parabolically toward the rocky substrate (Kaab, 2007). Rock glaciers move more slowly than ice glaciers and much faster than deposits that move as a result of gelifluction. Considering their functionality, we can distinguish several types, including **active rock glaciers,** which have an external 2 meter layer of coarse detritus and a nucleus of finer material with ice lenses in which motion can be determined by the advance of the front. Active rock glaciers are part of permafrost; they indicate continuous or discontinuous permafrost on a regional scale (French, 2007). Recently active rock glaciers have been found at 2,600 to 2,950 meters in the central-eastern Pyrenees (Spain); they move 22 cm/year (Serrano and Agudo, 1998; Serrano *et al.*, 2002). Recent geophysical studies using radar (GPR) on various rock glaciers throughout the world have established relationships between internal structure and surface morphology. Lobate sedimentation is related to catastrophic episodes of rock falls from cirque walls. Longitudinal GPR profiles suggest that the morphological difference between tongue and lobate rock glaciers may be due to cirque geometry, the frequency and location of discharges of detritus, and the magnitude of the valley gradient **Inactive rock glaciers** do not move but still have a frozen nucleus and an external layer that exceeds a thickness of 10 m (Barsch, 1988). The first symptoms of inactivity are plant colonization on the front and the development of a detrital fan at the foot of the frontal slope. Finally, in **relict or fossil rock glaciers** (Barsch, 1977b), the internal ice has melted and thus they may have cryokarst collapse features. Rock glaciers develop best in continental and semiarid climates because extremely wet conditions generate glaciers (Hollerman, 1983; Kaab, 2007). Research shows that active rock glaciers are less than 10,000 years old; thus they formed after the last glaciation of the Wurm, that is, during the Holocene or Postglacial (Barsch, 1996). In the South Shetland Islands, west of Antarctica, Serrano and Lopez- Martinez (2000) studied nine rock glaciers and categorized them into three groups: (a) those that postdate the last major glacial retreat; (b) those that have an age of 2000 year BP but predate the Little Ice Age, and (c) those that formed during the Little Ice Age. The age estimated by OSL (Optical Stimulated Luminescence) for block streams in the Falkland Islands ranges from 54 to 16 Ka, which suggests that they have been active since 16 Ka (Kaab, 2007; Hamson *et al*., 2008). Andre *et al*., (2008) think they formed by denudation of an alteration profile of subtropical or temperate origin, possibly of Tertiary age, and were later reworked by periglacial processes.

Grèzes litées

These are slope deposits of stratified angular clasts. Some authors call these deposits *eboulis ordonnes* although others use this term for stratified screes formed by gravity. In **grèzes litées** the strata are layers of equigranular pebble-sized clasts alternating with finer material that is predominantly sand, silt, and clay (Tricart, 1952a). If the deposit contains numerous boulder-sized fragments, it is called *groise litee*. In the grezes, the large clast layer has a grain-supported structure; that is, there is no matrix in the interstices, and as a result, the clasts are loose. When these deposits are composed of limestone components, the layer of fine material is usually cemented because it retains percolation water. The transition from one layer to another is usually clear and channel-shaped structures can sometimes be observed. Unlike stratified screes, stratification in a grezes litees deposit has nothing to do with the present-day topographic surface. Grezes litees occasionally fossilize glacial moraines as in the Andes of Ecuador and Bolivia (Heine, 1995) or in Santa Elena at the head of the Tena Valley in the Central Pyrenees (Spain). Grezes litees occur in any orientation although they are most commonly located on slopes with southern exposure. They are found on slopes that range from 7 to 45°; they can be 40 m thick. The layers tend to be thicker downslope (Guillien, 1951). Grezes litees are relatively common in mid latitudes but are unknown in polar zones. Their formation requires thermal oscillations around 0°C that enable congelifraction to generate gelifracts, enough water to sort the deposit, and scant or non-existent vegetation (DeWolf, 1988). Stratification and sorting are interpreted to be due to meltwater from snow accumulations, which is the essential agent for sedimentation of coarse material and the eluviation of the smallest fines (Guillien, 1951). Francou (1988, 1990) and Lautridou and Francou (1992) observed typical stratified slope deposits in the Andes of Peru at 4,400 to 4,900 m of elevation. These deposits have gradients of 33 to 35°; they are characterized by large solifluction lobes that move several centimeters a year. A cross section of the deposit along the lobe margin reveals its internal structure; the lobes move very slowly and develop like a carpet that buries clasts under the lobes. At the same time the ice push/lifting separates fines from clasts, which move by ice *creep* and *pipkrake* at higher velocities than solifluction lobes. Van Steijn *et al*. (1995) show that research on the morphology and sedimentology of stratified slope deposits can be used to infer the processes that formed them. Vari ous origins have been proposed, which may act alone or in combination: solifluction laminae or lobes, flows of dry particles or particles cemented by ice, *debris flows*, and ice-eolian transport. Researchers also discuss the problem of convergence of landforms. In the central Pyrenees (Spain) a large number of grezes litees occur primarily at elevations of 600–800 m. Researchers think there may be more than one phase of formation although the principal stage is the Late-Glacial (Pena *et al*., 1998) date of stratified debris at Eripol, 20,060 ± 180 BP. Garcia Ruiz *et* *al*. (2000, 2001) think the deposits formed by gelifluction lobes and debris flows during the coldest stage of the upper Pleistocene and during the Late-Glacial.

Nivation landforms

Nivation niches are found on slopes beneath a thin film of snow. Meltwater penetrates into bedrock material and at night when temperatures reach the freezing point, frost weathering occurs. The particles generated by this process can move downslope by gelifluction and ice creep aided by runoff; they can be seen as debris accumulations at the foot of the niche (Thorn, 1988b). Particle leaching inside the nivation niche generates periglacial pavements. This evidence has led to the idea that the most important nivation process is the freeze-thaw cycle. In low latitude areas such as the Cordillera of the Andes, more than 300 annual freeze cycles have been recorded (Troll, 1944); these are areas where nivation may be potentially significant. The continuous activity of nivation processes enlarges the niches, which begin as small depressions about a meter across; later they can reach dimensions of almost a kilometer if they coalesce with other niches. They enlarge along margins where melting and subsequent congelifraction are most active because the protective action of the snow in the center prevents the freeze-thaw process. These margins are modified throughout the year as a function of seasonal variations. Niches that reach large dimensions are called **thermocirques** (Selby, 1985). Nivation niches have a scarp or arched amphitheater and a terrace or low gradient slope at their base, which aids in the removal of detritus (Christiansen, 1998). They form more rapidly on rocky outcrops than in vegetated soils; they have been observed to form in a few years in nivation niches beneath snow patches. The niches are embryonic cirques where glaciers gradually take the place of nivation processes when snow that accumulates in the interior of the niche persists long enough to be converted to ice (Tricart and Cailleux, 1962). **Nivation ridges** can form at the foot of snow accumulations that are present at the base of scarped slopes; these are gelifract ridges (Butzer, 1976; Luckman, 2007) or **protalus ramparts** (Washburn, 1979). Clasts from the upper rocky part of the slope slide on top of the snow and accumulate at its margin. When the margin melts, the ridge remains behind and gets disconnected from the slope profile. Ridges that are parallel to each other have occasionally been observed; this indicates parallel slope retreat (Sharpe, 1938).

Slope evolution

Slopes that form in periglacial environments are classified according to their shapes and the processes that generate them. The primary types are 1) cliff-talus slopes, which are concave and shaped like talus accumulations and detritus cones, 2) benched slopes formed by cryoplanation processes, and 3) convex-concave slopes with continuous cover of solifluidal material (French, 1996). Cliff-talus slopes evolve by parallel scarp retreat; over time the slope lengthens and its angle and the length of the detritus layer decrease (Carson and Kirkby, 1972). In general, scarps in periglacial environments retreat more slowly than in other morphoclimatic zones although this is a function of lithologic and structural characteristics and microclimate, such as orientation (Souchez, 1966). Benched slopes result from cryoplanation processes that soften and flatten the profile. Continuous erosion of high areas and sedimentation in low areas, primarily in valley bottoms, involves parallel scarp retreat by congelifraction and enlargement of benched areas. These ideas were expressed by Peltier (1950) in his periglacial "erosion cycle". Concave-convex slopes have a convex-concave profile that lacks transitions; they are covered by a continuous deposit of gelifracts and are affected by gelifluction, surface washing, and rills in high gradient areas. These slopes may evolve into **tripartite slopes** or **triangular slope facets** (Budel, 1982) when they are deeply dissected by gullies in the middle part of the slope where tributary gullies flow together. Material that has been eroded from the lower parts of slopes is deposited by gullies as small alluvial fans. Once the triangular facet has formed, denudation processes change so that each facet constitutes a relict landform that functions independently of higher parts of the slope. The alternation of accumulation and dissection stages creates a sequence of tripartite slopes separated from each other and in turn parallel to the scarp. Research on triangular facets on Seymour Island (Antarctica) shows that the transition from the accumulation stage to the dissection stage is a geomorphic threshold related to glacio-isostatic adjustment and/or climate change (Gutierrez *et al*., 2011b).

Fluvial landforms

Braided channels with significant bed load, high inclination, and high width/depth ratio are the most common channel type in periglacial environments. During summer flood periods, these channels do significant mechanical erosion on channel margins. In addition, runoff water melts the permafrost and undermines river margins; this creates **thermo- erosional niches** (Czudek and Demek, 1970b; Walker, 1973) or thermal-fluvial erosion (French, 2007). This erosion may extend to heights of 3 meters and depths of 1–3 meters. When the niches collapse, they release large masses that provide new load to the river. This explains why thermal erosion is more effective than mechanical erosion but this is true only of non-cohesive sediments (Jahn, 1975). The same erosion occurs along coastal cliffs. The morphology of valleys that form in periglacial zones has not been well studied; most research is related to the problem of valley asymmetry. The scientific literature on Pleistocene valleys is more abundant; these are relict forms in areas that were subject to periglacial processes such as wide areas of central and western Europe. **Asymmetric valleys** are valleys in which one side is notably more inclined than the other. They are common in middle latitude periglacial environments. The asymmetry is not specific to periglacial environments; it can result from various factors, primarily lithological and structural. However, in periglacial regions in the northern hemisphere, the most inclined slope is oriented toward the north. More insolation on sunny slopes causes rapid and long-term thawing, which generates a larger amount of water to trigger gelifluction and mass movement. The eroded materials direct fluvial currents toward shady areas and cause basal undermining; as a result, these slopes will be more steeply inclined (Carson and Kirkby, 1972). In addition, north-oriented slopes experience less thawing during the summer so permafrost helps preserve their steeper slope (Gravis, 1969; Czudek, 1973b; Karrasch, 1993). Another reason that may explain valley asymmetry is that dominant wind activity promotes snow accumulation on leeward slopes (French, 1971). During summer, melting provides a large amount of water that accelerates mass movement on slopes. Low order segments of fluvial basins in non-mountainous regions of middle latitudes have different morphologies. One landform is the **U-shaped valley** (Tricart, 1967), which is U-shaped in cross section and has a longitudinal profile with slight reverse slopes, indicating that runoff plays a minor role. The slopes have a convex-concave morphology because gelifluction is the dominant agent of detrital transport. The valley bottoms are primarily composed of unsorted gelifluction deposits from the slopes. These valleys are common in old periglacial regions of Europe and most are now dry. U-shaped valleys generally do not exceed 3 km in length and toward the bottom they flow into a valley with stream flow or they may become **flat-bottom valleys**, which are the most characteristic landforms in Budel's (1977) ectropical zone. In Iceland they are 5–100 m wide or more, 5–20 m deep and have a gradient of 4–11% (Schunke, 1975). Their bed is primarily composed of non-sorted detritus from slopes and materials that have undergone some longitudinal transport. The connection between the flat bottom and the slopes is concave; it shows the lateral solifluction supply and the inability of the river to remove these deposits. In other places, the junction between the slope and valley bottom is a basal angle that shows significant erosion at the foot of the slopes during flooding (Tricart, 1967). Like U-shaped valleys, these flatbottom valleys are primarily dry valleys today (Jahn, 1975). Dry valleys are also called *dells*; they have been described in central-western Europe

(Washburn, 1979; French, 2007). This dryness may be due to the considerable thickness of the bottom deposits where water can percolate easily by flowing through alluvium. However, dry valleys form in other environments such as karst areas.

Thermokarst or cryokarst

Thermokarst is the term for landforms generated by land surface subsidence resulting from thawing of the upper part of permafrost. The resulting landscape may be highly irregular and even chaotic in places (thermokarst relief in badlands of French, 2007). The term thermokarst is hardly appropriate since the forms produced are not at all similar to exokarst landforms present in carbonate terrains (Pissart, 2004). Thermokarst morphologies occur primarily in low-relief areas such as alluvial sediment plains because the resulting thaw lakes rarely occur in zones of high relief (Williams and Smith, 1989). The landforms are usually active and as a result, periglacial landscapes are rapidly modified in short periods of time (Harry, 1988). Thawing results from multiple causes. Change of climate toward milder temperatures is the most significant cause affecting large areas. Modifications of plant cover due to fires, deforestation, and construction are locally significant because the loss of cover can trigger thawing due to removal of the insulating vegetation layer. Other geomorphic causes are thermal erosion resulting from standing water, slumping, and mudflows that erode the active layer (French, 1987). **Thaw lakes** are the most common forms. They are shallow, less than 4 m deep, and are not wider than 2 km (French, 2007). Thaw lakes form from melting of the permafrost, which triggers ground subsidence that later fills with water. They occur primarily in plains and fill rapidly with sediments and peat. Thaw lakes usually do not persist more than 3000 years. Groups of **oriented lakes** are rare. The best example occurs in the coastal arctic plain of Alaska (Carson and Hussey, 1962). They are elliptical and sometimes rectangular in shape, two or three times longer than wide, with longer axes up to 5 km and depths of 2–6 meters (Black and Barksdale, 1949). In recent work, Hinkel *et al*. (2005) show them to be 15 km long and 6 km wide. They may have complex morphologies due to coalescence of several lakes. Their origin is uncertain; they may be due to the dominant wind activity that deposits sediments on shores perpendicular to the wind directions. The deposits isolate the thawing shores and protect them from wave action; for this reason, beaches have a very smooth gradient. From this time forward, reflux toward the extremities of the lake causes a stronger current in these parts and has an abrasive action that enlarges the lake. Measured retreats are rapid, on the order of 25 m/year (Carson and Hussey, 1962). Even so, con- siderable controversy surrounds explanations for the orientation of these lakes. **Pingos and collapsed palsas** can form closed subcircular depressions; they are shallow and have elevated borders composed of sediments that dip toward the depression (Pissart, 2000, 2002). They fill with lacustrine sediments and peat (De Gans, 1988). These morphologies occur in Europe and Asia in areas that were previously periglacial regions. They are difficult to distinguish from other thermokarst depressions or other fluvioglacial kettles. When the frozen soil crops out and melts, **thaw collapses** form; these are arc-shaped depressions that open in the downslope direction. They may unite to form a festooned scarp. As they thaw, the surface permafrost becomes saturated with water and flows, leaving a scar several meters high. These cracks retreat at a rate of 7–10 m/year (French and Egginton, 1973), one of the most rapid erosional processes in the periglacial environment. When these collapses occur in areas of ice wedging they create large collapses. An **alas** is a closed depression that forms in low relief regions; it has abrupt margins and a bottom covered with grass and water. The name comes from central Yakutia (Siberia) (Soloviev, 1962, 1963 in Jahn, 1975; Czudek and Demek, 1970 b; Soloviev, 1973; Yershov, 1998). Alases are rounded or elliptical and often contain a shallow lake at their center. Their diameter ranges from 0.1 to 15 km and their depth from 3 to 40 m. Alases form as a result of permafrost thawing; the coalescence of several depressions may form alas valleys tens of kilometers long. As thawing progresses, different morphologies appear. Melting of ice wedges leads to linear depressions separated by bulged cells called *baydjarakhs*, which are 3–4 m high and 3–20 m in diameter (stage I). The progressive melting of ice produces a closed depression with disintegrating conical hills known as *duyoda* (stage II). Stage III of alas formation occurs when the bottom is devoid of trees and covered with grass, abrupt margins are present, and lakes perhaps occupy the interior. If the permafrost is reactivated, ice veins and pingos can form at the bottom of the depression (*Khonu*); these features may collapse later (stage IV). In Yakutia, 40–50% of the original surface has been modified by alases; they took several thousand years to form and are related to the Climatic Optimum (2500–9000 BP). At present more than 10% of this territory is active thermokarst (Guttierez, 2013).

UNIT-8: ELEMENTS OF SLOPE AND DIFFERENT APPROACHES TO STUDY SLOPE DEVELOPMENT

Slopes cover most of the Earth's surface (Young, 1972) and thus are a fundamental component of relief. The study of their shape, significance, and evolution is one of the basic subjects of geomorphology (Douglas, 1977). Comparatively little research was done on slopes until the last five decades. However, research on slope processes has now proliferated in both experimental field plots and the laboratory, models have been developed to assess danger, and procedures have been established for analyzing slope stability. Even so, the research is difficult due to the complexity of the phenomenon because the mechanisms at work are not known in a detailed way (Carson and Kirkby, 1972). In addition, the features used to recognize landslides are rapidly covered by vegetation in humid climates, making it difficult to identify them. The principal exceptions to this comment about slopes are found on the extensive plain that makes up the Colorado Plateau, the paramos of Castile (Spain), the hamadas of the Sahara, erosional surfaces of the Brazilian, African, and Australian shield, and the wide alluvial plains along the Indus-Ganges and the Po. Early geomorphologists gave special emphasis to the long-term slope evolution. According to Davis and as discussed in various publications, particularly Davis' (1899) first theory of general evolution of relief, slope evolution consisted of a gradual decrease of slope angles in valleys. Penck's (1924) morphological analysis system showed that if fluvial excavation accelerates, the resulting slope profile is convex whereas if downcutting is slow, the slope is concave. King (1953) proposed a third theory of landscape evolution based on slope research by Wood (1942). King proposed that slopes retreat in a parallel manner, a theory that Powell had already published in his 1875 Exploration of the Colorado River of the West, although it had its origin in the work of Fischer (1868, 1872). Researchers have occasionally discussed the contribution of landslides to landscape evolution. Cendrero and Dramis (1996) propose a ratio between the rate of landslide movement and the temporal magnitude of the lowering of the ground surface. Analysis of available data from various European regions shows that in some cases mass movements are the principal process of topographic evolution. Strahler (1950b) proposes a slope classification based on the result of a frequency distribution analysis and field studies of slopes in numerous locations. He classifies them as three types. **Highly cohesive slopes,** those composed of clays or strong massive rocks such as granite, schist, or gneiss, have angles of 40° and 50° and are affected by landslides; rivers undercut the bases of their slopes. The second type, **slopes in repose**, which are called block-controlled slopes by Bryan (1922) and gravity slopes by Meyerhoff (1940), have angles of 30°–35°. They are composed of large clasts controlled by the angle of repose. Laboratory research on the angle of repose of fragmentary materials shows characteristics similar to valley slopes (Van Burkolow, 1945). The third type, **slopes reduced by wash and creep**, called rain-washed slopes by Bryan (1922) and wash slopes by Meyerhoff (1940) correspond to the *haldenhange* of Penck (1924). The activity of water and creep reduces slope inclination to less than the angle of repose. The morphogenetic forces acting on a slope are perpendicular and parallel components (Jahn, 1954). The perpendicular component includes weathering processes and pedogenesis, which tend to form a weathering film. The parallel component refers to the displacement of material by gravity (landslides, gullying, solifluction, creep, etc.). As a result, the two components are interdependent and the evolution of a slope is a function of the magnitude of both. Jahn (1954) calls the relationship between the two **the denudation balance**; Tricart (1957) prefers the term **morphogenetic balance**. Schumm and Chorley (1966) use similar reasoning in their work on the Colorado Plateau; they call the quotient between these two components the **weathering relationship**. When the perpendicular component exceeds the parallel, the regolith deepens and does not move totally downslope. On the contrary, if the parallel component exceeds the perpendicular, the regolith is rapidly denuded and the slope stabilizes when the denudation reaches coherent rock. If it is soft, badlands form. If the relationship is equal to 1, it indicates that the amount of altered material is equal to the evacuated material. This notion corresponds to Hack's (1960) **dynamic equilibrium**. This morphogenetic balance depends on the inclination of the slope, the nature of the rock, and climate (Tricart, 1957). If the gradient is significant, more detritus will be exported and the slope gradient will increase. Coherent rocks require prior weathering (perpendicular component). Climate intervenes directly or as a result of vegetation and affects both components. In addition, at the present, "we find ourselves in an unprecedented situation in the four thousand five hundred million years of life on the planet, that for the first time there is a species (man) with the capacity to affect the various natural systems on a planetary level with a qualitative and quantitative importance that could equal and even exceed these natural agents" (Cendrero, 2003). We need to understand the processes of slope development for many of man's activities including agriculture, forest practices, construction of linear projects (highways, channels, and railroads), and construction of buildings; geomorphologists and engineers do this fundamental work (Dunne and Leopold, 1978; Clark and Small, 1982). Slopes are composed of two types of materials: rock and soil, which have different mechanical properties and evolve in different ways. Rocks are consolidated and coherent materials that are affected by planes of weakness (stratification, schistosity, and joints) that weaken the rock and favor fragmentation and disaggregation. Soils (regolith, alterite, surface deposits) are unconsolidated materials with low mechanical strength and high porosity. Mass movements are processes that move materials by gravity; they can be dangerous and even disastrous when they affect human life and property. They are of concern to geotechnical and geologic engineers (Bromhead, 1986). Geomechanics and related disciplines provide an understanding of landslides from an "internal" perspective whereas geomorphology contributes an "external" perspective (Crozier, 1986). Landslides pose a risk when they threaten human life and property. This risk has increased during the last century due to population increase and the increased area used for resource extraction; these factors have sometimes caused people to use areas of potential risk (Crozier, 1986). In addition, other activities such as emigration toward large urban areas and the construction of numerous engineering works that modify the landscape and cause slope instability generate landslides that can cause injuries and fatalities, property loss, and environmental damage. Landslides pose significant danger in tropical areas because of the presence of a thick regolith and high precipitation in these climates. These conditions favor surface and deep mobilization of the altered layer, which occurs most often with storm rainfall. For this reason, it is necessary to conduct detailed geomorphic cartography and install reconnaissance boreholes (Tricart, 1974b). According to the calculations of Ayala (1994), since the year 1000 about 300,000 landslides throughout the world have each killed at least 100 people. In Spain the most significant disaster took place in Azagra (Navarre) where a landslide on a gypsum cliff killed more than 100 (Faci *et al*., 1988). In Spain landslides cause losses on the order of 180 million euros and several deaths every year (Suarez and Regueiro, 1997). Cendrero *et al*. (1997) consider mass movements in Spain to be the third largest geomorphologic risk; they estimate losses for the period of 1986–2016 to be 7.6 billion dollars (Ayala *et al*., 1987a). The San Francisco Bay, California, geologic map showed one landslide at the beginning of the 20th century. In 1970, the map showed 1200 landslides; by 1980 the number had risen to 70,000 mapped landslides. This is due in part to the strong pressure that humans place on the environment (Brabb, 1989). In the United States, annual damages from landslides were estimated at 1.5 billion dollars in 1985 (Schuster and Fleming, 1996) and Bromhead (1986) puts the number of deaths at about 25 people per year.

Slope shapes

A slope shape that shows the dip of the slope is a **slope profile** (Douglas, 1977). Slope angles and distances can be measured on a section perpendicular to the slope. A series of irregularities is generally observed, such as concavities and convexities that depend primarily on the lithology of the substrate and processes that affect slope morphology; these factors play a significant role in the morphoclimatic environment (Kirkby, 1976; Toy, 1977). Geomorphologists characterize a slope by analyzing its topographical profile and assigning to it one or more elementary geometrical shapes. These geometric shapes were established by Savigear (1952, 1956) and later expanded and quantified by Young (1964). Young's work combines Savigear's convex, concave, and rectilinear elements into units, segments, and elements of parts of a slope. The morphology of a profile is analyzed in this way. The techniques used are compiled in King (1966), Goudie (1981a), and Gardiner and Dackombe (1983). Another classification proposed by Dalrymple *et al*. (1968) divides the slope into nine units based on shape and pedogenic processes. Slope measurements make it possible to distinguish nine possible shapes of slope units (Ruhe, 1975). For **slope cartography**, Young (1972) distinguishes morphological maps, slope angle maps, and genetic morphological maps. Morphological maps are drawn at a detailed scale to determine the surface units of the slope. This technique, which has been developed by numerous authors, is based on the assumption that relief can be divided into uniform morphological units bordered by morphological discontinuities (Curtis *et al*., 1965). There are two types of morphological slope angle maps: isoclinal maps that show the slope angle at a point; they are generally made from contour maps (Gregory and Brown, 1966). The other type, called maps of average slope, uses contours to indicate average slope angles over distances of a kilometer. The third type, genetic geomorphic maps, are most commonly used in geomorphic cartography. Landforms are shown according to their possible origin, taking into account present and past processes and the age of the landform. Numerous legends have been proposed for maps at different scales including those of Tricart (1965, 1972), Verstappen and van Zuidam (1968, 1991), St-Onge (1968), Demek (1972), Panizza (1972), Demek and Embleton (1978), Barsch and Liedke (1985), and Pena (1997). These legends generally follow morphological differences present in various countries; for this reason geomorphologists develop systems or legends appropriate for the needs of particular countries (Parsons, 1988). Structural shapes predominate in the preparation of small-scale geomorphic maps whereas climatic landforms are used in large-scale geomorphic mapping (Guttierez, 2013).

UNIT-9: CONCEPT OF BASIN HYDROLOGY AND RUN OFF CYCLE; UNIT HYDROGRAPH, RATING CURVE AND THEIR APPLICATIONS

The flow in river channels exerts hydraulic forces on the boundary (bed and banks). An important balance exists between the erosive force of the flow (driving force) and the resistance of the boundary to erosion (resisting force). This determines the ability of a river to adjust and modify the morphology of its channel. One of the main factors influencing the erosive power of a given flow is its discharge: the volume of flow passing through a given cross-section in a given time. Discharge varies both spatially and temporally in natural river channels, changing in a downstream direction and fluctuating over time in response to inputs of precipitation. Characteristics of the flow regime of a river include seasonal variations in discharge, the size and frequency of floods and frequency and duration of droughts. The characteristics of the flow regime are determined not only by the climate but also by the physical and land use characteristics of the drainage basin. In this chapter you will learn about:

- The pathways taken by water as it travels to the channel network
- How the flow in a river responds to inputs of precipitation
- Seasonal variations in flow that characterise different climatic zones
- The size (magnitude) and frequency characteristics of floods
- Flows that are significant in shaping the channel

FLOW GENERATION

Hydrological pathways

Inputs of water to a drainage basin are the various forms of **precipitation** that fall over its area. These include rain, snow, sleet, hail and dew. If you look at a drainage basin on a map, you will see that river channels cover only a very small part of the total area. This means that most of the water reaching the ground surface must find its way from the hillslopes and into the channel network. A number of pathways are involved, and a given 'parcel' of water arriving at a stream channel may have taken any number of them. Not all the incoming precipitation actually makes it to the basin outlet, since a certain percentage is evaporated back to the atmosphere. Water that falls on the leaves of vegetation and artificial structures like buildings is **intercepted**. Some of this water does fall to the ground, as anyone who has sheltered under a tree will know, but much of it is evaporated (water resource managers refer to this output as 'interception loss'). The amount of interception that occurs is dependent on such factors as the extent and type of vegetation (leaf size, structure, density and the arrangement of foliage), wind speed and rainfall intensity (Jones, 1997). Evaporation also takes place from the surface layers of the soil, and from lakes and wetland areas. A related process is **transpiration**, whereby plants take up water through their roots and evaporate it through pores, called stomata, on the underside of their leaves. This means that water can be 'lost' to the atmosphere from some depth below the surface. Because evaporation and transpiration are difficult to monitor separately, they are usually considered together as the combined process of **evapotranspiration**. Water reaching the soil surface may either enter the soil by a process called **infiltration** or remain at the surface, moving down-slope as **overland flow.** Infiltrated water either travels through the soil, parallel to the surface, as **throughflow** (2), or slowly percolates downwards to the saturated zone, travelling as **groundwater flow** (3). Rates of water movement are fastest for overland flow (typically between 50 m and 500 m per hour), between 0.005 m and 0.3 m per hour for throughflow and 0.005 m to 1.5 m per *day* for groundwater flow (Ward and Robinson, 1990). How fast a river rises after rainfall is very much dependent on the relative proportion of water taking faster (surface and near-surface) pathways, and slower (subsurface) pathways. An important control on this is the rate at which water infiltrates into the soil, which is determined by the **infiltration capacity**. This is defined as the maximum rate at which water enters the soil when it is in a given condition (Horton, 1933). This is an important definition: if the rainfall intensity exceeds the infiltration capacity, the soil cannot absorb all the water and there is excess water, which ponds at the surface. This moves down slope as overland flow, more specifically, **Hortonian overland flow** (1). As Horton's definition implies, soil properties control the infiltration capacity; these include soil permeability, the presence of vegetation and plant roots and how much water is already in the soil. Following a dry period, the infiltration capacity is highest at the start of a storm, rapidly decreasing as rainfall continues, until a constant infiltration rate is reached (Horton, 1933). A second type of overland flow, **saturation overland flow** (4), is generated when the soil is totally saturated and therefore cannot take any more water (Kirkby and Chorley, 1967). Where the water table is relatively shallow, for example in valley bottoms, rainfall can cause it to rise to the ground surface. Where this occurs, a saturated area forms around the channel, increasing in extent as the storm progresses. The saturated area acts as an extension to the channel network, meaning that a significant volume of water is transferred in a short time. These saturated areas are known as **variable source areas** or **dynamic contributing areas** and are highly significant in humid environments, where this is the main way in which storm run-off is generated (Hewlett and Hibbert, 1967; Dunne and Black, 1970). Hortonian overland flow is rarely observed in humid environments, unless the surface has a low infiltration rate, for example where there are outcrops of bare rock or artificially paved surfaces. However, in dryland environments, a combination of factors mean that Hortonian overland flow is the dominant mechanism. Rainfall, when it occurs, typically has a high intensity and exceeds the low infiltration capacity of sparsely vegetated soils (Dunne, 1978). In addition, dryland soils often develop a crust at the surface, which further reduces the infiltration capacity. Water that infiltrates the soil may travel at various depths below the surface. While it is fairly obvious that water should move downwards under the force of gravity, it might seem counter-intuitive that it also flows through the soil as throughflow. This happens because a preferential flow path is set up. Soil permeability decreases with depth, meaning that the downward movement of water is slowed. During rainfall, this leads to a backing up effect in the more permeable surface layers. This has been likened to the flow of water down a thatched roof: it is easier for the water to move parallel to the slope of the roof, along the stems of the straw, than to move vertically downwards through it (Ward, 1984; Zaslavsky and Sinai, 1981). Where they exist, **soil pipes** provide a very rapid throughflow mechanism, and rates of flow can be comparable to those in surface channels. Soil pipes are hydraulically formed conduits that can be up to a metre or more in diameter. They are found in a wide range of environments and are sometimes several hundreds of metres in length (Jones, 1997). As such, they can act as an extension to the channel network, allowing the drainage basin to respond rapidly to precipitation inputs (Jones, 1979). Even during rainless periods, flow will be maintained, as long as the level of the water table does not fall below that of the channel bed. A rather different situation exists for many dryland channels, where the water table may be several metres, or tens of metres, below the surface. In this case, the direction of flow is reversed, as water is lost through the bed and banks of the channel, percolating downwards to recharge the aquifer. This is termed a **losing stream** and, although not exclusive to drylands, many examples are found in these environments. It is not unusual for a river to be gaining and losing flow to groundwater along different parts of its course, while seasonal fluctuations in water table levels can mean that losing streams become gaining streams for part of the year. The Euphrates in Iraq provides a good example, with most of the flow being generated in the headwaters in northern Iraq, Turkey and Syria. Further downstream, at the Hit gauging station (150 km west of Baghdad) the river loses flow to groundwater for much of the year (Wilson, 1990). The loss of flow from a channel, due to downward percolation and high evaporation rates, is referred to as **transmission loss**.

The storm hydrograph and drainage basin response

Flow discharge (also known as ***Q***) is the volume of water passing through a given channel cross-section in a given time. The units of discharge most commonly used are cubic metres per second (m3 s−1), known as ‘cumecs’, although for very small flows litres per second may be used. In the United States cubic feet per second, or ‘cusecs’, are used instead of cumecs. During a particular rainfall event, there is a delay between the onset of rainfall and the time at which the discharge starts to increase. The initial increase is due to water falling directly into the channel and close to it, though as the storm progresses, water travelling from greater distances reaches the channel. Water taking the fastest pathways – overland flow, shallow throughflow and pipe flow – is rapidly transferred to the channel, contributing to the **quick flow** component of the storm hydrograph. **Base flow** contributions come from water taking the slower subsurface routes, taking much longer to reach the channel. This means that water continues to enter the channel as base flow for some time after rainfall has ceased, keeping rivers flowing during dry periods. Glaciers, lakes, reservoirs and wetlands also contribute to base flow. The relative proportions of quick flow and base flow determine the size of the hydrograph peak and the time delay, or **lag time**, between peak rainfall and peak flow. For instance, where quick flow dominates, lag times are relatively short and peak flows relatively high. Rivers dominated by base flow respond more slowly and the peak flow is lower. This is because most of the precipitation falling over basin 2 infiltrates and takes slower subsurface pathways to the stream channel. There is also a marked difference in the summer low flows, with the greater base flow component for basin 2, which sustains higher summer flows.

ANNUAL FLOW REGIMES

The **annual flow regime** of a river describes the seasonal variations in flow that are observed during an ‘average’ year. As you might expect, this is influenced by the seasonal distribution of rainfall, and the balance between rainfall and evaporation at different times of year. For example, some tropical rivers experience a marked wet and dry season, drying up completely for part of the year and carrying high flows during the wet season. Climate also has an important influence on the type and density of vegetation, soils and land use, all of which act as controls on the processes of runoff generation. Several climate characteristics are important in determining the flow regime. These include whether it is humid or arid, if it is predominantly warm or cold, the annual range of temperatures, and whether precipitation is seasonal or occurs all year round. At high latitudes and in some mountain environments, the timing and length of glacial ablation and snowmelt is a dominant factor. Figure 3.3 shows a selection of typical flow regimes, which characterise different climatic zones. These come from a classification scheme developed by Beckinsale (1969) from an existing climate classification. The different regimes are categorized using a system of two letters. The first letter relates to the mean annual precipitation and annual temperature range: A: Warm, moist tropical climates, where the mean temperature exceeds 18°C for all months of the year. B: Dry climates, where rates of potential evaporation1 exceed annual precipitation. C: Warm moist temperate climates. D: Seasonally cold climates with snowfall, where the mean temperature is less than –3°C during the coldest month. The second letter indicates the seasonal distribution of precipitation: F: Appreciable rainfall all year round. W: Marked winter low flow. S: Marked summer low flow. For example, the regime of the Pendari River is influenced by a tropical climate with a marked winter low flow, and would be classified as AW.

Soils and geology

● Soil type and thickness. Soil texture (relative proportion of sand, silt and clay particles) affects infiltration rates. Sandy soils have high permeability whereas clay soils do not. In arid areas a crust can form on the soil surface, decreasing the permeability. Soil thickness affects how much water the soil can absorb.

● Geology. Drainage basins underlain by a permeable geology tend to have a slower response to precipitation, although the flow is sustained for a longer time during dry periods. Drainage basins underlain by impermeable materials have a faster, or more 'flashy' response.

Vegetation and land use

● Vegetation type and density. Vegetation reduces the impact of raindrops and allows a more 'open' soil structure, meaning that infiltration rates are higher. Vegetation also affects interception rates and evapotranspiration losses from the basin.

● Urban areas. Depends on the proportion of the drainage basin that is urbanised. Large areas of paved surfaces, drains and culverts rapidly transmit water to river channels, leading to an increase in peak flow and a shorter lag time.

● Grazing and cultivation. When deforestation occurs, rates of overland flow tend to increase. Heavy machinery and trampling by animals compact the soil, reducing permeability, although ploughing can increase infiltration rates. Flow may be concentrated in plough furrows that run up and down the slope.

● Land drainage. The installation of field drains allows rapid transfer of runoff into the nearest stream channel.

Physiographic characteristics

● Drainage basin size and shape. In larger basins the travel times are longer, as flow has to travel greater distances to reach the outlet. The total volume of runoff increases with the drainage area. Elongated drainage basins have a response that is initially more rapid but with a lower, more gentle peak.

● Drainage density. Where the density of stream channels is high, the average distance over which water has to travel to reach the channel network is reduced, leading to a more rapid response.

● Drainage basin topography. Travel times are increased over steep slopes. In upland areas, steep slopes are often associated with thin soils and the response tends to be flashy. Rainfall may be affected by altitude and aspect with respect to storm tracks.

Channel characteristics

● Channel and floodplain resistance. The velocity of flow in river channels is affected by the roughness of the bed and banks and the shape of the channel. Overbank flows are slowed by the roughness of the floodplain surface.

● Floodplain storage. When channel capacity is exceeded, water spills out onto the surrounding floodplain, where it is stored until the floodwaters recede. If floodplain storage is limited, a greater volume of water travels downstream.

● Conveyance losses. In dryland environments the channel may lose flow due to high rates of evaporation and 'leakage' by exfiltration through the channel boundary.

Meteorological factors

● Antecedent conditions. The conditions in the drainage basin prior to the onset of precipitation. Where recent or prolonged previous rainfall has occurred, the soil may be near saturation, meaning that a relatively small input of rainfall could lead to a rapid runoff response. Where snow is lying on the ground, subsequent rainfall can cause it to melt, which may lead to flooding downstream.

● Rainfall intensity. Rainfall intensity is expressed in millimetres per hour (mm h^{-1}). The more intense the rainfall, the more likely it is that the infiltration capacity of the soil will be exceeded.

● Rainfall duration. This is the period of time over which a given rainfall event takes place. As the storm progresses, runoff contributing areas at greater and greater distances from the channel network become active. The channel network may also extend upstream as normally dry channels start to carry flow.

Downstream variations in discharge

As well as varying through time, discharge also changes along the course of a river. At any location, channel form is dependent on the discharge and supply of sediment from upstream. In most cases, discharge increases downstream as the area of the drainage basin increases and tributaries join the main channel. There is also a general increase in the size of the channel, with discharge acting as a control on the gross dimensions (Knighton, 1998). The quantitative description and understanding of the nature of these downstream changes have been the focus of much research and is explored further in Chapter 8. Although there is a *general* downstream increase in channel dimensions, local influences lead to considerable variation, even over short distances. Downstream changes in dryland channels can be very pronounced (Tooth, 2000). For example, infrequent flooding occurs along the ephemeral streams draining the Barrier Range in arid western New South Wales, Australia. Away from the uplands, high transmission losses lead to a rapid downstream reduction in discharge and channel size (Dunkerley, 1992). Downstream reductions in discharge and cross-sectional area are also observed in the piedmont and lowland zones of rivers draining the northern plains of arid central Australia (Tooth, 2000). There is an eventual termination of channel flow and bedload transport. However, during large floods, flows continue out across extensive unchannelled surfaces called 'floodouts' (Tooth, 2000). Tooth highlights the complex interactions between discharge, sediment transport, channel slope, tributary inflows, bank sediments and vegetation. These give rise to considerable variations in the downstream channel changes observed for dryland rivers.

FLOODS

Although the flow regime shows seasonal variations in river flow, it does not provide detailed information on the magnitude (size) and frequency of floods and droughts. Floods are of most interest here because they are capable of carrying out large amounts of geomorphological work and are thus significant in shaping the channel. The term 'flood' is hard to define. In general terms, a flood is a relatively high flow that exceeds the capacity of the channel. While more frequent flows are confined within the channel, periodic high flows overtop the banks and spill out onto the surrounding floodplain. Significant here is the **bankfull discharge** (Qb), defined as 'that discharge at which the channel is completely full' (Knighton, 1998). Although these definitions may sound straightforward enough, it is actually quite difficult to define bankfull discharge in the field because the height of the banks varies, even over short distances. This means that overtopping of the banks does not occur simultaneously at all points along the channel. Floodplain relief can be quite variable, with variations of between 1.7 m and 3.3 m observed on three Welsh floodplains (Lewin and Manton, 1975). Along the Alabama River, United States, flooding has been observed to occur more frequently at the apexes of actively migrating meander bends. This is associated with the development of floodplain features called levees. These are raised ridges that form along the banks when material is deposited during overbank flows. Levee development is impeded at actively migrating bends because the deposits are eroded as the channel migrates. Levees are better developed (higher) along less actively migrating sections of channel, where flooding occurs less frequently (Harvey and Schumm, 1994).

Flood magnitude and frequency

Floods of different sizes are defined in terms of high water levels or discharges that exceed certain arbitrary limits. The height of the water level in a river is called its **stage**. For a given river, there is a relationship between the size of a flood (in terms of its maximum stage or discharge) and the frequency with which it occurs. Floods of different sizes do not occur with the same regularity: large floods are rarer than smaller floods. In other words, the larger the flood, the less often it can be expected to occur. Floods are therefore defined in terms of their magnitude (size) and frequency (how often a flood of a given size can be expected to occur). You have probably heard reference to the 'twenty-year flood' or the '100-year flood'. This **return period** is an estimate of how often a flood of a given size can be expected to occur and, since less frequent floods are more extreme, the 100-year event would be bigger than the twenty-year flood. The return period (T) can also be expressed as a probability (P) by taking the inverse of the return period, i.e.: Using this, the probability of a 100-year flood taking place in any one year can be calculated as 0.01 (i.e. 1 per cent), and for the twenty-year flood, 0.05 (5 per cent). The probability that a flood with a particular return period will occur is the same every year and does not depend how long it was since a flood of this size last occurred – the twenty-year does not occur like clockwork every twenty years. However, if a period of several years is considered, the likelihood of a given flood occurring during this time increases. For example, if someone bought a house on the 100 year floodplain and lived there for thirty years, the probability of that property being flooded in any one year would be 0.01. This increases to 0.3 (probability × number of years), or 30 per cent, for the thirty-year period.

The frequency of bankfull discharge

Although bedrock channels are mainly influenced by high magnitude flows, those formed in alluvium can be adjusted by a much greater range of flows. This is reflected by the morphology and size of alluvial channels. Over the years, much research has focused on the bankfull discharge (defined above), since it represents a distinct morphological discontinuity between in-bank and overbank flows. Leopold and Wolman (1957) suggested that the channel cross-section is adjusted to accommodate a discharge that recurs with a certain return period. From an examination of active floodplain rivers, they found that the bankfull discharge had a return period of between one and two years. This is corroborated by later observations made for stable alluvial rivers (for example, Andrews, 1980; Carling, 1988). However, the concept of a universal return period for bankfull discharge that can be applied to all rivers is controversial. Williams (1978) observed wide variations in the frequency of bankfull discharge, which ranged from 1.01 to 32 years, and concluded that this was too variable to assume a uniform return period for all rivers. Even along the same river, there can be marked variations in the frequency of bankfull discharge (Pickup and Warner, 1976). The concept of a uniform frequency for bankfull discharge assumes that all channels are '**in regime**'. This means that the morphological characteristics of a given channel, such as size, fluctuate around a mean condition over the time scale considered (Pickup and Reiger, 1979). This is not true for all rivers and there are many examples of **non-regime**, or **disequilibrium**, channels. An example would be where channel incision is taking place through erosion of the channel bed. This results in a deeper channel, which requires a larger, and therefore less frequent, discharge to fill it. The Gila River in Arizona, United States, was greatly enlarged when past events had led to large floods. The enlarged channel is not adjusted to the contemporary flow regime, which means that the bankfull discharge for the enlarged channel has a much lower frequency (Stevens *et al*., 1975). The material forming the bed and banks is also significant. In cases where the boundary is very erodible, the bankfull discharge may simply reflect the most recent flood event (Pickup and Warner, 1976).

The geomorphological effectivenes of floods

Given that many rivers exceed their channel capacity and flood on a fairly regular basis, it would not be unreasonable to ask why they do not shape channels that are large enough to convey all the flows supplied to them. While it is true that high-magnitude events lead to significant changes in channel morphology, the comparative rarity of these large floods must also be taken into account. The cumulative effect of smaller, more frequent floods can also be significant in shaping the channel. The effectiveness of any given discharge over a period of time is therefore something of a compromise between its size and how often it occurs. The basic question is: are a number of smaller floods as effective as one large flood?

Regional flood frequency curves

The flood frequency–magnitude relationship differs between regions. Despite the low annual rainfall in dryland environments, precipitation can be highly variable and the twelve largest floods ever recorded in the United States all occurred in semi-arid or arid areas (Costa, 1987). During flash floods, such as the one shown in Colour Plate 14, floodwaters rapidly inundate the dry channel. Not all dryland rivers are prone to flash flooding however, and there is considerable variation in the size, type and duration of flooding. Regional flood frequency curves are shown in Figure 3.4. The return period is plotted on the horizontal axis using a logarithmic scale, with the relative flood magnitude on the vertical axis. A *relative* flood magnitude has been used to allow comparison between floods for a number of rivers in different regions. Because these all drain different areas, a direct comparison of flood magnitudes would not be very meaningful. Instead, for each river included in the analysis, the ratio between the magnitude of each flood on record and a low magnitude 'reference flow' – the mean annual flood – has been used.

Reconstructing past floods

Palaeoflood hydrology is a new and developing area of hydrology and geomorphology, which reconstructs past flood events in order to extend the flow record. Due to problems associated with monitoring major floods and the relatively short duration of most gauged records, extreme floods are very rare in the observational record. By reconstructing palaeofloods, the flood record can be extended, allowing increased accuracy in the estimation of floods for risk analysis. Evidence of past flood events is provided by geological indicators such as flood deposits, silt lines and erosion lines along the channel and valley walls (Benito *et al*., 2004). Historical records are also used and include documents, chronicles and flood marks inscribed on bridges and buildings. Using this evidence, it is possible to determine the size of the largest flood events over periods of time ranging from decades to thousands of years (Benito *et al*., 2004). As well as identifying the largest floods, evidence of floods above or below above specified flow stages can also be reconstructed (Stedinger and Baker, 1987). Although time-consuming, it is possible to reconstruct a complete record, chronicling the largest flood, together with the size and number of intermediate palaeofloods (Benito *et al*., 2004).

CHAPTER SUMMARY

Inputs of precipitation falling over the area of a drainage basin are transferred to the channel via a number of different pathways. These include surface overland flow, throughflow (through the soil) and deeper groundwater flow. Rates of movement vary considerably: overland flow and shallow throughflow are generally much more rapid than groundwater flow. An important control is the infiltration capacity, which determines how quickly water can be absorbed by the soil. If the rainfall intensity is greater than the infiltration capacity, excess water builds up at the surface, leading to overland flow. Overland flow also occurs when rain falls on saturated areas. The volume of water flowing through a given cross-section in a given time (discharge) fluctuates in response to inputs of precipitation. A hydrograph shows how discharge changes over the course of a year (annual hydrograph), or one rainfall event (storm hydrograph). The shape of the hydrograph is affected by the physical, land use and climatic characteristics of the drainage basin. These variables all determine the relative proportion of water taking faster and slower pathways to reach the channel. Climate is a very important control on the annual flow regime of a river, which reflects the precipitation amount, seasonal distribution and annual temperature variations. Another important characteristic of the flow regime is the frequency and magnitude (size) of flood events. As the size of a flood increases, the frequency with which it occurs (return period) decreases. The relationship between frequency and magnitude differs from region to region. In dryland environments, large, low frequency floods are much more extreme than those with a similar return period in humid areas. The bankfull discharge is that flow at which the channel is completely filled. Wide variations are seen in the frequency with which the bankfull discharge occurs, although it *generally* has a return period of one to two years for many stable alluvial rivers. The geomorphological work carried out by a given flow depends not only on its size but also on its frequency of occurrence over a given period of time (Charlton, 2008).

UNIT-10: STORM WATER AND FLOOD MANAGEMENT: STORM WATER MANAGEMENT, DESIGN OF DRAINAGE SYSTEM, FLOOD ROUTING THROUGH CHANNELS AND RESERVOIR, FLOOD CONTROL AND RESERVOIR OPERATION

Industrial, agricultural and urban development put ever-growing pressures on land and water resources. Traditional management priorities have met human needs, with little regard for ecosystems. On a global scale the resulting impacts on the physical, chemical and ecological condition of river systems have been truly profound. There are many examples of severe environmental degradation, and few rivers can be described as being in a pristine or near-pristine condition. However, over the last part of the twentieth century, growing environmental awareness led to a shift in management priorities. This was coupled with developments in the understanding of river behaviour, dynamics and change. The role of the river manager now includes new challenges, where environmental considerations must be integrated with the development of water resources and the management of hazards, such as flooding. River management is now multidisciplinary, involving experts from a number of different fields, including geomorphology, ecology, engineering and economics. Geomorphologists have an important role to play in assessing the condition of rivers and their catchments, their sensitivity and future response to change. In this chapter you will learn about:

- The ways in which river channels have traditionally been managed to meet human needs
- The adverse impacts of past land and water management on the physical, chemical and ecological condition of channels and floodplains
- Changes in management philosophy towards more environmental approaches
  Environmentally sensitive management techniques
- Restoration of degraded channel reaches.

TRADITIONAL ENGINEERING TECHNIQUES

Why rivers are engineered

Human settlements have long been located along river channels, which provide a supply of water and power, fertile floodplain soils, fisheries and a potential means of navigation. Rivers can also be hazardous and many urban areas are increasingly at risk from flooding as they expand onto floodplains. This risk is further increased by larger flood peaks associated with land use change within the drainage basin – upstream deforestation, land drainage and urban development can all significantly increase flood peaks further downstream. Flood control works involve artificially increasing the channel cross-section, constructing flood embankments, straightening channels and removing vegetation and other obstacles. More recent techniques include the construction of flood diversion channels and flood storage reservoirs. Many lowland rivers are maintained for navigation; these include the Rhine, Danube, Mississippi, Missouri, Ohio and Arkansas. The aim is to maintain a minimum depth of water along the navigable length of the river by means of dredging, removal of shoals and other obstacles and river training (see below). Weirs and locks are also used to extend the navigable length, providing a minimum depth for larger vessels at all times, despite natural variations in discharge. Increased industrialisation and urbanisation place growing demands on water supply systems. The world's largest cities consume water at a rate that is exceeded only by the flow of a few major rivers. Meeting these demands involves constructing dams to store and regulate flow. It is also necessary to integrate supply from a number of different surface and subsurface sources, which can involve transferring water over large distances in river channels, pipelines and canals. On a global scale, the largest demand comes from irrigated agriculture in arid and semi-arid environments.

Advances in irrigation technology allowed a huge expansion in the total irrigated area during the Green Revolution of the 1960s. This was also the decade that saw the construction of the greatest number of dams, the scale of which had been increasing since the construction of the Hoover Dam in 1935. By 1986 there were 39,000 large dams over 15 m in height (International Commission on Large Dams (ICOLD), 1988) and today few of the world's major rivers are unregulated. In the second part of the twentieth century there was a growing trend towards multi-purpose dams, whose roles include water supply, flood control and hydroelectric power generation. Channel modifications are often involved in land reclamation and the drainage of wetlands and low lying areas. Many channels in lowland areas have been deepened and straightened to convey the increased volume of water resulting from the installation of field drains. Local modifications are also made to channels where channel instability might cause problems, for example, to prevent bank erosion at the site of bridges and other structures.

Channelisation and flow regulation

Channelisation is the modification of natural river channels for the purposes of navigation, flood control, land drainage and erosion control (Brookes, 1988). Re-sectioning and realignment Re-sectioning describes the modification of the channel cross-section to provide adequate depth for navigation and to increase the channel capacity for land drainage and flood control. This may involve the removal of a few bars or shoals, or deepening all, or part, of the crosssection. Some channels may be enlarged further by widening. Depending on the size of the channel and the purpose of the engineering works, re-sectioning is carried out using dredging, or by means of river training. **Realignment** involves the straightening of river channels for purposes of navigation and flood control. It is also carried out where channels share the valley with roads and railways, to reduce the number of bridges that have to be constructed. In navigable rivers with a high natural sinuosity, the removal of meanders greatly reduces the distance that has to be travelled by vessels moving up and downstream. However, the increase in gradient can lead to instability as a result of increased stream power in the straightened section. Increased erosion in this section can lead to problems of deposition further downstream. Dredging Dredging is the removal of sediment from the bed of the channel for flood control and to maintain or deepen existing navigation channels. It is also carried out when sand and gravel are mined from the river bed. Dredging has been carried out for thousands of years and was practiced by the Egyptians, Sumerians, Chinese and Romans, by means of mass labour and manual tools (Petersen, 1986). Over time, developments in dredging technology and mechanisation have enabled these operations to be carried out at increasing scales. In smaller, non-navigable channels, dredging is carried out from the bank using a dragline or bulldozer. Bank-side vegetation is often removed to enable access. In larger channels, the dredger is mounted on a floating platform. Mechanical dredgers remove material by lifting it from the bed in a bucket or dipper, whereas hydraulic dredgers use suction pumps to remove material, via a pipeline, to a disposal site. Rotating cutter heads and explosives are used to remove resistant bed sediment and bedrock outcrops. Dredging for channel maintenance has to be carried out on a regular basis, at considerable cost. This is because it treats the problem (sediment accumulation) rather than its causes (sediment sources). Excessive mining of sediment from the river bed can lead to serious problems of erosion, both upstream and downstream from the site. Snagging and clearing Flow resistance in the channel is increased by woody debris (fallen trees, logs, branches), large rocks and urban debris. In-channel debris presents additional hazards to navigation and may threaten bridges and other structures. The purpose of snagging and clearing is to remove this material, and these operations are usually part of the routine maintenance carried out every few years on many engineered channels. Trees and bushes at the edge of the channel may also be cleared at the same time. This is because bank-side vegetation can increase resistance to high flows, reducing velocity and potentially increasing the flood risk. Other reasons for removing vegetation are to allow access and to reduce the amount of woody debris entering the channel. As with dredging, snagging has to be carried out on an ongoing basis. Levees and embankments Levees are artificial embankments which are built alongside or close to the channel margins of lowland rivers. Their purpose is to increase the channel capacity at high flows and protect the surrounding floodplain from inundation. Levees are found extensively along many major rivers, including the Nile and Mississippi. Traditionally, levees have been constructed of earth, and many still are. In urban areas, where the potential human and economic losses are greater, levees and floodwalls are usually made of concrete. It is not feasible in economic or practical terms to construct levees that would contain all the floods that could possibly occur. Levees are therefore built to withstand a certain **design flood**, such as the twenty-year event. If this flow is exceeded the levees will be overtopped. The depth of flow contained within the levees is greater than it would be if no levees were present and water was able to inundate the floodplain. Since shear stress increases with flow depth, increased erosion of the channel bed is possible. Bank protection Banks are protected using various types of revetments and resistant lining materials. **Revetments** provide armouring, in the form of loose rocks and boulders, or container systems, such as wire baskets filled with rock (gabions). Banks can also be lined with concrete, asphalt, paving slabs or, where the engineered crosssection is rectangular, using vertical sheet steel piling. Plate 10.1 shows a heavily engineered urban channel which has been lined with concrete to prevent erosion. Lining a channel with concrete can lower the resistance to flow, possibly leading to problems of scour further downstream. For larger rivers, banks can be protected by laying down mattresses of concrete slabs connected by tough steel cables. This method has been extensively used along the lower Mississippi river, United States. **Spur dykes** or **groynes** can be used to protect banks by deflecting flow away from vulnerable zones. They are built at an angle from the bank and are constructed from various materials, including stone, boulders, earth, gabions or pile. Thousands of stone spur dykes have been constructed along the lower Mississippi. Bed protection Armouring is often used to protect the channel bed from erosion. In-channel **grade control structures** can also be installed. These are of two basic types: sills and weirs. **Sills** are low, submerged structures, which are built at right angles to the direction of flow. They provide local fixed points that control the channel bed slope and water surface elevation to prevent degradation and headcutting. **Weirs** act as hydraulic controls, dissipating excess energy and reducing the energy slope. Protecting the bed from erosion reduces the available load and may lead to scour further downstream by sediment-hungry water. River training River training techniques have been used since as long ago as the late sixteenth century on the Yellow River in China (Przedwojski *et al*., 1995). In Europe, aggrading glacially fed braided channels were among the first to be ‘corrected’. One of the earliest and most successful examples was the work carried out on the Alpine Rhine (Switzerland) in the early nineteenth century. Before then, the active braided channel had occupied a width of several kilometres. High rates of channel migration combined with frequent flooding meant that the floodplain could not be fully utilised. There was also a high incidence of waterborne disease, including malaria. The main idea behind the works was to ‘train’ the river to flow in a deeper channel, reducing the incidence of flooding. This was achieved by confining the flow to a straight, single channel, using embankments and groynes to encourage deposition at the channel edges and to stabilise the channel in one position. Flow was concentrated along the centre of the channel, leading to deepening and an increase in channel capacity. This allowed floodwater and sediment to be rapidly transported downstream. By 1845, 12.5 million ha of the floodplain marsh had been drained, allowing increased rates of agricultural production (Downs and Gregory, 2004). Today the area is intensively cultivated. Extensive training works have been carried out on many other rivers, including the Rhone, Danube and Mississippi. Spur dykes were constructed along the lower Mississippi to encourage deposition at the edge of the channel. This concentrates erosion of the channel bed, allowing the depth of the shipping channel to be maintained. In order to protect the opposite bank from erosion, it was necessary to install bank protection in the form of extensive concrete mattresses. Bed degradation is controlled using fixed weirs and sills. Structures can also be installed to alter flow patterns on a more localised scale.

Dam construction

Dams are constructed for power generation, flood control, and for supplying water to irrigation schemes and urban centres. Today, very few large rivers remain unregulated, and the global volume of water stored in reservoirs now exceeds the volume of flow along rivers (Brierley and Fryirs, 2005). The scale of dams varies, from relatively small structures on tributaries to large dams that exceed 15 m in height. Gregory (1995) estimates that over 200 large dams are completed each year. However, there has been a decline in the rate of dam building in the industrialised world because many potential dam sites have now been developed. Flow regulation dramatically alters the flow and sediment regimes. Flood peaks are reduced in magnitude and most of the sediment load is trapped in the reservoir behind the dam. Downstream from hydroelectric power stations there are often rapid fluctuations in discharge. Inter-basin transfers involve the movement of water across drainage divides, with the result that there is a net gain to some river systems and a net loss from others.

Emplacement of locks and weirs

The Rhine, Mississippi and Arkansas rivers have all been canalised to ensure a minimum depth of water for shipping and to increase the navigable channel further upstream. Canalisation involves installing dams across the channel to create a series of slackwater pools. Sluice gates, weirs and other control structures regulate the flow of water and vessels pass up and downstream through locks or, occasionally, ship lifts.

ECOLOGICAL REQUIREMENTS

In common with the geomorphological system, ecological systems are characterised by flows of energy and materials. They are also highly dynamic, being strongly conditioned by the hydrological and geomorphological environment and biological processes such as predation, competition, dispersal, migration, colonisation, ecological succession and extinction. **Biological communities** The biological communities that live in rivers, riparian zones and floodplains include many different types of plants and animals. Microbial organisms are important in primary production – the growth of plant material as chemical energy is stored through the process of photosynthesis – and in the breakdown of organic material and recycling of nutrients. The 'slime' that forms on submerged boulders and pebbles in river channels is a complex mixture of algae, bacteria, fungi and fine particles of organic and inorganic material (Closs *et al*., 2004). Many plant species are anchored by their roots to the substrate, although floating plants can be found in backwater areas. Macroinvertebrates include insect larvae and nymphs, crustaceans such as freshwater shrimps, bivalve molluscs, gastropods (e.g. snails) and worms, while fish are the most notable vertebrates. Fauna that are associated with both terrestrial and freshwater environments include amphibians, reptiles and waterfowl, together with mammals such as otters, beavers or platypuses (depending on geographical location). Habitats Different biological communities are found in different **habitats**, the habitat being defined by the physical conditions. In-stream habitats are defined by channel morphology, bed and bank sediment, flow regime, sediment transport rates, in-channel and riparian vegetation, and woody debris. Many of these conditions fluctuate over short periods of time so it is the *range* of conditions that prevails, rather than an average condition that is important. For example, the physiology of many species of fish and invertebrates means that they cannot tolerate extremes of temperature. In order for a particular species to thrive in a given habitat there must be a tolerable range of conditions, a supply of food and an environment that can support all stages of its life cycle. Important wetland habitats are maintained by overbank flows and interactions between groundwater and surface water. Within a reach of channel there are many small-scale variations associated with bedforms, scour pools, gravel patches, bars and shelter behind tree roots or large boulders. Associated with these differences are variations in flow depth and velocity. This creates a mosaic of **mesohabitats**, each characterised by different assemblages of species. Many species move between the different mesohabitats to feed, reproduce, find shelter during high flows, or hide from predators. Species diversity and abundance are usually higher where a range of different mesohabitats can be found. Downstream changes Along the length of a 'typical' river, recognisable changes can be seen in the nature and structure of biological communities as physical conditions change, from shallow, turbulent upland streams to deeper lowland channels. In the headwaters, variations in depth and velocity are more extreme and only those species that are able to withstand large variations in velocity and depth are able to survive. These make use of physical adaptations, such as a means of attachment (plants and macroinvertebrates). Behavioural strategies include sheltering behind boulders or in pools (fish), or moving into the spaces in a gravel substrate (macroinvertebrates). Food resources are limited, because unstable substrates mean that few plants are able to establish themselves. The main source of food is provided by plant remains and the corpses and faeces of animals falling into the channel from adjacent slopes and riparian vegetation. Further downstream, plants are more able to establish themselves (although in deep channels with high turbidity this may only be at the edges of the channel) and biological communities become more diverse and abundant.

The fluvial hydrosystems concept

The fluvial hydrosystems concept (Petts and Amoros, 1996) recognises the important transfers of energy, materials and biota that occur within fluvial systems. These include longitudinal transfers along the channel, lateral transfers between the channel and floodplain, and vertical transfers between the channel and underlying alluvial aquifer (Figure 10.1). ● ***Longitudinal transfers*** are vital to many species that need to be able to move through the channel network. Migratory fish, such as salmon, swim downstream in their juvenile stage and live most of their lives in the ocean, before making the long journey back upstream to their spawning grounds. There are also downstream transfers of energy and biological material from the catchment area and through the channel network. These allow more productive and diverse biological communities to be sustained further downstream. ● ***Lateral transfers*** occur during overbank flows. Channels and floodplains become connected by floodwaters, enabling the lateral movements of water, sediment and biota that are vital to many species. In many fish species the timing of reproduction coincides with the timing of floods, when fish move onto the floodplain. Here they enter the sheltered environment provided by flooded hollows and abandoned channels, where spawning takes place. These environments also provide a refuge for many other aquatic species during flood flows. Periodic inundation of the floodplain creates unique wetland habitats that support rich biological communities. ● ***Vertical transfers*** between the alluvial aquifer and the overlying channel and floodplain are also important. This subsurface zone is important in supporting various groundwater fauna. Abandoned channels, which are isolated from the main channel at the surface, may still be connected periodically to the alluvial aquifer below the surface, allowing vertical transfers to take place.

ENVIRONMENTAL DEGRADATION

Human activity has led to the environmental degradation of numerous river systems. The fragmentation of river systems by dams, weirs and other structures seriously disrupts the natural functioning of physical and ecological processes. Together with declining water quality, this leads to a dramatic reduction in species diversity. The effects of human activity result in direct and indirect impacts on river systems. **Indirect impacts** are brought about by changes within the drainage basin that affect the flow and sediment regimes. Examples include deforestation, changing agricultural practices, urbanisation, mining and road construction. **Direct impacts** result from dam construction and channelisation, where modifications are made to the channel itself. **Basin-scale impacts** Deforestation, associated with agricultural development, has affected river systems for thousands of years. However, there has been an acceleration in the development of land and water resources over the last 100 to 500 years. Across much of Western Europe the intensity of agricultural production increased dramatically after the Second World War. This intensification has included changes from grazing to arable land, the clearance and cultivation of riparian zones and increases in stocking densities. This has often increased the supply of fine sediment to river channels, leading to problems of siltation. Agrochemicals such as pesticides, herbicides and fertilisers have all increased pollution from agricultural runoff. Municipal sewage, industrial effluents and urban runoff also contribute to water pollution. With increasing urbanisation and the move to a more industrialised society, the floodplains of major rivers in Europe and elsewhere are changing from agricultural to urban land use. This affects the flood hydrograph, increasing peak flows and necessitating further flood protection works. At the same time, because a greater proportion of incoming rainfall is rapidly diverted to rivers via drains, gutters and sewers, there is a reduction in groundwater recharge rates. **Impacts of dams** Changes to the flow regime One of the most profound influences is the alteration of the flow regime by dams and other flow control structures. The life histories of many species have evolved in response to natural flow regimes. As you have seen in previous chapters, flow affects all aspects of the physical habitat, including the shape and size of the channel, the spacing of riffle and pool habitats and nature of the channel substrate. Even at small scales, variations in velocity and shear stress across the channel bed affect the distribution of plants and macro-invertebrates within the channel. For most fish species, the timing of life events such as reproduction and spawning can be linked to the flow regime. The timing of rising flows are important for fish that move out onto the floodplain to spawn. Other triggers include day length and temperature. Under an altered flow regime, these may no longer be synchronised with natural variations in flow. The impacts of flow regulation on a given river system are dependent on several factors, including the number and size of dams, the distance downstream from an impoundment and the proportion of the upstream drainage basin area that is regulated. Further downstream, the effects of flow regulation may be reduced to some extent by unregulated tributaries joining the main channel. The type of operation, for example hydroelectric power generation, is also significant. One of the main impacts is the reduction in flood peaks, which play a vital role in the life cycle of many species. Irrigation schemes can even result in a reversal of the flow regime, when seasonal flood flows are impounded and later released to water crops during dry months. Downstream from hydroelectric plants the flow can vary considerably in just a few hours, as electricity is generated to meet daily fluctuations in energy demand. Such releases can have a serious impact on the temperature regime. During relatively warm conditions, water released from the base of a dam originates from the cooler depths of the reservoir. This can result in severe thermal shocks to fish and other species downstream from the dam. In July 1976 high mortalities occurred among grayling (a fish species) in the River Ain, a tributary of the Rhone, because of twice-daily releases of cold water from upstream reservoirs (Bravard and Petts, 1996). Seasonally modified temperature regimes also affect life cycle patterns and coldwater releases have been found to delay spawning by up to thirty days in some fish species (Zhong and Power, 1996). In the Western world, **inter-basin transfers** are becoming more common as most potential dam sites have now been developed. This involves diverting water from one drainage basin to another, via a pipeline or canal, to enhance flow in the receiving drainage basin for water supply or irrigation. Downstream from the abstraction point, a '**compensation flow**' is released into the channel to meet the needs of downstream users, and to provide sufficient water to dilute sewage and industrial effluents that are discharged into the channel. However, this is usually much less than the natural flow and may not be sufficient to meet environmental needs. In dryland environments, the receiving river can be transformed from an ephemeral to a perennial channel. This is the case with the Great Fish River in South Africa, which has changed from a series of unconnected pools in the dry season to a perennial river as a result of the Orange–Fish inter-basin transfer. Such transfers have major implications for existing plant and animal communities that have evolved to a highly variable flow regime. In fact, non-native 'exotic' species may thrive under the new flow regime. Regulation of flows in some Australian rivers is thought to favour nonnative carp and mosquitofish (Bunn and Arthington, 2002). Despite precautions, exotic organisms, including parasites, bacteria and fish, are often unintentionally transferred between basins via connecting pipelines. This is how Orange River fish have made their way into the Great Fish River in South Africa. Reduced connectivity Longitudinal connectivity is greatly disrupted by dams, locks and weirs. In many cases migratory pathways are blocked, even by relatively small structures. Migratory species including shad, lamphrey and eels have disappeared from the Rhone in France (Bunn and Arthington, 2002). Fish ladders are often installed at the site of dams to allow fish to bypass the dam by swimming up through a flight of pools. For various reasons these structures are not always successful, meaning that reduced numbers of fish are able to make the journey upstream. Lateral connectivity is reduced by the dramatic reduction in the frequency, extent and duration of overbank events. Bravard and Petts (1996) cite the case of the Volga where the duration of floodplain inundation has decreased, from fifty to seventy, to ten to fifteen days a year, since flow regulation began. This has had a major impact on fish populations, since a minimum of forty days is needed for the growth and development of juvenile fish before their descent to the Caspian Sea. A reduction in flood inundation also threatens wetland areas, particularly where land drainage is carried out. The Macquarie Marshes in eastern Australia, a wetland reserve for birds, has been reduced to between 40–50 per cent of its original size by flow diversions and weirs (Kingsford and Thomas, 1995). Lateral connectivity is further reduced by levees and embankments and artificially deepened channels (see below). Vertical transfers are also affected, since recharge of the underlying alluvial aquifer takes place when floodwaters inundate the floodplain. Reduced rates of aquifer recharge lead to a drop in groundwater levels, further contributing to the drying out of wetland areas. **Impacts of channelisation** Instability problems Channelisation programmes have significantly modified tens of thousands of kilometres of river channels (Brookes, 1985). These modifications often lead to instability within the engineered reach and in the reaches upstream and downstream from it. Changes to the channel slope, width, depth or roughness all affect channel hydraulics. Feedback mechanisms lead to adjustments as the channel tries to find a new equilibrium. For example, in a channel that has been enlarged for flood control, there will be a reduction in velocity and unit stream power at low flows. This will result in deposition along the reach, meaning that the channel has to be re-dredged on a regular basis to maintain its capacity. Modifications to the channel slope, made by creating artificial cutoffs, gravel mining or dredging, can have the most dramatic effects because of the resultant increase in stream power. Incision often occurs upstream from the artificially steepened reach. Erosion is concentrated at the break in slope between the gentler upstream reach and steeper engineered reach. Upstream incision then takes place as a series of headcuts migrate upstream, although the rate of incision decreases in an upstream direction. In severe cases, banks become unstable, resulting in collapse and channel widening. Bridges and other structures can also be undermined. The additional sediment that is produced causes further problems of aggradation downstream. Major instability problems resulted from the removal of sixteen meander bends along the lower Mississippi between 1929 and 1942. It shortened the channel by 220 km and led to excessive channel erosion, necessitating further intervention. Prior to the installation of bank protection on a massive scale, erosion was removing 900,000 m3 of bank material a year (Bravard and Petts, 1996). Upstream degradation has also led to extensive deposition in the engineered reach, in the form of bars. Geomorphological response times vary, being dependent on the type of work carried out, and the extent to which unit stream power, sediment supply and vegetation cover are affected. In some cases it may take up to 1,000 years for the channel to reach a new equilibrium form (Brierley and Fryirs, 2005). Ecological impacts The abundance and diversity of different species tends to be greatly reduced in engineered channels as a result of limited connectivity and habitat availability. Dredging and snagging remove geomorphic structures such as riffles, pools and bars, and disturb the structure of bed sediment. The uniformity of engineered channels provides little variety, affecting the viability of certain species. This is particularly true of concrete-lined channels, which have very little ecological value. During high flows, stream velocities may be higher than some species can withstand. Deepened channels and levees increase channel capacity, greatly concentrating high flows. Opportunities for shelter are reduced within the channel, and levees prevent access to calmer waters on the floodplain. Water temperatures can increase to intolerable levels during low flows. Enlarged channels may not provide a sufficient depth of flow, a problem that is exacerbated by the removal of pools and the shading effects of riparian vegetation. Aesthetic impacts As well as affecting the morphology and behaviour of river channels, channel engineering works often have a negative impact on the appearance and amenity value of the channel. Part of what makes natural rivers pleasing to look at is the amount of variety one can observe, even over a short distance. Variations in depth, velocity, slope and sediment size, associated with forms like riffles and pools, bends, bars, rapids, trees and other vegetation, all combine to create an interesting environment. By contrast, engineered reaches can be very monotonous in appearance, with their straight channels and uniform cross-section, cut off from the adjacent floodplain by embankments and levees.

ENVIRONMENTAL APPROACHES TO CHANNEL ENGINEERING

Changing management priorities

Environmental awareness increased substantially over the last part of the twentieth century. This was brought about by increased understanding of the environment, together with concern for loss of habitat, loss of species diversity and the implications of ecosystem decline for human health. River systems play a vital role in sustaining human health, and their degradation also has socio-economic implications for the value of river systems in commercial and amenity terms. The ‘health’ of a river can be described in terms of the quality of its habitat structure, hydrological characteristics, ecological function and water quality. Growing concern for river health is reflected in statutory requirements, where the preservation of ecosystems is a priority, together with human requirements. Central to this is a holistic approach where, at whatever scale is being considered, the functioning of the whole drainage basin is taken into consideration. Basin management involves prioritising and dealing with many different, and often conflicting, demands. These include meeting ecological requirements; managing channel instability; providing a supply of water for agriculture, urban centres, industry and power generation; flood management; navigation; recreation; and the needs of local communities. Appropriate management strategies are also dependent on the condition of the river itself, in terms of both river health and channel dynamics. On a global scale, very few rivers are in a pristine or near-pristine state. Even where channelisation has not been carried out, land use changes within the drainage basin have often resulted in altered flow and sediment regimes, as well as a decline in water quality. The first priority should be in the **preservation** of channel reaches that are in a pristine or near-pristine condition. Second, in modified reaches where there is potential to return the river to a more natural state, various channel **restoration** strategies can be implemented. The lowest priority are cases of severe environmental degradation, where the only option may be to leave the channel in a state of **dereliction** and focus resources elsewhere. **Information requirements** In order to integrate the many different priorities and work with, rather than against, the natural functioning of river channels, it is essential to start by understanding the condition that they are in. Environmental assessment is a very important aspect of channel management and many countries have legislation that requires preproject assessment to be carried out. A range of environmental assessment techniques have been developed for different purposes and for varying levels of detail. There are usually three levels of pre-project environmental assessment (Downs and Gregory, 2004). These are briefly outlined below. Level 1. Basin-scale assessment of river channel condition The first stage involves compiling basin-scale inventories of channel morphology and habitat condition along the length of the channel. This allows identification of reaches of high conservation value, the location of past channelisation works, areas of instability, the morphology of different reaches and the habitats associated with them. Assessment usually includes a stream reconnaissance survey and a baseline habitat survey. The **stream reconnaissance survey** involves combining basic measurements of channel geometry with qualitative assessments of channel conditions, including valley characteristics, land use, connectivity, and instream and riparian vegetation. Standard formats have been devised for collecting this information, such as the *Stream Reconnaissance Handbook* (Thorne 1998), although skilled interpretation is required in the field. Field observations are supplemented with information from maps and aerial photographs. **Baseline habitat surveys** provide further information on the 'health' of the river in terms of its water quality, geomorphic structure and connectivity. These are based on the hierarchical links that exist between channel processes, form, habitat and associated biota. The character of different in-stream habitats is determined by interactions between channel morphology; bed and bank sediments; the flow regime; flow hydraulics (mainly by the combinations of velocity and depth conditions); sediment transport characteristics; and riparian vegetation (Downs and Gregory, 2004). Certain groupings of biota are associated with different in-stream habitats. Surveys can be 'top down', starting from morphological surveys, or 'bottom up', where data from biotic surveys are grouped (Downs and Gregory, 2004). Top-down approaches are based on the concept of **hydraulic** or **physical biotopes** (Rowntree and Wadeson, 1996; Newson *et al*., 1998). These are habitat 'units' that are defined by combining the physical features of the channel with flow character. On this basis, hydrogeomorphological units such as 'riffle', 'rapid' and 'pool' can be distinguished. In this way, links can be made between habitats, flow types and species requirements (Newson *et al*., 1998). Level 2. Basin historical analysis River channels are dynamic, and it is vital to assess each reach in terms of the controls on channel adjustment, sediment dynamics (supply, transport and storage), whether the reach is stable or unstable and its potential sensitivity to change. In order to understand the condition of a given channel reach, it is essential to link a local site, such as a channel reach, with the wider context of the drainage basin and its history (Gilvear, 1999). Basin history can be reconstructed using flow records, aerial photographs, maps, historical records, floodplain stratigraphy and vegetation. Interpretation of this information provides insight into the ways in which changes in the flow and sediment regimes reflect changing environmental conditions within the basin. Through this understanding it is possible to determine the sensitivity of a given reach to change and to develop appropriate management strategies. Using a **fluvial audit**, the condition of the reach – channel stability, sediment movement and morphological stability – can be related to sediment dynamics in the catchment (Sear *et al*., 1995). This is constructed using information on sediment sources, pathways and characteristics. Level 3. Reach-scale analysis of the sensitivity of channel hazards and assets In order to determine the risk associated with channelrelated hazards, such as flooding or lateral erosion, detailed reach-scale analyses are carried out. This requires considerable expertise and intensive fieldwork, so is feasible only for a small number of reaches. These are carefully selected on the basis of the preceding catchment and historical assessments, and should be representative of similar reaches. Examples of the field measurements carried out include planform mapping, surveys of bed topography, measurement of velocity fields, sediment transport rates, rates of lateral erosion, and analysis of the structure, composition and dominant failure mechanisms of the banks. **Reducing the impacts of channel engineering works** Channel engineering can now be carried out using more environmentally sensitive techniques. These reduce the impact on the natural function and morphology of the channel, and promote greater ecological diversity. Distant flood banks and multi-stage channels Where space allows, a more environmentally sensitive approach is to use distant flood embankments, which are set back from the channel. These allow the river to inundate its floodplain but provide flood protection at the same time as maintaining important channel–floodplain interactions. Distant embankments do not need to be as high as those built adjacent to the channel to provide the same level of protection and the channel itself remains undisturbed. Two-stage channels are designed to accommodate a range of flows, with the low-flow channel being preserved within an excavated flood channel. The berms allow enhanced interactions between surface water and groundwater because the frequency of inundation is increased, and also because the water table is closer to the lowered surface. Problems can arise if vegetation growth on the berms remains unchecked, as this will increase the roughness of the flood channel, reducing velocity and channel capacity. Another problem can be the build up of silt on the berms. Partial dredging Rather than dredging along the entire length of the engineered reach, partial dredging reduces the impacts on in-stream ecology. This involves dredging only the central part of the cross-section at intervals along the channel. For example, the depth could be increased at shallower riffles, leaving parts of the bed undisturbed. Careful design is required to ensure that instability problems do not result. Partially dredged sections of channel can rapidly become infilled with sediment, necessitating regular maintenance. Revetments using 'soft' engineering techniques A wide range of bank protection measures can now be used. These range from 'hard' engineering structures such as rip-rap and concrete to 'softer' or 'biotechnical' engineering solutions. For situations where bank erosion is a serious threat – in urban areas and where channels run close to roads or railways – hard engineering structures are most appropriate. However, in less critical settings, soft engineering provides a more environmentally sensitive alternative. In high-energy environments the placement of a stone 'apron' protects the toe of the bank from erosion. The spaces between the interlocking stones allow plants to become established and provide habitats for instream fauna. Where the specific stream power is lower, natural plant materials can be used to protect banks by placing a line of stakes along the bank and interweaving them with branches. Fast-growing species such as willow may be used as stakes to create a living barrier. Geotextiles include materials such as jute, coconut fibre and brush mattresses (woven from brushwood). These provide bank protection whilst stabilizing vegetation becomes established. Reeds and grasses also provide some degree of protection and encourage deposition of fine sediment at the channel margins. Although soft engineering is less expensive than building traditional hard engineering structures, it can be labour-intensive to install, takes time to establish, and requires regular monitoring and maintenance. Biotechnical revetments are less durable than more traditional approaches, meaning that these approaches are not suitable for higher-energy environments. Controlled flooding of wetland areas and flood basins An alternative method of flood control is to divert floodwaters into temporary storage on floodplains, wetland areas or artificial flood basins, in order to reduce the size of the flood peak further downstream. Wetland areas can be reinstated by changing floodplain land use to grazing or other low-risk activities. Control structures such as sluice gates may be used to divert floodwater onto the floodplain to reduce the flood risk to urban centres and valuable agricultural land further downstream. There are also many examples of flood basins, which are constructed to provide a temporary store for floodwater. Controlled by sluice gates, water is then gradually released as the flood recedes.

RIVER RESTORATION

River restoration is defined as the restoration of a river to a former or original condition (Downs and Gregory, 2004: 240). This is where techniques are applied to assist recovery and accelerate the re-establishment of natural physical and ecological processes. It is often not possible to restore a river fully to its former condition, owing to changes in the drainage basin, which alter the flow and sediment regimes. However, a new condition can be established in which natural function is restored. Restoration techniques can be passive or active. **Passive restoration** involves addressing factors that are preventing recovery, such as activities within the basin that adversely affect water quality, sediment and flow regimes. Examples of passive restoration include the establishment of buffer strips – zones of natural floodplain vegetation alongside the channel – and reinstatement of the flow regime. **Active restoration** is when specific modifications are made to accelerate recovery, such as the morphological reconstruction of meanders, riffles and pools. **Buffer strips and riparian zone management** Natural riparian vegetation fulfils a number of important roles, including bank stabilisation, nutrient regulation, filtering of sediments, shading, a source of large woody debris and a nesting site for birds (Gordon *et al*., 2004). However, intensification of agricultural practices has resulted in the widespread removal of riparian vegetation to provide access to water for grazing animals, and to increase the cultivated area. Livestock can have a significant influence on river channels when banks are trampled and overgrazed. This leads to instability and channel widening, with a resultant increase in the volume of sediment entering the channel. It is estimated that livestock grazing has altered approximately 80 per cent of stream and riparian ecosystems in the western United States (Belsky *et al*., 1999). By fencing off riparian areas, livestock access can be restricted to allow natural vegetation to re-establish. Buffer strips are strips of channel-adjacent land that run parallel to the river. Natural vegetation is encouraged to re-establish in these zones, which, together with the channel, provide corridors across the landscape. Buffer strips have been shown to filter fine sediments, which may carry contaminants (Chapter 5), and allow dissolved chemicals to be taken up by plants (Large and Petts, 1994). Flood peaks are reduced when floodwaters are stored on the floodplain, and there are also significant benefits for groundwater quality due to increased recharge. The effective width of buffer strips varies from channel to channel and is partly dependent on the size of the channel, while wider areas are needed to trap fine sediments such as silt. Typical widths range from 15 m to 80 m (Large and Petts, 1994). In South Africa, exotic riparian vegetation, such as the Australian Eucalyptus, creates major problems because of the vast quantities of water it consumes. Since the Working for Water programme began in 1995 over a million hectares of thirsty invasive alien plants have been cleared. If these plants were not removed, 30 per cent of runoff would be lost to rivers within ten to twenty years, and 74 per cent within twenty to forty years (DWAF, 2001). The programme has been highly successful in restoring flow to rivers and in creating large numbers of jobs. **Restoring the flow regime** Flow regulation is widely recognised as being one of the main causes of environmental degradation in rivers and is a priority issue in river management worldwide. Flow restoration involves adjusting regulated flow regimes and returning diverted flows to rivers. This can be controversial, particularly in water-stressed regions, where the available water supply is barely adequate to meet competing demands for human needs. Determining what is an 'ecologically acceptable' flow regime is a difficult task. In addition to determining a minimum acceptable flow to sustain biological communities, flow variability is necessary to maintain ecosystem structure and function. Flood pulses are a particularly important aspect of natural flow regimes. Flood flows with higher return periods are needed to rejuvenate the floodplain system, while more frequent, channel forming floods maintain the morphological features of the channel (Petts and Maddock, 1996). As a general rule, this corresponds to a flow with a oneand- a-half-year return period for stable alluvial channels, although there are wide variations (see Chapter 3, pp. 32–33). Considerations such as the dam outlet design impose limitations on the magnitude of artificial floods. To maintain salmon fisheries in regulated rivers, **pulse releases** are made from dams to stimulate migration and to flush out fine sediment from spawning gravels. The timing of such flows has to be carefully planned to fit in with life cycles. Flow regulation can also be used to enhance natural floods. A natural one-in-five-year event on the Murray River in Australia, which occurred in 2000, was enhanced by three releases from a headwater dam. This increased the duration of the flood peak and extended the recession curve, maintaining suitable conditions for breeding birds in the Barmah–Millewa Forest. As a result there was an increase in the number of species and bird numbers (Gordon *et al*., 2004). An interesting experiment was carried out on the Colorado River, United States, in April 1996, when a controlled flood flow was released into the Grand Canyon, from the Glen Canyon Dam, over a seven-day period. The aims of the experiment included remobilising sediment that had built up where unregulated tributaries joined the main channel, and building up sandy beaches. The controlled flood provided research opportunities for scientists from a range of fields and was closely observed. **Dam removal** may be a viable option, and many smaller structures have been decommissioned in the United States and elsewhere. A major problem associated with dam removal is the large volume of sediment that is trapped behind the dam, since releasing it into the channel would have serious consequences downstream. **Morphological restoration** Passive restoration – environmentally sensitive flow regimes and buffer zones – enables the full or partial recovery of formerly degraded channel reaches and river systems. 'Recovery' rarely involves a return to the original state, since land use and other changes in the catchment are likely to have altered the flow and sediment regimes. However, the morphology and function of the channel are restored, providing the diversity needed to sustain river health. In the case of lowland channels with limited stream power, slow rates of adjustment mean that natural recovery could take several centuries (Gordon *et al*., 2004). The process of recovery can be accelerated through **structural restoration**, the main aim of which is to restore structural diversity to a given reach. Morphological reconstruction is also used on upland reaches, where previous channel modifications have led to instability. Structural restoration is expensive and labour intensive, meaning that it is feasible in only a limited number of cases. There are also adverse impacts for in-stream and riparian biota, which will experience a certain amount of disturbance during the construction phase. A number of different techniques are employed, but all require a good understanding of fluvial dynamics, and morphological structure and function. Careful design is necessary to ensure that instability does not result, and that the newly modified channel is selfmaintaining. Particular attention needs to be paid to flow resistance and rates of sediment transport. In-stream structures In-stream structures are installed to recreate physical diversity in channelised reaches. These structures create areas of flow divergence and convergence, modifying conditions at the bed and encouraging localised erosion and deposition. Prior to installation, a detailed assessment of the reach is essential, in order to understand the causes of channel degradation and the needs of fish and in-stream biota. Care is required in installing these structures as they can work *against* the river if incorrectly applied (Downs and Gregory, 2004). For example, in higher-energy settings, it is important to ensure that erosion does not occur where flow is deflected towards the bank. **Weirs or low dams** create an area of ponding immediately upstream from the structure. As the flow cascades over the weir, scouring of the bed occurs at its base. Weirs are constructed using natural materials, such as wood or stone, and can also be angled to encourage localised areas of scour and deposition, as shown in the diagram. Obviously it is important to ensure that these structures do not impede the upstream migration of fish and other species. Such structures need careful design and must be 'keyed in' to the bed or bank so that they are not undermined by erosion (Downs and Gregory, 2004). This could lead to subsequent failure of the structure. **Flow deflectors** are widely used to promote localised scour and deposition, to deepen the channel and to divert flow away from vulnerable banks. At the same time, recirculating flow deposits material at the edges of the channel. Other configurations can be used to create different patterns of flow convergence and divergence, according to the specific needs of the application. A critical design consideration is setting the height of flow deflectors relative to the normal range of water levels. For example, if live willow logs are set too high, excessive growth could lead to flood obstruction (River Restoration Centre, 1999). Bank-attached **groynes** or **spur dikes** are often used for bank protection. These project at an angle from the bank and deflect the *thalweg* away from the eroding bank. Constructed from materials that include stone or woven fencing, these structures encourage deposition. Over time the colonisation of vegetation stabilises the deposits. **Sills** are fully submerged, even during low flow conditions, and modify near-bed flows. Depending on the characteristics of the specific reach, boulders can be placed to provide shelter for fish and anchorage for plants. Boulders also affect flow patterns, to create small-scale variations in the channel substrate and provide a means of dissipating excess energy. The significance of large woody debris is becoming increasingly recognised. Emplacement of bank-attached tree trunks and root boles greatly increases microhabitat availability and provides shelter for fish at high flows. Modifying channel dimensions Rivers that have been engineered for navigation or flood control have often been over-deepened and/or widened. Various techniques are used to reduce channel width. For example, in-stream structures, such as flow deflectors, can be installed to encourage deposition along one or both banks of the channel. Where rates of sediment supply are low, **aquatic ledges** or **berms** are installed. As well as narrowing the channel, these provide a valuable wetland habitat during low flows. The steep, canyon-like banks of over-deepened channels can be reprofiled mechanically, using a bulldozer or similar. Excavation may also be used to restore asymmetric variations in channel cross-section. Overdeepened reaches are sometimes partially infilled with gravel, care being taken to ensure that it is of a size that will not be flushed downstream. To restore channel–floodplain interactions, excavation of the floodplain surface on one or both sides of the channel creates a new, lower elevation floodplain adjacent to the channel. With all these designs, it is also important to recreate the natural variability that occurs along channel boundaries, including overhangs, variations in bank angle and irregularities along the line of the bank. Meander restoration

As well as providing a variety of flow conditions, meander reconstruction lengthens the channel and thus increases the area of channel habitat. Meanders also reduce channel gradient and, in turn, velocity and transport capacity. Re-meandering, if carefully designed, can therefore reduce the instability of artificially straightened reaches. Meanders have been reconstructed on a number of previously straightened lowland channel reaches in Germany, Denmark, the UK and elsewhere. Since 1994, restoration of over 20 km of the River Brede in South Jutland, Denmark, has re-created a meandering channel that flows across a 500 m wide floodplain. A smaller-scale example of meander restoration in an urban setting is the 1 km re-meandered reach of the River Skerne, England (Box 10.2). Careful design of meander bends is essential if instability is to be avoided. In some cases, the former meandering course can be reconstructed using old maps, aerial photographs, or other evidence. For example, the original course of the River Cole in Oxfordshire, England, was traced from historic maps, from remnants of meanders on the floodplain and avenues of willows that lined its former course. Another method is to copy a new meandering course from one or more **reference reaches** elsewhere along the same, or a similar, river. A rough idea of meander dimensions can be obtained using established relationships between channel width and meander wavelength. However, other variables such as slope, channel substrate and rates of sediment transport also need to be taken into account. After meander construction there may be some migration and adjustment of the channel, so bank protection may be necessary where erosion is undesirable. Recreation of pools and riffles Riffles and pools provide a means of dissipating excess energy as well as increasing habitat diversity. Various methods are used, including the excavation of bed sediment to create pools, the placement of gravels or cobbles to form riffles or through the use of appropriately spaced in-stream structures. For any of these methods to be successful, an understanding of morphologic controls and flow-sediment interactions is essential to prevent newly created pools from silting up, riffles from being eroded, or the spaces between newly laid gravels becoming infilled with fine sediment. Steps and pools These structures help to oxygenate the water, but are only appropriate in steep reaches where steps and pools would naturally occur. They may also impede the upstream migration of fish, so may not always be suitable. Weirs are constructed from stone walls, gabions, wooden poles or tree trunks. Water ponds up behind the weir before cascading over the crest to scour another pool at the base of the weir. Bank protection is usually necessary to prevent weirs from being removed by erosion.

CHAPTER SUMMARY

The development of land and water resources has had an adverse effect on many rivers, modifying flow regimes, water quality, morphology, and physical and ecological function. Direct impacts on river channels result from dam construction and channelisation. Indirect impacts are brought about by activities within the drainage basin, such as agricultural development, urbanisation, and land drainage schemes. Traditional management has focused on human needs, with little regard for the physical, chemical and biological condition of river channels. River channels are increasingly fragmented by dams, weirs, levees and other structures. These block migratory paths and prevent important transfers between channels and floodplains. Altered flow regimes have many adverse effects, such as disrupting the timing of life-cycle events and reducing the extent and duration of flooding. Channelisation greatly reduces the diversity of habitats within river channels, meaning that few species can survive. The links between ecosystem health and human health are increasingly recognised, and growing environmental concern has led to a shift in management priorities. New 'soft' engineering techniques have been developed to reduce the impact of channel interventions, at the same time as addressing human needs. There are now many examples of formerly degraded rivers that have been restored by applying techniques to assist recovery. Careful design is necessary to ensure that the structures installed are appropriate for the specific application. Channel instability problems can result if these techniques are incorrectly applied. Further research is also needed to improve on existing designs. It is rarely possible to return a river system to its original condition, but its natural function can potentially be restored. This can be achieved by restoring flow regimes, managing land use and reconstructing channel morphology (Charlton, 2008).

UNIT-11: DROUGHT MANAGEMENT: DROUGHT ASSESSMENT AND CLASSIFICATION, DROUGHT ANALYSIS TECHNIQUES, DROUGHT MITIGATION PLANNING

Introduction

Drought is a complex, slow-onset phenomenon of ecological challenge that affects people more than any other natural hazards by causing serious economic, social and environmental losses in both developing and developed countries. The period of unusual dryness (i.e. drought) is a normal feature of the climate and weather system in semi-arid and arid regions of the tropics, which covers more than one-third of the land surface and is vulnerable to drought and desertification. A drought is an extended period where water availability falls below the statistical requirements for a region. Drought is not a purely physical phenomenon, but instead is an interplay between natural water availability and human demands for water supply. There is no universally accepted definition of drought. It is generally considered to be occurring when the principal monsoons, i.e. southwest monsoon and northeast monsoon, fail or are deficient or scanty. Monsoon failure causing crop failure, drying up ecosystems and shortage of drinking water results in undue hardship to the rural and urban communities.

In contrast to aridity, which is a permanent feature of climate and is restricted to low rainfall areas, a drought is a temporary aberration and it is known to cause extensive damage and affects a significant number of people. It is named by some authors as a creeping phenomenon. Drought is a recurring natural phenomenon associated with a deficit availability of water resources over a large geographical area and extending along a significant period of time. Approximately 85% of the natural disasters are related to extreme meteorological events with drought being the one that causes most damages. Drought is considered as a major natural hazard, affecting several sectors of the economy and the environment worldwide. It affects almost all the determinants of the hydrological cycle starting from precipitation and ending with stream flow in the surface water systems or the recharge and storage in the groundwater aquifers. Therefore, our analysis depends upon which part of the hydrological cycle we are interested to focus.

Droughts produce a complex web of impacts that span many sectors of the society, including economy and may reach well beyond the area experiencing a drought. Wilhite & Glantz and the American Meteorological Society classify the droughts into four categories: meteorological, hydrological, agricultural and socioeconomic. According to statistics 35 countries will face severe water shortage by the year 2020.

The identification of drought was realized by using drought indices. Spatial and temporal extents and severity of drought could be determined using these indices. Drought indices are quantitative measures that characterize drought levels by assimilating data from one or several variables (indicators) such as precipitation and evapotranspiration into a single numerical value. It can provide decision makers with information on drought severity and can be used to trigger drought contingency plans, if they are available. Since the initiation of the drought indices notion, multiple drought indices have been developed (more than 150). They include the Palmer Drought Severity Index (PDSI), Rainfall Anomaly Index (RAI), Deciles, Crop Moisture Index (CMI), Bhalme and Mooly Drought Index (BMDI), Surface Water Supply Index (SWSI), National Rainfall Index (NRI), Standardized Precipitation Index (SPI), Reclamation Drought Index (RDI) and Reconnaissance Drought Index (RDI). The soil moisture drought index (SMDI) and crop-specific drought index (CSDI).

Generally the Deciles are the one of the simplest meteorological drought indices. The Standardized Precipitation Index is being used for defining and monitoring the drought due to its low data requirement and its ability to analyze the various aspects of drought, which is based on varying time-scales. The Reconnaissance Drought Index is developed to be used in semi-arid and arid regions over the world. This drought index is based on both cumulative precipitation (P) and potential evapotranspiration (PET). It is quite useful drought indicator with low data requirements, flexibility and high sensitivity. This paper aims to monitor and to assess the severity and duration of the meteorological drought in Wadi Djelfa- Hadjia sub-basin by evaluating the performance of the three indices.

What is a Drought?

The term drought is widely used but no unique definition exists across disciplines. A consequence is the difficulty to understand drought characteristics across time and space. In general terms, IPCC (2012) defined drought as "a period of abnormally dry weather long enough to cause a serious hydrological imbalance". It results from a shortfall of precipitation over a certain period of time and/or from a negative water balance due to an increased atmospheric water demand following high temperatures or strong winds. This situation may be exacerbated by antecedent conditions in soil moisture, reservoirs and aquifers, for example, and typically lasts from months to a few years. Extreme "Megadroughts" can persist for decades, while so-called "Flash Droughts" are short periods (< 3 months) of high temperatures, resulting in a fast depletion of soil moisture that can lead to major impacts (Mo and Lettenmaier, 2016).

Indeed, droughts are a recurring feature of all climates and are defined with respect to the long-term average climate of a given region (e.g., Heim Jr., 2002; Dai, 2013). They are to be distinguished from aridity, a seasonally or fully dry climate (e.g., desert) and from water scarcity, a situation where the climatologically available water resources are insufficient to satisfy long-term average water requirements (e.g. van Lanen et al. 2017; Tallaksen and van Lanen, 2004).

Depending on the prevailing effects on the hydrological system and the resulting impacts on society and environment, meteorological, soil moisture, and hydrological droughts (groundwater, stream flow, and reservoirs) are distinguished. The definition of a drought and the assessment of the related risk, therefore, will depend on the sector analyzed and the related processes and impacts.

Assessing the drought risk

Figure 1 presents the three components of drought risk and Figure 2 the resulting global drought risk map. As described in Carrão et al. (2016), the three components of risk were aggregated following a multivariate and non-parametric linear programming algorithm (Data Envelopment Analysis). The values for each component are not an absolute measure, but a relative statistic that provides a regional ranking of potential impacts (hotspots) with which on can prioritise actions to reinforce adaptation plans and mitigation activities. Figure 2 shows that on a global scale drought risk is generally higher for highly exposed regions - mainly populated areas and regions extensively exploited for agriculture - such as South and Central Asia, North Eastern China, the Southeast South American plains, Southern, Central and Eastern Europe and the Midwestern United States.

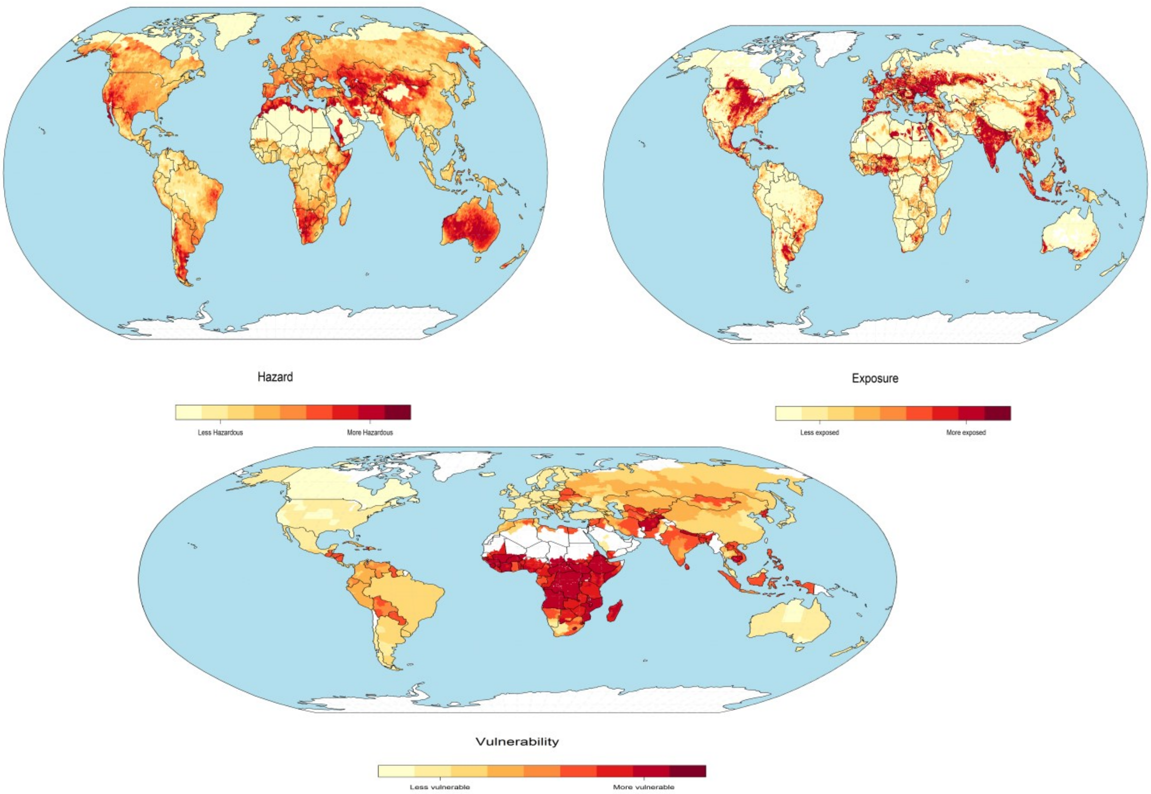


Figure 6. Global Drought Hazard according to the Weighted Anomaly of Standardized Precipitation (WASP) Index (upper left), Exposure (upper right), and Vulnerability (Lower panel).

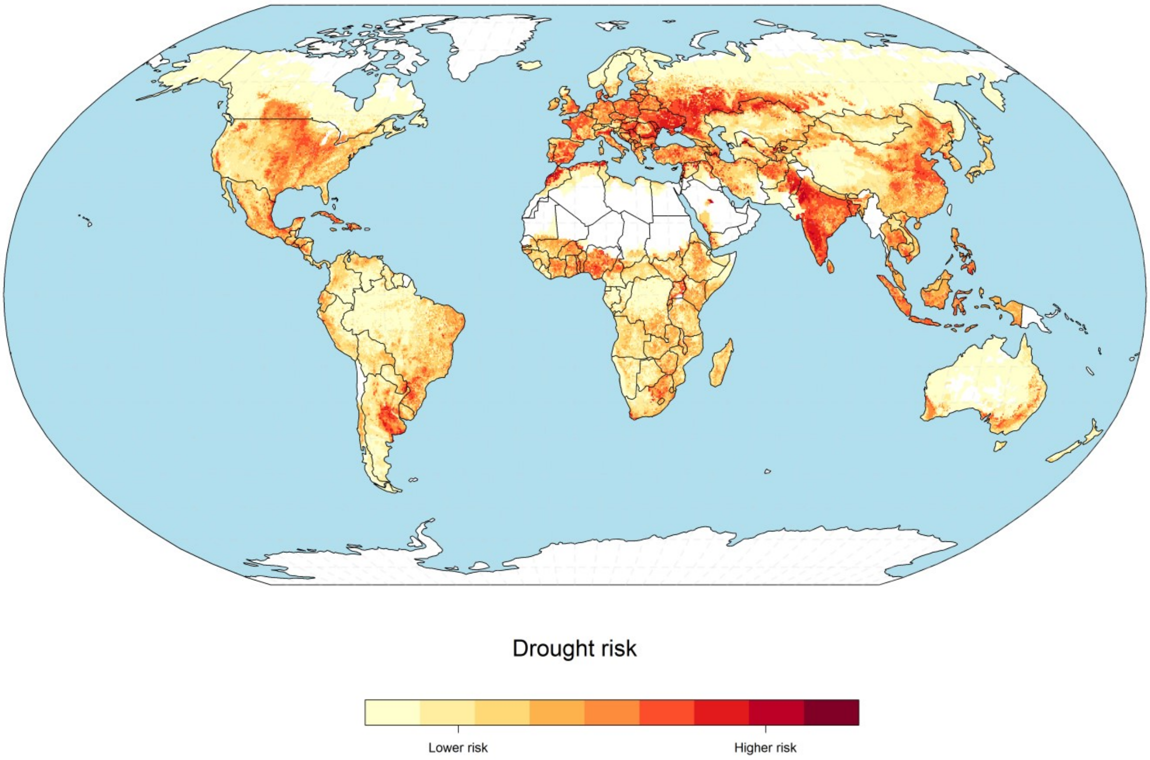


Figure 2. *Drought Risk based on the risk components shown in Figure 1.*

Several factors of uncertainty have to be considered in such analysis, since the metrics involved are partially subjective and conditioned by the data availability at global scale. Indeed, as exemplified above, agricultural drought can be quantified by a number of different indicators, each one able to provide a valid estimate of the different components of drought risk. As an example, Figure 3 depicts the drought hazard map according to the soil moisture-based Yearly Drought Severity Index (YDSI), which quantifies the simultaneous occurrence of a soil water deficit and extremely rare dry conditions (Cammalleri et al., 2016). It could replace or be combined with the WASP index used in the upper left panel of Figure 1

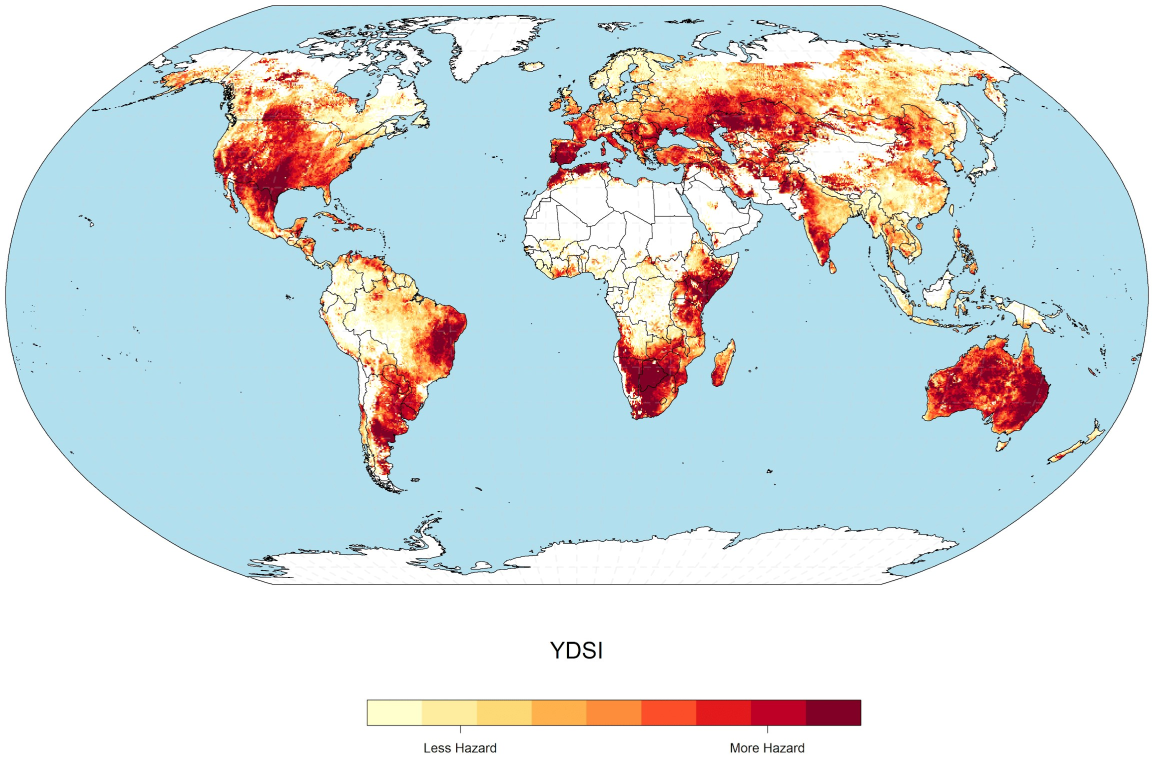


Figure 8. *Drought Hazard according to the Yearly Drought Severity Index (YDSI).*

Even if it is possible to observe analogies in the general patterns between the drought hazard map in Figure 1 (upper-left panel) and the one in Figure 3, it is evident how different conclusions at local scale can be obtained by using one indicator rather than the other.

Classification of Drought

Wilhite and Glantz (1985) categorized the droughtin terms of four basic approaches: meteorological, hydrological, agricultural and socioeconomic. The first three approaches deal with ways to measure drought as a physical phenomenon whereas the last deals with drought in terms of supply and demand, tracking the effects of water shortfall as it ripples through socioeconomic systems.

A. Meteorological Drought

Meteorological drought, also termed climatological drought, is based on precipitation's departure from normal average over a certain period of time and region. It is defined usually on the basis of the degree of dryness (in comparison to some normal or average amount) and the duration of the dry period. Definitions of meteorological drought must be considered as region specific since the atmospheric conditions that result in deficiencies of precipitation are highly variable from region to region.

B. Hydrological Drought

Hydrological drought is associated with the deficiency of water on surface or subsurface due to shortfall in precipitation. Although all droughts have their origin from deficiency in precipitation, hydrological drought is mainly concerned about how this deficiency affects components of the hydrological system such as soil stream-flow, moisture, groundwater and reservoir levels.

It takes longer for precipitation deficiencies to show up in components of the hydrological system such as soil moisture, stream-flow and groundwater and reservoir levels. As a result, these impacts are out of phase with impacts in other economic sectors. For example, a precipitation deficiency may result in a rapid depletion of soil moisture that is almost immediately discernible to agriculturalists, but the impact of this deficiency on reservoir levels may not affect hydroelectric power production or recreational uses for many months. Also, water in hydrologic storage systems (e.g., reservoirs, rivers) is often used for multiple and competing purposes (e.g., flood control, irrigation, recreation, navigation, hydropower, wildlife habitat), further complicating the sequence and quantification of impacts. Competition for water in these storage systems escalates during drought and increases the conflicts among water users.

C. Agricultural Drought

Agricultural drought links various characteristics of both meteorological and hydrological drought to agricultural impacts, focusing on precipitation shortages, differences between actual potential evapotranspiration, soil water deficits and reduced groundwater or reservoir levels. Plant water demand depends on prevailing weather conditions, biological characteristics of the specific plant and its stage of growth and the physical and biological properties of the soil.

A good definition of agricultural drought should be able to account for the variable susceptibility of crops during different stages of crop development, from emergence to maturity. Deficient surface soil moisture at planting may hinder germination, leading to low plant populations and a reduction of final yield. However, if surface soil moisture is sufficient for early growth requirements, deficiencies in sub-surface soil moisture at this early stage may not affect final yield if sub-surface soil moisture is replenished as the growing season progresses or if rainfall meets plant water needs.

D. Socio-economical Drought

Socioeconomic definitions of drought associate the supply and demand of some economic good with elements of meteorological, hydrological, and agricultural drought. This type of drought mainly occurs when the demand for an economic good exceeds its supply due to weather related shortfall in water supply.

It differs from the other above mentioned types of drought because its occurrence depends on the time and space processes of supply and demand to identify or classify droughts. The supply of many economic goods, such as water, forage, food grains, fish and hydroelectric power, depends on weather. Because of the natural variability of climate, water supply is ample in some years but unable to meet human and environmental needs in other years. Socioeconomic drought occurs when the demand for an economic good exceeds supply as a result of a weather-related shortfall in water supply. During 1988–89, in Uruguay, drought resulted in significantly reduced hydroelectric power production because power plants were dependent on stream-flow. Reducing hydroelectric power production required the government to rely upon more expensive petroleum and implement stringent energy conservation measures to meet the nation’s power needs.

Drought Analysis Techniques

Drought monitoring and assessment is done through analyses of variables such as rainfall, stream flow and soil moisture data on a variety of timescales. There are several methods that are applicable in this process. Frequency or probability-based methods analyse the low flows or low flow volume during a specific period. Regression-based methods bring out the relationship between the droughts parameters such as geomorphic and/or climatic factors, crop-yield factors, etc. with severe drought events In the theory of runs-based methods; the probabilistic structure of duration (run length) and severity (run sum) of a drought are analysed using the notion runs. Drought parameters such as longest duration and largest severity are analysed, based on the time series of random or Markovian variables. Discrete autoregressive and moving average processes model the variability of wet and dry years Group theory-based methods show the duration and lengths, as groups and cluster of groups. Datasets are analysed to develop drought prediction and forecasting techniques utilizing the concept of pattern recognition and neural networks.

The PDSI-based methods, the time series of PDSI are synthesized to identify and characterize the severity of droughts. Since PDSI series display a Markovian structure, such indices and their derivatives have been the focus for forecasting of agriculture droughts. Moisture adequacy index (MAI)-based methods: It is a measure of the degree of soil-moisture availability for plant growth. The Food and Agriculture Organization (FAO) of the United Nations has developed an algorithm to generate MAI time series for characterization of agricultural droughts and their severity.

Remote sensing and geoinformatics application

The process of resource exploitation and land-use patterns, migration and environmental degradation are responsible for the changing patterns of drought. Accurate, efficient and reliable information on drought hazard with spatial and temporal coordinates and attributes is necessary to communicate the potential risk to the specific vulnerable parts of the society.

Early warning and alert messages based on scientific monitoring techniques and methods would minimize the severity of the tragedy. The advancements made in the orbital satellite technology could aid in mapping the disaster area, prediction/forecasting of impending disaster, and disaster relief management. A number of satellites are available for weather forecasting, earth surface observations, monitoring and assessment.

The information from the NOAA, METEOSAT, INSAT and GMS, NOAA/AVHRR and IRS/WiFS, SPOT 4, DMSP/SSMI and IRS–P4/MSMR, TRMM, RESOURCESAT, MODIS and MERIS and LANDSAT, IRS and SPOT satellites are being used for prediction, vegetation-cover monitoring/early warning, drought information monitoring and drought management purposes.

GIS provides wider application of merging cartography, statistical analysis and database technology, for collation and interpretation, mapping and overlaying the attributes available as rater or vector data or non-spatial data on various aspects of drought risk and vulnerability. The benefit of using GIS over other conventional methods is mainly in handling large amounts of data in various scales and in bringing these on a map. GIS is significantly useful in combining spatial data from different sources together to identify and describe spatial associations present in the data and use the models for analysis and prediction. A number of methods are used to achieve this, viz. database query, overlay, proximity analysis, network analyses, digital terrain model, and statistical and tabular analyses.

Drought Mitigation Planning: -

Mitigation actions, programs, and policies are implemented during and before drought to reduce the magnitude of risk to human life, property, and productive capacity. Emergency response will always be a part of drought management, because it is unlikely that government and others will anticipate, avoid, or reduce all potential impacts through mitigation programs. A future drought event will also exceed the "drought of record" and the capacity of a region to respond. However, emergency response will be used lesser and only, if it is consistent with the longer-term drought policy goals and objectives.

Considering the increase in the frequency of droughts in different parts of the country, it is necessary that there is a shift in public policy from drought relief to drought mitigation measures. These measures are important for adapting to climate change, restoring ecological balance, and bringing development benefits to the people.

The services of Village Resources Centres being established by Indian Space Research Organization (ISRO), ICAR, State Agricultural Universities and other Organizations, will be effectively used towards management of droughts.

Drought mitigation and preparedness

To reduce the recurring impacts of drought, some countries are striving to obtain a higher level of mitigation and preparedness through development of national disaster reduction platforms as part of efforts such as the Hyogo Framework for Action and the United Nations Convention to Combat Desertification, or as part of separate national initiatives (ISDR, 2007). Additional drought risk reduction work is also being implemented at the grassroots level by local NGOs, state or provincial governments, and concerned citizens.

Such efforts are aimed at identifying and implementing strategies in advance of drought to reduce the likelihood of negative effects when drought does occur (i.e., mitigation), and developing contingency action plans to implement as drought conditions evolve in order to effectively respond to the situation (i.e., preparedness).

Some of the factors that have contributed to the trend toward drought mitigation and preparedness policies are spiraling costs of impacts associated with drought, the complexity of impacts on sectors well beyond agriculture, increasing concerns over social and environmental well-being, rising water conflicts between users, and an increasing number of professional practitioners and policy makers that recognize the importance of drought mitigation and planning.

Another factor that has helped promote the trend for more emphasis on drought mitigation and preparedness is the availability of drought planning methodologies. These methodologies have helped provide guidance to drought planners. For example, methodologies have been developed by Wilhite (1991) and Knutson et al. (1998) of the National Drought Mitigation Center (NDMC) in the United States to assist planners in preparing for drought. These strategies were merged by Wilhite et al. (2000, 2005) to place a greater emphasis on risk analysis in drought planning. The resulting strategy provides a set of guidelines or a checklist of the key elements of a drought plan and a process through which they can be adapted to any level of government (i.e., local, state or provincial, or national) or geographical setting as part of a natural disaster or sustainable development plan, an integrated water resources plan, or stand-alone drought mitigation plan.

Climate change, with its threat of an increased frequency of drought events in the future, has also caused greater anxiety about the absence of drought preparedness. According to the Intergovernmental Panel on Climate Change (IPCC, 2007), freshwater availability in Central, South, East and Southeast Asia, particularly in large river basins, is projected to decrease as a result of climate change, which, along with population growth and increasing demand arising from higher standards of living, could adversely affect more than a billion people by the 2050s. Furthermore, it is projected that crop yields could decrease up to 30% in Central and South Asia by the middle of the 21st century. Taken together and considering the influence of rapid population growth and urbanization, the risk of hunger is projected to remain very high in several developing countries. Endemic morbidity and mortality due to diarrheal disease primarily associated with floods and droughts are also expected to rise in East, South and Southeast Asia because of projected changes in the hydrological cycle associated with global warming. The IPCC findings also indicate dramatic changes in temperature and precipitation conditions for the Mediterranean region and a likely increase in drought frequency and intensity. Concomitant impacts on water supply and demand would also be expected to occur.

CAUSES OF DROUGHT OVER INDIA: -

Drought is a regional manifestarion *of* a general climatic fluctuariom associated with rhe abnormal atmospheric circulation patterns caused by extra-terrestrial or terrestrial factors.

(a) In extraterrestrial forces, there is a *Il-year* cycle in the variation of annual mean sunspot number, but evidence of such a cycle in the energy output of the sun arc lacking. If such cycle exists, the atmospheric circulation would be in a state of constant re-adjustment in accordance with the greater or lesser amoullt of solar energy rcaching the earth causing drought first in one region then in another.

(b) Terrestrial Factors pertain to changes on the earth and following attributes lead to large-scale drought.

(i) Volcanic activity: Injection of large amounts of ash and dust into the atmosphere by violent volcanic actjyjry may alter the earth's radiation balance and, thereby, create compensating circulation adjustments. which includes climatic fluctuations.

(ji) Composition of atmosphere: Water vapour, carbon dioxide and ozone would modify the heat balance of the earth and in-cum produce fluctuations *in* large-scale cjrculacion pattern resulting in drought. The timing, intensity and frequency of meteorological extremes such as drought may be increased and influenced by the enhanced concencration of atmospheric green house gases, such as CO" merhane and nlrrous oxide resulring info global warming (Shukla. 1991 and 1998).
(iii) ImerdCtion in Earth-Ocean-Atmosphere complex: The enormous heat-storage capacity of the ocean and energy exchanges which take place between the atmosphere and the ocean, make airsea inreraction likely cause of climate fluctuations. The biggest natural climate fluctuations occurring globally on time scales of a few years are the Southern Oscillarion and associated El Nino phenomenon (ENSO). Major EN SO episodes lead (Q massive displacement of rainfall regions of the rropics, causing drought or torrential rains. ENSO has an irregular period, but normally occurs once in every 2 co 7 years. FNSO shows strong link with droughts over India, Indonesia, Australia and torrential rains over coastal areas of Ecuador and northern Peru.

References: -

- Alam, N. M., Ranjan, R., Jana, C., Singh, R. J., Patra, S., Ghosh, B. N., & Sharma, N. K. (2014). Drought classification for policy planning and sustainable agricultural production in India.
- Gupta, A. K., Tyagi, P., & Sehgal, V. K. (2011). Drought disaster challenges and mitigation in India: strategic appraisal. *Current science*, 1795-1806.
- Narain, P., Rathore, L. S., Singh, R. S., & Rao, A. S. (2006). Drought assessment and management in arid Rajasthan.
- Sundnes, K. O., & Birnbaum, M. L. (2003). Health disaster management: Guidelines for evaluation and research in the Utstein style. *Prehospital and Disaster Medicine*, *17*(Supplement 3).
- Wilhite, D. A., & Knutson, C. L. (2008). Drought management planning: Conditions for success. *Options Mediterraneennes Series A*, *80*, 141-148.

UNIT-12: METHODS OF WATER CONSERVATION: RAINWATER HARVESTING AND WATERSHED MANAGEMENT

Introduction

Water, Paani, Jal, Tanni, L'eau, Wasser, Acqua perhaps the most familiar and widely used word in the world. Water needs no introduction, the importance of this is known to one and all. However, despite water being the basic human need, this precious resource is being wasted, polluted and getting depleted. Every drop of water is precious but we continue to waste it like it is a free natural commodity. 98% of water on this planet is salty and is not fit for human consumption. Out of the 2% of fresh water reserves, 1% is locked up in form of ice in various regions around the world. Hence, only 1% of total water reserves are available for our domestic & industrial use. Many cities in India and around the world are already facing sever water shortages due to reduced rainfall, man-made climatic changes, reduction in ground water levels, population explosion, industrialization and staggering amount of water wastages because of negligence by users & dilapidated water supply systems. The importance of water in a country's economic growth should not be undermined.

Water pollution, unavailability of drinking water, inadequate sanitation, open dumping of wastes, loss of forest cover are some of the problems faced by many parts of India. Heavy toll of infant mortality due to water borne diseases, the daily struggle for procuring water, mismanagement of waste water, improper sanitation are common features and are leading to serious consequences on human health and the economy of the country. The situation demands immediate intervention in the management of these rapidly growing problems, especially through an integrated approach for water, sanitation and related issues.

Water Conservation and Management

Since there is a declining availability of freshwater and increasing demand, the need has arisen to conserve and effectively manage this precious life giving resource for sustainable development. Given that water availability from sea/ocean, due to high cost of desalinisation, is considered negligible, India has to take quick steps and make effective policies and laws, and adopt effective measures for its conservation. Besides developing water-saving technologies and methods, attempts are also to be made to prevent the pollution. There is need to encourage watershed development, rainwater harvesting, water recycling and reuse, and conjunctive use of water for sustaining water supply in long run.

Prevention of Water Pollution

Available water resources are degrading rapidly. The major rivers of the country generally retain better water quality in less densely populated upper stretches in hilly areas. In plains, river water is used intensively for irrigation, drinking, domestic and industrial purposes. The drains carrying agricultural (fertilizers and insecticides), domestic (solid and liquid wastes), and industrial effluents join the rivers. The concentration of pollutants in rivers, especially remains very high during the summer season when flow of water is low. The Central Pollution Control Board (CPCB) in collaboration with State Pollution Control Boards has been monitoring water quality of national aquatic resources at 507 stations. The data obtained from these stations show that organic and bacterial contamination continues to be the main source of pollution in rivers.

The Yamuna river is the most polluted river in the country between Delhi and Etawah. Other severely polluted rivers are: the Sabarmati at Ahmedabad, the Gomti at Lucknow, the Kali, the Adyar, the Cooum (entire stretches), the Vaigai at Madurai and the Musi of Hyderabad and the Ganga at Kanpur and Varanasi. Groundwater pollution has occurred due to high concentrations of heavy/toxic metals, fluoride and nitrates at different parts of the country. The legislative provisions such as the Water (Prevention and Control of Pollution) Act 1974, and Environment Protection Act 1986 have not been implemented effectively. The result is that in 1997, 251 polluting industries were located along the rivers and lakes. The Water Cess Act, 1977, meant to reduce pollution has also made marginal impacts. There is a strong need to generate public awareness about importance of water and impacts of water pollution. The public awareness and action can be very effective in reducing the pollutants from agricultural activities, domestic and industrial discharges.

Recycle and Reuse of Water

Another way through which we can improve fresh water availability is by recycle and reuse. Use of water of lesser quality such as reclaimed wastewater would be an attractive option for industries for cooling and fire fighting to reduce their water cost. Similarly, in urban areas water after bathing and washing utensils can be used for gardening. Water used for washing vehicle can also be used for gardening. This would conserve better quality of water for drinking purposes. Currently, recycling of water is practiced on a limited scale. However, there is enormous scope for replenishing water through recycling.

1.7 SELF-ASSESSMENT TEST

1. What is Lane balance between stream power and sediment supply?
2. Distinguish between aggradation and degradation.
3. Give an account of the driving variable and boundary conditions that control the form of a channel.
4. What is time scale of adjustment of channel form?
5. How adjustments are made in channel form?
6. Give an account of the different channel geomorphic units.
7. Differentiate between straight and meandering course?
8. Discuss the mechanism of flow and sediment transport in meander bends.
9. Give an account of the process of meander migration.
10. What is braided river?

1.8 SUMMARIES AND KEY POINTS

Following points have been discussed in this section:

- Plate tectonics as a unified theory of global tectonics
- Tectonic geomorphology: Influence of tectonics in landscape evolution
- Concepts in Geomorphology: spatial scale, temporal scale, systems, feedback, equilibrium and threshold
- Catchment process and fluvial processes; Factors regulating entrainment, transportation and deposition
- Adjustment of channel forms and patterns to morphodynamic variables
- Coastal morphodynamic variables and their influence in evolution of landforms
- Periglacial processes and landforms
- Elements of slope and different approaches to study slope development
- Concept of basin hydrology and run off cycle; Unit hydrograph, rating curve and their applications
- Storm water and flood management: storm water management, design of drainage system, flood routing through channels and reservoir, flood control and reservoir operation
- Drought management**:** drought assessment and classification, drought analysis techniques, drought mitigation planning
- Methods of water conservation: rainwater harvesting and watershed management

1.9 STUDY TIPS

Charlton, R. (2008). *Fundamentals of Fluvial Geomorphology.* Oxon: Routledge.

Gabler, R. E., Petersen, J. F., Trapasso, L. M., & Sack, D. *Physical Geography.* Broooks/Cole.

Guttierez, M. (2013). *Geomorphology.* CRC Press.

Strahler, A. (2013). *Introducing Physical Geography.* USA: Wiley.

Disclaimer: This self-learning material is based on different books, journals and web-sources.

Post-Graduate Degree Programme (CBCS)

in

GEOGRAPHY

Semester-I

Paper: GEO/CC/T-102

Self-Learning Material

DIRECTORATE OF OPEN AND DISTANCE LEARNING

UNIVERSITY OF KALYANI

Kalyani, Nadia
West Bengal, India

COURSE MATERIALS COMPILED BY			
Sl. No.	Names	Designation and Affiliation	SLM
1.	Dr. Sahina Khatun	Assistant Professor Department of Geography, University of Kalyani Kalyani-741235, Nadia District, West Bengal	**GEO/CC/T-102 Climatology**
2.	Supriya Ghosh	Ph.D. Research Scholar Department of Geography, University of Kalyani Kalyani-741235, Nadia District, West Bengal	
3.	Dr. Ayan Rudra	Assistant Professor (Contractual) Directorate of Open and Distance Learning University of Kalyani, West Bengal	

December, 2021

Directorate of Open and Distance Learning, University of Kalyani
Published by the Directorate of Open and Distance Learning,
University of Kalyani, Kalyani-741235, West Bengal and Printed by
East India Photo Composing Centre, 69, Sisir Bhaduri Sarani, Kolkata-700 006

December, 2021

SYLLABUS

Paper– GEO/CC/T-102

Semester-I;
Paper– GEO/CC/T-102: Climatology (Total Credit-4, Total Marks–50)
GEO/CC/T-102: Climatology
(Marks - 50: Internal Evaluation – 10, Semester-end Examination - 40)

Unit-1: Adiabatic and isothermal processes

Unit-2: Atmospheric stability and instability

Unit-3: Air-masses

Unit-4: Tri-cellular model

Unit-5: Atmospheric fronts: classification, formation and characteristics

Unit-6: Formation of precipitation; Rainmaking

Unit-7: El Nino, Southern Oscillation and La Nina

Unit-8: Monsoon: theories of its origin; Recent trends of monsoon in India

Unit-9: History of global climate change

Unit-10: Climate change: causes and evidences

Unit-11: Global warming: consequences and adaptations

Unit-12: Weather forecasting: short, medium and long range

Content

Page No.

SEMESTER-I; PAPER – GEO102T; GROUP – GEO102T.1: CLIMATOLOGY
(Credit – 2; Marks - 50: Internal Evaluation – 10, Semester-end Examination - 40)

Introduction

In this lesson, different aspects of climatology and climate will be discussed. By subdividing the branches of climatology, the content and scope is elaborated. Air masses and related aspects like adiabatic lapse rate, stability and instability, their relationship and resultant influence on climate is discussed. Atmospheric circulation and origin of monsoon are explained with help of few models and theories and brief overview of weather forecasting is incorporated in this lesson. Some abnormalities like El Nino, La Nina, global warming, climate change are discussed.

Learning Objectives

After going through this unit, the learners will able to

- Describe the content of the climatology
- Elucidate the concept of air mass, adiabatic lapse rate, stability and instability.
- Explain the mechanism of atmospheric circulation, monsoon, ENSO
- Write overview of the global warming and climate change

Assessment of prior knowledge

1. What do you mean by climate?
2. Highlight the major components of weather.
3. Differentiate the climate from weather.
4. State the relationship between pressure belts and planetary winds.

Learning Activities

Climatology deals with the climate, its elements, regional variations and influence on plants, animal and human activities. Going through the different climatic aspects, learners will be able to assess different climatic phenomena, their dynamics and cause-effects relationships of atmospheric conditions. In this regard, learners may collect the temporal and spatial data on different elements of climate and thereby, examine the trend of temperature, rainfall, humidity etc. as well as their spatial distribution. Besides, synoptic charts and different statistical simulation are to be prepared based on climatic data to predict the future weather conditions, climate change etc.

Feedback of Learning Activities

As feedback of learning activities, learners will consult with the groups and instructors regarding their objectives, methods and other related problems of the climate related projects and acquire necessary instructions and solutions. From the completion of these activities, they will able to endow their knowhow both from theoretical and practical backgrounds.

EXAMPLES AND ILLUSTRATIONS (SUBJECT CONTENT)

UNIT – 1: ADIABATIC AND ISOTHERMAL PROCESSES

INTRODUCTION

The word 'isothermal' means constant temperature. An isothermal process is a process occurring at a constant temperature. The word 'adiabatic' means isolated from surroundings. Adiabatic process means a process that neither allows the heat to transfer inside nor let the heat out of the system. For example, a reaction that takes place in a Dewar Flask is adiabatic. The conventional thermal cycles under use so far are based on the Carnot engine, that are quadrilateral cycles in which ideally heat is absorbed at high temperature (top temperature) to be expanded while performing mechanical work undergoing temperature decrease from the top temperature to approach the bottom temperature under a quasi isentropic transformation.

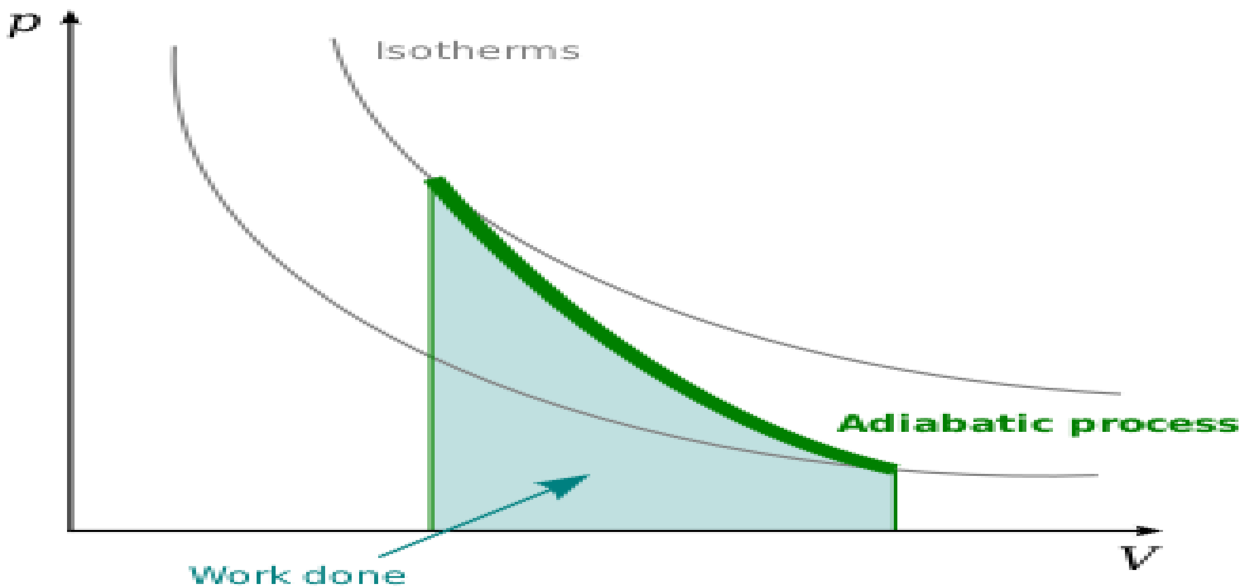
ADIABATIC AND ISOTHERMAL

Thermodynamics uses the concepts of the isothermal process, isochoric process, isobaric process, and adiabatic process to explain the behaviour of a thermodynamic system and its relationship to the temperature changes. An isothermal process is a process that happens when there are no variations in the temperature of the system, but other parameters (volume, pressure) regarding the system can be changed accordingly. Adiabatic process describes a process that remains aloof of its surroundings. It is a process in which no heat transfer occurs between a system and its surroundings. Here, the temperature of the system can vary in order to avoid any heat transfer. This indicates that the main difference between isothermal and adiabatic processes is that the isothermal process takes place under constant temperature whereas the adiabatic process occurs under changing temperature.

If a material undergoes a change in its physical state (e.g., its pressure, volume, or temperature) without any heat being added to it or withdrawn from it, the change is said to be adiabatic. Suppose that the initial state of a material is represented by the point A on the thermodynamic diagram below, and that when the material undergoes an isothermal transformation it moves along the line AB. If the same material undergoes a similar change in volume but under adiabatic conditions, the transformation would be represented by a curve such as AC, which is called an adiabat.

In meteorology and oceanography, adiabatic cooling produces condensation of moisture or salinity, over saturating the parcel. Therefore, the excess must be removed. There, the process becomes a *pseudo-adiabatic process* whereby the liquid water or salt that condenses is assumed to be removed upon formation by idealized instantaneous precipitation. The pseudo adiabatic process is only defined for expansion because a compressed parcel becomes warmer and remains under saturated.

Fig: - An isotherm and an adiabat on a p–V -diagram.



The adiabat AC is steeper than the isotherm AB. The reason for this is easily seen.

During the adiabatic compression (dα < 0) the internal energy increases:

$$dq = du + p\, d\alpha \text{ and } dq = 0 \Longrightarrow du = -p\, d\alpha > 0$$

and therefore the temperature of the system rises:

$$du = c_v\, dT > 0 \Longrightarrow T_C > T_A$$

However, for the isothermal compression from A to B, the temperature remains constant: $T_B = T_A$. Hence, $T_B < T_C$. But $\alpha_B = \alpha_C$ (the final volumes are equal); so

$$p_B = \frac{RT_B}{\alpha_B} < \frac{RT_C}{\alpha_C} = p_C$$

that is, $p_B < p_C$. Thus, the adiabat is steeper than the isotherm.

ADIABATIC ISOTHERMAL ISOBARIC ISOCHORIC

ADIABATIC PROCESS

An adiabatic process is a thermodynamic process that can take place without any heat transfer between a system and its surrounding. Here, neither heat nor energy is not transferred into or out of the system. Therefore, in an adiabatic process, the only way the energy transfer takes place between a system and its surrounding is the work.

An adiabatic process can be quickly maintained by doing the process. For example, if we quickly press the piston in a cylinder filled with gas, there is not enough time for the system to transfer heat energy to the surroundings. In adiabatic processes, the work done by the system alters the internal energy of the system.

ISOTHERMAL PROCESS

An isothermal process is a thermodynamic process that takes place at a constant temperature. It means that an isothermal process occurs in a system where the temperature remains constant. However, to keep the temperature of the system constant, heat must be transferred into the system or shifted out of the system.

Apart from that, many factors of the system also vary during the continuation of an isothermal process such as internal energy. To maintain the temperature of the system, it can be kept in a tight cylinder. Then, by regulating the temperature of the cylinder, we can control the temperature of the system to an optimal level.

Below are the examples of the isothermal process:

1. A phase change of matter
2. Melting of matter, and
3. Evaporation, etc.

A common industrial use of the isothermal process is the Carnot heat engine.

To maintain the temperature of the system, work should either be done on the system or be done by the system on the surrounding; doing work on the gas increases the internal energy, which, in turn, increases the temperature.

However, if the temperature rises more than the required range, then work is done by the system on the surrounding. However, when the temperature of the system decreases, the energy is released to the surroundings in the form of heat.

ISOBARIC PROCESS

An isobaric process has the word ‘bar’, where the bar is the unit of pressure. Another three letters added ‘iso’ make a process called the isobaric process. An isobaric process is a process that takes place under constant pressure.

DIFFERENCE BETWEEN ISOTHERMAL PROCESS AND ADIABATIC PROCESS

Isothermal Process	**Adiabatic Process**
Transfer of heat occurs in this process.	No heat goes inside or leaves the system.
At a given volume of the substance, the pressure remains high.	At a given volume of the substance, the pressure remains low.
In an isothermal process, the temperature remains invariant.	The temperature varies because of the internal system changes.
Heat can be added or released to the system to keep the temperature constant.	There is no addition of heat nor the release of the heat because maintaining constant temperature doesn’t matter here.
The isothermal process has a slower transformation flow.	The adiabatic process has a faster transformation flow.
In an isothermal system, work done is because of the change in the net heat content of the system.	In an adiabatic process, the work done is because of the change in internal energy.

UNIT - 2: ATMOSPHERIC STABILITY AND INSTABILITY

STABILITY AND INSTABILITY

Stability and instability are two atmospheric conditions based on the vertical motion of air parcel. The nature of vertical motion depends upon the relationship of normal lapse rate (6.5^{0}C/km) and adiabatic lapse rate.

Stability is that atmospheric conditions in which vertical motion of air parcel is restricted. If ascending air parcel is colder than surrounding environment, uprising of air is restricted and it starts to descend. When dry adiabatic lapse (DALR) rate is greater than environmental lapse rate (ELR), stability prevails in the atmosphere. For example, the temperatures of ascending air parcel and environment on ground are 30^{0}C and 25^{0}C respectively. At the elevation of 2km, the former gets 10^{0}C temperature due to adiabatic cooling (10^{0}C/km) and latter's temperature is 15^{0}C as it cools by normal lapse rate (5^{0}C/km). Thus, at this height, rising air parcel is colder than surrounding environment and tends to move downwards.

Sometimes, rising air still bound to sink down even after condensation at dew point when wet adiabatic rate (WALR) is greater than environmental lapse rate. Such stability is called *absolute stability*.

ABSOLUTE STABILITY (ELR<WALR<DALR)

Height (Mts.)	**Temperature of environment (ELR-5^{0}C/km)**	**Temperature of air parcel (DALR-10^{0}C/Km and WALR-6^{0}C/Km)**	**Remarks**
Suface	20^{0}C	20^{0}C	Dry adiabatic lapse rate
1000	15^{0}C	10^{0}C	
2000	10^{0}C	0^{0}C	Condensation level
3000	5^{0}C	-6^{0}C	Wet adiabatic lapse rate
4000	0^{0}C	-12^{0}C	

On the other hand, continuous vertical movement of air parcel having higher temperature than its surroundings is called atmospheric *instability*. Instability or rising of air parcel ceased when temperature of rising air parcel coincides with the temperature of the surrounding environment. When dry adiabatic rate is lower than environmental lapse rate, the ascending air parcel always remains warmer than surrounding and continues to rise upward. Such continued vertical movement is termed as *absolute instability*.

When moist air has a environmental lapse rate between dry and wet adiabatic rate, instability occurs only after attaining the condensation level and it is called *conditional instability*. ‘Condition’ is used because only if the air parcel is forced to attain condensation level, it becomes unstable and convergence, turbulence, orographic lifting help the air parcel to attain such level.

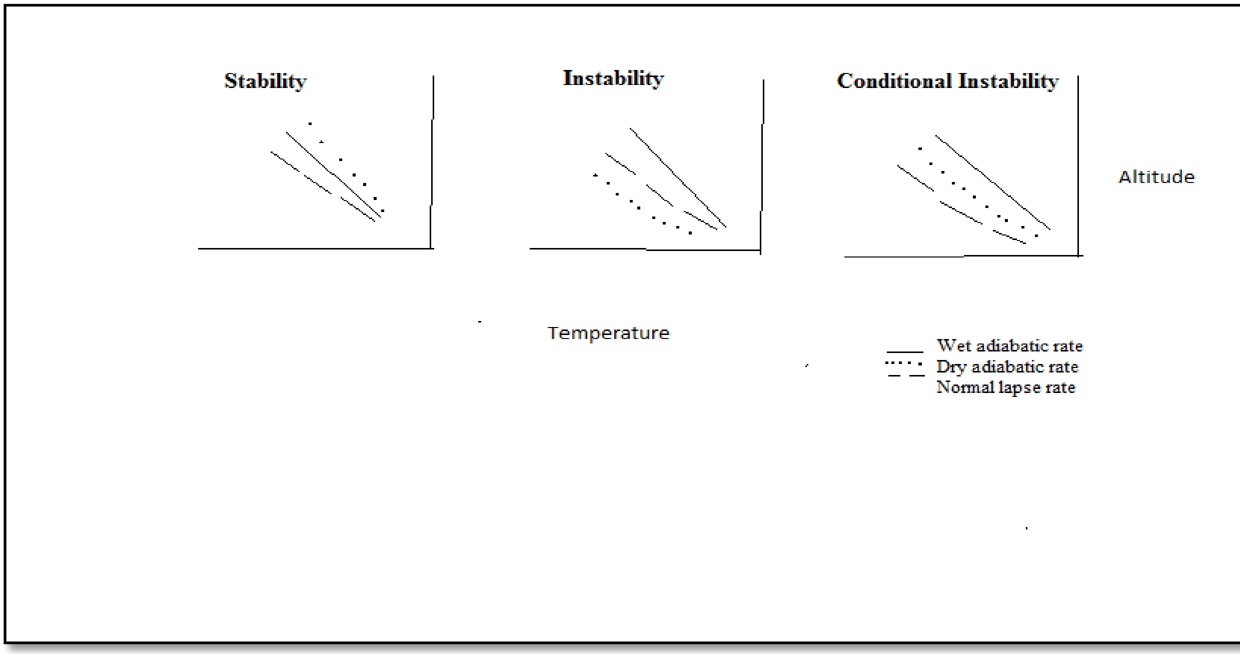
Various weather phenomena are associated with atmospheric stability and instability. Stable condition leads to subsistence and formation of inversion of temperature, fog, anticyclone etc. while instability helps to form vertically thickened clouds and heavy downpours.

Though the causes of these phenomena are still mysterious, the impact of these are very prominent both in tropical and subtropical region. Rhythm of Indian monsoon is also influenced by these shifting conditions of atmosphere.

ABSOLUTE AND CONDITIONAL INSTABILITY

	Absolute Instability (ELR>DALR>WALR		**Conditional Instability (DALR>ELR>WALR)**		**Remarks**
Height (Mts.)	Temperature of environment (ELR-12^{0}C/km)	Temperature of air parcel (DALR-10^{0}C/Km and WALR-5^{0}C/Km)	Temperature of environment (ELR-8^{0}C/km)	Temperatur e of air parcel (DALR-10^{0}C/Km and WALR-6^{0}C/Km)	
Suface	40^{0}C	40^{0}C	40^{0}C	40^{0}C	
1000	28^{0}C	30^{0}C	32^{0}C	30^{0}C	Dry adiabatic lapserate
2000	16^{0}C	20^{0}C	24^{0}C	20^{0}C	
3000	4^{0}C	15^{0}C	16^{0}C	14^{0}C	Condensation level
4000	-8^{0}C	10^{0}C	8^{0}C	8^{0}C	
5000	-20^{0}C	5^{0}C	0^{0}C	2^{0}C	Wet adiabatic lapserate
6000	-32^{0}C	0^{0}C	-8^{0}C	-4^{0}C	

CAPARISON OF STABILITY, INSTABILITY AND CONDITIONAL INSTABILITY OF ATMOSPHERE



UNIT: - 3 AIR-MASSES

DEFINITION OF AIR MASSES

According Barry and Chorley (1968) an air mass may be defined as a large body of air whose physical properties, especially temperature and moisture content and lapse rate (of temperature) are more or less uniform horizontally for hundreds of kilometers. A.N. Strahler and A.H.Strahler (1978) define air mass as a body of air in which the upward gradients of temperature and moisture are fairly uniform over a large area is known as air mass. According to Oliver and Hidore (2003) air mass is a large, horizontal, homogeneous body of air that may cover may cover thousands of squares kilometers and extend upward for thousands of meters. Its uniformity is principally one of temperature and humidity. So, an air mass is an extremely large body of air having more or less homogeneity properties of temperature and moisture content (humidity) both horizontally and vertically.

CHARACTERISTICS OF AIR MASSES

1. An air mass originates when atmospheric condition remain stable and uniform over a extensive area for fairly long period.
2. An air mass is seldom stationary over the source region rather it travels to other regions.
3. Contrasting air masses tend to remain separate by distinct sloping boundary surface, called front.
4. Air mass is modified by path, time and destination region.

Source Regions of Air mass: Source Regions are simply geographic areas where an air mass originates. Extensive and uniform surface and stable atmospheric condition etc. are some required characteristics of source region. Source region may be classified as equatorial, tropical, polar, arctic on the basis of latitude and continental and maritime on the basis of nature of surface.

MODIFICATION OF AIR MASSES

Air mass modification broadly classified into two viz. *thermodynamic modification* by exchange of heat with underlying surface or by changes in moisture content due to changing evaporation or by increase of moisture content from precipitation of overlying moist air and *mechanical modification* by the process of uplift or subsidence, divergence or convergence of air or by turbulent mixing. Thermodynamic modification of an air mass depends on initial temperature and moisture character of the air mass, nature of land or water surface, path of movement and time of travel. On the other hand, mechanical modification depends on convection, advection, cyclonic, anti-cyclonic condition etc.

CLASSIFICATION OF AIR MASSES

Air masses are classified on the basis of source region and their subsequent modification as moves away from source regions. Systematically, the classification of air masses can be organized as follows-

1. Geographical Classification (based on source region)
2. Thermodynamic Classification (based on modification)
3. Composite Classification (based on combination)

1. Geographical Classification: Petterssen has identified five types of air masses on the basis of the source region viz.
2. *tropical air mass, polar air mass, arctic, anterctic and equatorial* . But, Trewartha, on the basis of geographical location of air masses, has classified into two groups i.e. *polar air mass*(formed pole-ward of 60º) which is denoted by capital 'P' and *tropical air mass* (formed within 30º of the equator), denoted by capital'T'. He also opined that Arctic, Anterctic and Equatorial air masses are considered to bemodified form of Tropical and Polar air masses.

These two types of air masses have been further subdivided on the basis of nature of surface of their source region i.e. land or water surface- continental air mass (c)- formed over large land masses and maritime (m) -formed over the oceans. So, on the basis of source region as well as the nature of surface, the principal types of air masses are-

- cP – continental polar air mass
- mP- maritime polar air mass
- cT – continental Tropical air mass
- mT – maritime tropical air massContinental polar (cP) air masses:

Origin: North America and Eurasia in the sub -Arctic zone i.e. Siberia, northern Canada, Greenland and Antarctica.

Characteristics:

- Low specific humidity
- Very cold in winter
- The depth of these air mass is almost 3km for heavy coldness
- Due to coldness, cPWs type of air mass is formed at the centre of the source region.
- cP lost its characteristics when passes over the ocean because it takes both heat and moisture from the ocean and turn into mPKu air mass.

Maritime poplar (mP) air masses:

Origin: mid latitude oceans

Characteristics:

- Consist of cool moist air with steep laps rate
- Orographic type of precipitation.Continental tropical (cT) air masses: *Origin*: Subtropical

deserts *Characteristics:*

- less water vapour content
- Low relative humidity
- At summer time temperature is 50 º -55ºC
- Hottest air mass at the surface
- cT changes into mT air mass when cT leaves its source region and begins to hold moisture.
- cT air mass dominates over North America in winter and South Africa in summer.

Maritime tropical (mT) air masses:

Origin: warm oceans in the tropical and subtropical zones

CHARACTERISTICS

- high specific humidity
- Heavy precipitation
- Cumulus and cumulonimbus type of cloud

3. Thermodynamic Classification: on the basis of under the Influence of Underlying Surface air mass may be warm air mass-warmer than underlying surface (W) and cold air mass-colder than underlying surface (K) and on the basis of heat exchange between air mass and underlying surface air mass may be stable (s) and unstable (u)

4. Composite Classification: Based on the geographical and thermodynamic classifications combinedly, 16 types of air masses are classified.

A. Tropical Air masses

1. cTWs – Continental Tropical Warm Stable
2. cTWu – Continental Tropical Warm Unstable
3. cTKs – Continental Tropical Cold Stable
4. cTKu – Continental Tropical Cold Unstable
5. mTWs – Maritime Tropical Warm Stable
6. mTWu – Maritime Tropical Warm Unstable
7. mTKs – Maritime Tropical Cold Stable
8. mTKu – Maritime Tropical Cold Unstable

B. Polar Air masses

1. cPWs– Continental Polar Warm Stable
2. cPWu– Continental Polar Warm Unstable
3. cPKs – Continental Polar Cold Stable
4. cPKu– Continental Polar Cold Unstable
5. mPWs– Maritime Polar Warm Stable
6. mPWu– Maritime Polar Warm Unstable
7. mPKs– Maritime Polar Cold Stable
8. mPKu– Maritime Polar Cold Unstable

UNIT – 4: TRI-CELLULAR MODEL

DEFINITION OF TRI-CELLULAR MERIDIONAL CIRCULATION

According to the old concept of the mechanism of general circulation of the atmosphere the movement of air is temperature dependent. In other words, temperature gradient causes air circulation on the earth's surface. According to the advocates of thermal school of the mechanism of general circulation of the atmosphere the tropical areas receive maximum amount of solar energy which substantially decreases poleward.

Thus, there is latitudinal imbalance of solar radiation from lower to higher latitudes. Consequently, there is transfer of heat through horizontal air circulation from the areas of high solar radiation (low latitudes) to the areas of low solar radiation (high latitudes) in order to balance the heat energy so that there does not exist too much heat energy in the low latitudes and too low heat energy in the high latitudes.

This old school considers only the horizontal component of the atmospheric circulation and does not consider the potential energy generated by unequal heating of the earth and its atmosphere and its continuous transformation into kinetic energy. It may be pointed out that the potential heat energy is continuously transformed into kinetic energy by the upward movement (ascent) and downward movement (descent) of heated and cold air respectively.

It may be remembered that the kinetic energy is also dissipated due to friction and small-scale atmospheric disturbances upward. Thus, it is necessary that there must exist balance between the rate of generation of kinetic energy and the rate of its dissipation due to friction. The modern concept of the mechanism of general circulation of the atmosphere, thus, includes both, the horizontal and vertical components of atmospheric circulation.

The modern school envisages a three-cell model of meridional circulation of the atmosphere, popularly known as tri-cellular meridional circulation of the atmosphere, wherein it is believed that there is cellular circulation of air at each meridian (longitude). Surface winds blow from high pressure areas to low pressure areas but in the upper atmosphere the general direction of air circulation is opposite to the direction of surface winds.

Apart from the surface and horizontal flows, winds have vertical and upper surface flows also. Such surface and upper surface flows combinedly form convective cell. Initially, assuming non rotating earth, unicellular convection of global atmospheric circulation extending equatorial to polar region was postulated but considering earth rotation, the atmospheric circulation is explained with the help of tricellur model. According to this model there is three cells circulation in both hemispheres. These are -

(a) Hadley cell or low latitude cell

(b) Ferrel cell or middle latitude cell

(c) Polar cell or high latitude cell

TROPICAL CELL

Tropical cell is also called as Hadley cell because G. Hadley first identified this thermally induced cell in both the hemispheres in the year 1735. The winds after being heated due to very high temperature at the equator ascend upward. These ascending warm and moist winds release latent heat after condensation which causes further ascent of the winds which after reaching the height of 8 to 12 kilometers in the troposphere over the equator diverge northward and southward or say poleward.

The surface winds in the name of trade winds blow from subtropical high pressure belts to equatorial low pressure belt in order to replace the ascending air at the equator. The upper air moving in opposite direction to surface winds (trade winds) is called antitrade. These upper air antitrades descend near 30°-35° latitudes to cause subtropical high pressure belt.

These antitrades after descending near 30°-35° latitudes again blow towards the equator where they are again heated and ascend. Thus, one complete meridional cell of air circulation is formed. This is called tropical meridional cell which is located between the equator and 30° latitudes. It may be pointed out that the regularity and continuity of the antitrade wind systems in the upper air has been refuted by a host of meteorologists on the basis of more upper air data being available during and after Second World War.

POLAR FRONT CELL OR MID-LATITUDE CELL

Polar front cell or mid-latitude cell-According to old concept surface winds, known as westerlies, blow from the subtropical high pressure belt to subpolar low pressure belt (60°-65°). The winds ascend near 60°-65° latitudes because of the rotation of the earth and after reaching the upper troposphere diverge in opposite directions (poleward and equator-ward).

These winds (which diverge equator-ward) again descend near horse latitudes (30°-35° latitudes) to reinforce subtropical high pressure belt. After descending these winds again blow poleward as surface westerlies and thus a complete cell is formed.

According to new concept of air circulation the pattern between 30°-60° latitudes consists of surface westerlies. In fact, winds blow from subtropical high pressure belt to subpolar low pressure belt but the winds become almost westerly due to Coriolis force. It may be mentioned that the regularity and continuity of westerlies are frequently disturbed by temperate cyclones, migratory extra-tropical cyclones and anticyclones.

Contrary to the existing view of upper air tropospheric easterly winds in the zones extending between 30°-60° latitudes Rossby observed the existence of upper air westerlies in the middle latitudes due to poleward decrease of air temperature.

According to G.T. Trewartha the middle and upper tropospheric westerlies are associated with long waves and jet streams. Warm air ascends along the polar front which is more regular and continuous in the middle troposphere. It may be pointed out that this new concept does not explain the cellular meridional circulation in the middle latitudes.

POLAR CELL

Polar cell involves the atmospheric circulation prevailing between 60° and poles. Cold winds, known as polar easterlies, blow from polar high pressure areas to sub-polar or mid-latitude low pressure belt. The general direction of surface polar winds becomes easterly (east to west) due to coriolis force.

These polar cold winds converge with warm westerlies near 60°-65° latitudes and form polar front or mid- latitude front which becomes the centre for the origin of temperate cyclones. The winds ascend upward due to the rotation of the earth at the subpolar low pressure belt and after reaching middle troposphere they turn poleward and equator-ward. The poleward upper air descends at the poles and reinforces the polar high pressure. Thus, a complete polar cell is formed.

Numerous objections have been raised against the concept of tri-cellular meridional circulation of the atmosphere. The temperature gradient should not be taken as the only basis for the origin and maintenance of cellular meridional circulation because not all the high and low pressure belts are thermally induced.

For example, the subtropical high pressure and sub-polar low pressure belts are dynamically induced due to subsidence and spreading of air caused by the rotation of the earth respectively. Upper air anti-trades are not uniformly found over ail the meridians. If the trade winds are exclusively of thermal origin, then the thermal gradient must be present boldly throughout the tropics but this is not true. At the height of 500 to 1000m in the atmosphere the winds become almost parallel to the isobars which are generally parallel to the latitude. If this is so, the meridional cell of air circulation may not be possible.

The pressure and winds in most parts of lower atmosphere are found in cellular form rather than in zonal pattern. These pressure and winds cells are elliptical, circular or semicircular in shape. These evidences (cellular form of air circulation) no doubt contradict the old concept of general pattern of atmospheric circulation but the cellular meridional circulation has not been fully validated.

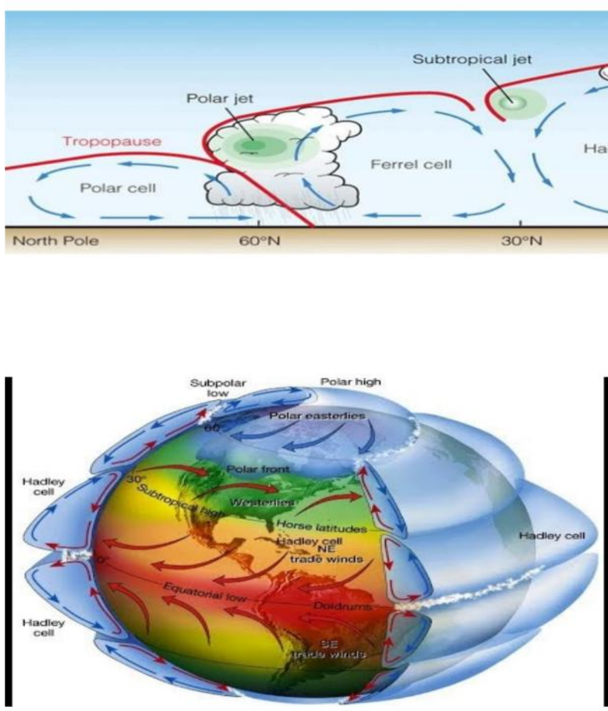


Fig: - tri-cellular Model & circulation winds on a rotating earth as a result of the three-cell model

UNIT – 5: ATMOSPHERIC FRONTS: CLASSIFICATION, FORMATION AND CHARACTERISTICS

INTRODUCTION

The literal meaning of the term 'front' is the forward margin of something. It is a sort of boundary between two air bodies. These two bodies are quite different from each other. It is not always possible to find definite boundaries between two different air masses. More often a gradual transition is observed for many kilometers. However when one air mass moves rapidly into another a distinct boundary becomes established. Such a boundary is called a front. Therefore, it can be stated that a narrow strip along two different bodies is generally termed as front.

Another word 'genesis' means beginning or method of formation or evolution of something. You must be aware that the air mass is divided into different groups based on different criteria. When two different bodies of air masses are coming into direct contact with each other, fronts are originated. Therefore, the process of the formation of front is known as frontogenesis. Frontogenesis is very commonly found in the higher latitudes. They have evolved a unique pattern over the globe and have modified the weather of the area under their influence very significantly. Opposite to frontogenesis, there is term called 'frontolysis' meaning the decaying of fronts. In this module, an attempt is made to study the process of the formation of fronts, their distribution as well as associated weather conditions.

DEFINITION OF FRONT

Fronts are three dimensional sloping boundary across which the horizontal gradients are sleek. In other words, fronts are defined as 'zone of mixing' or 'discontinuity between two air masses'. They have contrasting characteristics in terms of temperature, density, relative humidity, wind direction, and pressure. Fronts are interface between cold and warm air masses interacting along the temperate convergence zone or subtropical low which help maintaining latitudinal heat balance. Thus, it can be said the front is a line of discontinuity that represents the transitional zone between two air masses having contrasting characteristics. Frontal surface separates these two contrasting air masses.

The first real advancement in our understanding of the mid-latitude weather variations was made with the discovery that many day-to-day changes are associated with the formation and movement of boundaries or front, between air masses. The term 'front' was first proposed by a group of meteorologists which was led by Vilhelm Bjerknes who was working in Norway, during the First World War in 1918. The three meteorologists V.J. Bjerknes, H. Solberg, and T. Bergeron are considered important for creating an in-depth knowledge and advancement of the fronts. Their ideas are still a very integral part of the forecasting and weather analysis in the middle and the high latitudes. Later, in 1975 F.W. Cole defined fronts as an interface or transition zone between two air masses of different density. However, it should be noted that fronts are not only the transition zone between two air masses of different density. The transition zone might also be of different temperature, humidity, wind direction, and pressure.

FRONTAL ZONE

Frontal zone is the vast transitional zone between two converging air masses. It depicts the discontinuity zone of opposing and contrasting air masses. Frontal zones are inclined at a very low angle towards the ground and are never parallel or vertical to the ground surface. Frontal zone is a dynamic and mobile zone, therefore it drifts with prevailing winds.

FRONTOGENESIS

The term frontogenesis was introduced by Tor Bergeron (Norwegian meteorologist) for describing the formation of new fronts. However, later the term was extended to include the process of formation of both new fronts and the regeneration of the old and decaying fronts that already exists in the atmosphere. The region of frontogenesis represents the region that has the convergence of the contrasting air masses. On contrary, we use the term frontolysis means the destruction or dying of fronts. Hence, frontolysis is the opposite of frontogenesis. Fronts do not come into existence all of a sudden; rather they appear only after the process of frontogenesis has been in operation for quite some time. The act of weakening or vanishing of the existing fronts is not accompanied suddenly. The process of frontolysis must continue for some time in order to destroy an existing front.

CONDITIONS FOR FRONTOGENESIS

When the distribution of frontogenesis was studied, it was noticed that this process doesn’t occur everywhere. Frontogenesis requires certain characteristics for processes to occur. They are:

TEMPERATURE DIFFERENCE: The two opposing air masses that converge to form a front must have a contrasting temperature. If one air mass is warm, moist and light, a front can only be created when the other air mass is cold, dry and dense. What happens when these two contrasting air masses converge? The cold and dense air mass obviously invades the area occupied by the warm and light air mass. As a result, the warm and light air mass is pushed upward and this generates a front. This also explains why no frontogenesis takes place in the equatorial region. Two air masses also converge at the equator (trade winds) but the temperature of both these air masses is uniform. So, the temperature difference is the missing factor here because of which no frontogenesis takes place at the equator.

CONVERGENCE OF AIR MASSES: In the very definition of frontogenesis, the word 'convergence' has been used and thus it is understood that it is the pre-requisite for the frontogenesis process. When two air masses having different temperature converge, they try to invade the space of each other and this leads to the formation of the fronts. For frontolysis to occur, the air masses have to diverge or get diluted by mixing and the contrast being removed.

DEVELOPMENT OF FRONTS

When two contrasting air masses converge and spread horizontally along the axis of dialation, the probability of creation of front increases. However, there are other necessary conditions for the actual formation of the frontogenesis. When the convergence occurs, the formation of the fronts depends on the angle between the axis of dialation and the isotherms. Fronts are not formed when this angle is more than 45 degree. Frontogenesis is only activated when the angle between the axis of dialation and isotherms decreases. Due to this decrease in angle, a point comes when the isotherms become almost parallel to the axis of dialation and this is when the frontogenesis takes place.

The intensity and steepness of the front depend on the temperature gradient of the two air masses. A stationary front is formed when convergence between two air masses parallel to each other occurs. In such a situation, no upward displacement of the air occurs. Such kind of stationary front doesn't result in cloud formation and precipitation. Stationary fronts are rarely formed because a slope mostly occurs in the fronts due to the Coriolis force.

FRONTOLYSIS (DISSIPATION OF FRONTS)

Frontolysis (frontal decay) represents the final phase of a front's existence, although it is not necessarily linked with occlusion. When contrasting air masses lose their characteristics and difference decay occurs. In other words, when the air masses move away from each other or when the temperature contrast between the adjacent air masses diminishes due to one reason or another the fronts dissipate or start declining.

The dissipation of fronts takes place in three ways: (a) through front's stagnation over a similar surface; (2) as a result of both the air masses cold and warm moving on parallel tracks at the same speed; (3) by the system entering air of the same temperature. Frontolysis happens in the area of Siberia, Northern America etc.

FRONTAL CHARACTERISTICS

The character of frontal weather depends upon the vertical motion in the air masses. If the air in the warm sector, it is rising relative to the frontal zone. The fronts are very active and are termed *ana-fronts*. The term "ana-front" describes boundaries which show instability, meaning air rises rapidly along and over the boundary to cause significant weather changes. Sinking to warm air relative to cold air masses gives rise to less intense *Kata-fronts* (Berry *et at.*, 2010 pp-238). "*Kata*-front" is weaker, bringing smaller changes in temperature and moisture, as well as limited rainfall. Fronts may differ in their types, location or a real extent, but the following characteristics are more or less common to all of them.

Fronts are the zone of mixing of cold and warm air masses and it covers certain area. The area and extent which is covered by the fronts vary in time and space ranging from 3D boundary extending up to tropo-pause.

Apart from area, slope is one of the most important characteristics of the front. In general, cold fronts are steeper than warm fronts. The frontal slope depends upon latitude of the front, wind speed, and temperature difference between the air masses. Cold air always run under the warm air, when the slope between them is steep, the large vertical motion will occur because of the intense lifting of warm air mass. This will result in intense weather. The factors which are responsible for steeper slopes are large wind velocity, difference between air masses, temperature difference and latitude.

Isobars and isotherms develop ‘kink’ (bend) in the frontal zone. Because of an abrupt change in pressure as well as pressure gradient across a front, the isobars while crossing it from a kink and always bend towards the direction of low pressure. The wedge formed by the bending of isobars across a front always points towards the higher pressure. Mostly the fronts lie in the trough of low pressure. Temperature is the principal property used to identify air masses and fronts between them. Large differences in air temperature are recorded across a front. But the changes in temperature may be abrupt or gradual depending on the nature of opposing air masses. Further, the width of the frontal transition zone is dependent on the temperature contrast between two air masses. Larger the difference in temperature, thinner the frontal zone and vice versa. Winds always blow from higher to lower pressure crossing the isobars at an acute angel as wind movement is always controlled by pressure gradient and Coriolis force. It is, therefore, natural that an abrupt wind shift must occur at the fronts. On common parlance it can be stated that sudden shift in the direction of wind is observed (i.e. wind direction is not fixed). The frontal intensity is determined by the turbulence. Both turbulence and gustiness are required to make the front strong. Gustiness results into convective phenomena like strong winds and thunderstorm. Frontal activity is invariably associated with cloudiness and precipitation. Since warm air moves up along the frontal surface, it cools adiabatically which results in cloudy condensation and precipitation. The type of clouds and precipitation falling from them depend on the slope of the front and the amount of moisture in sending mass of air. The weather condition is also affected by the speed of the front. The rapidly moving fronts create intense weather conditions in comparison to slower moving fronts. Fronts are the agents of latitudinal exchange of heat and moisture as frontal zone is a mobile zone therefore it drifts with prevailing winds. Fronts are favorable place for cyclogenisis and have generally a life cycle of 4 to 7 days after that they start decaying.

CLASSIFICATION OF FRONTS

The types of front depend on the characteristics of the front, direction of the movement of the air mass, mechanism of frontogenesis and associated weather. Fronts are classified into four main types: -

- The Warm Front
- The Cold Front
- The Stationary front
- The Occluded Front

WARM FRONT: - Warm front is that gently sloping frontal surface along which warm and light air becomes active and aggressive and rises slowly over cold and dense air. The average slope of warm fronts in middle latitudes ranges between 1:100 to 1:400. The gradually rising warm air along the gently sloping warm front is cooled adiabatically, gets saturated and after condensation precipitation occurs over a relatively large area for several hours in the form of moderate to gentle precipitation.

COLD FRONT: Cold front is that sloping frontal surface along which cold air becomes active and aggressive and invades the warm air territory and being denser remains at the ground but forcibly uplifts the warm and light air.

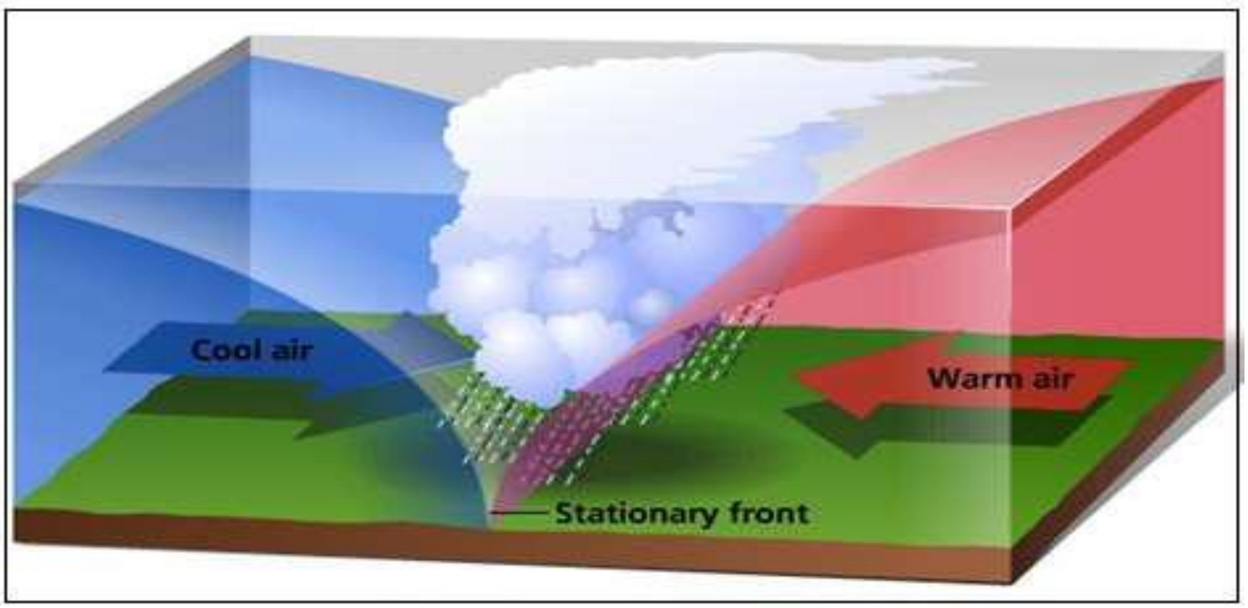
Since the air motion is retarded at the ground surface due to friction while the free air above has higher velocity and hence the cold front becomes much steeper than warm front. This is why the slope of cold front varies from 1:50 to 1:100 (which means the rise of the wedge of cold air at the rate of one kilometre for every 50 to 100 kilometres).

A cold front is associated with bad weather characterized by thick clouds, heavy downpour with thunderstorms, lightning etc. Sometimes, cold frontal precipitation is also associated with snowfall and hailstorms.

OCCLUDED FRONT: Occluded front is formed when cold front overtakes warm front and warm air is completely displaced from the ground surface.

STATIONARY FRONT: Stationary front is formed when two contrasting air masses converge in such a way that they become parallel to each other and there is no ascent of air. In fact, the surface position of stationary front does not move either forward or backward.

Figure: - Formation of Stationary front



Source: https://www.pmfias.com/wp-content/uploads/2016/01/stationary-front-.jpg

CHANGING WEATHER ASSOCIATED WITH FRONTS: Since fronts are formed due to convergence of two air masses of contrasting temperatures and hence contrasting weather conditions are found from north to south or south to north. Differences in terms of temperature, humidity, precipitation, cloudiness, and wind direction are experienced along different fronts e.g., warm and cold fronts.

WEATHER ASSOCIATED WITH WARM FRONTS: Warm air becomes active and aggressive along warm front as it invades cold air zone and thus being lighter it gradually rises over cold air and is cooled adiabatically from below. Cooling of warm air causes condensation and cloud formation followed by precipitation. If the aggressive warm air is stable and less humid, condensation occurs at great height and hence much lifting of air is required.

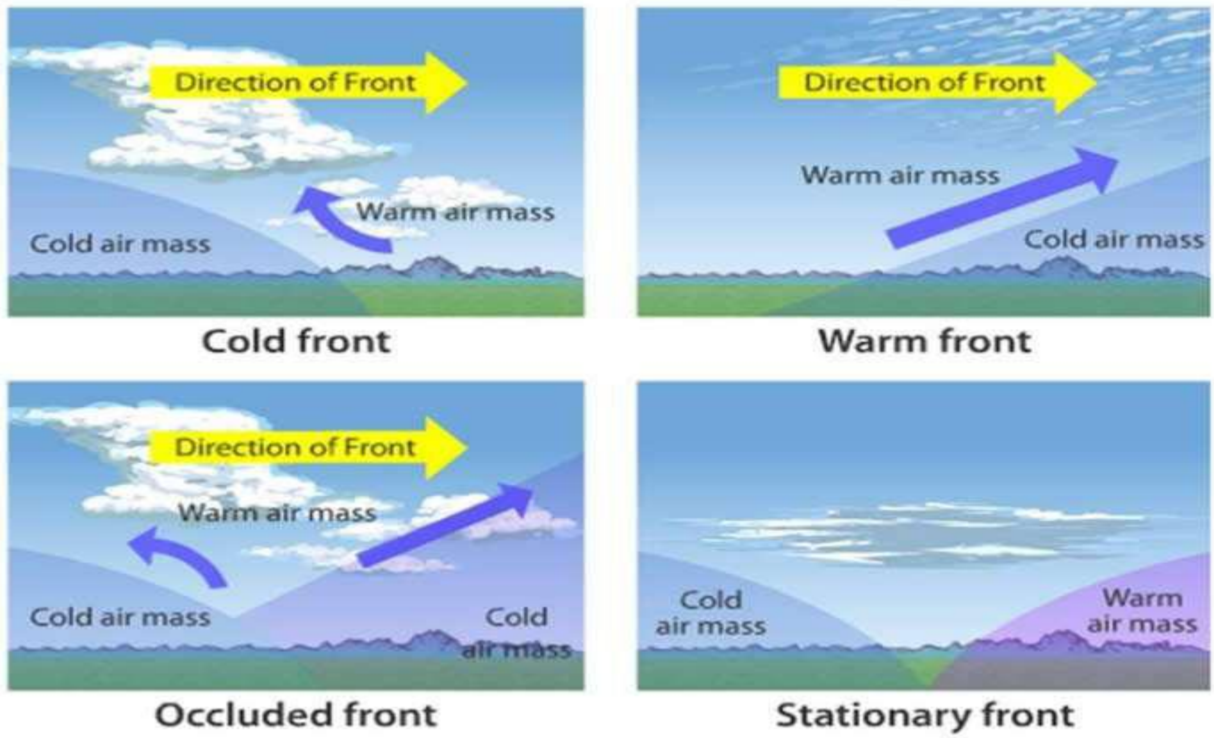
On the other hand, if the warm air is moist and unstable, only a slight lifting causes condensation and precipitation. The warm front precipitation is of long duration, moderate but widespread because of gentle slope of warm front. There are frequent changes in cloud types. The sequence of clouds from above downward comprises cirrus, cirro- stratus, alto-stratus and nimbo-stratus.

When warm front advances forward, the warm sector comes over the observation place. There is sudden change in weather conditions with the arrival of warm sector e.g., sudden increase in temperature and specific humidity, decrease in air pressure, disappearance of clouds, clear sky, and break in precipitation.

WEATHER ASSOCIATED WITH COLD FRONT: Cold and dense air becomes active and aggressive along cold front wherein cold air invades warm air region and pushes it upward while it, being denser, settles downward. If cold front passes away soon, weather also becomes clear soon, otherwise if the front becomes stationary, the sky becomes overcast with cumulo-nimbus clouds provided that cold air is moist and unstable, and frontal thunderstorms are formed. Heavy precipitation occurs but is of short duration.

The consequent weather is characterized by decrease in air temperature, increase in air pressure, decrease in specific and relative humidity and change in wind direction from 45° to 180°. Precipitation is accompanied by lightning and cloud thunder. Sometimes, rainfall is associated with hailstorms. After the passage of cold front, clouds disappear, precipitation terminates and weather becomes clear and north-west cold winds set in.

FIG: - TYPE OF FRONTS



Source: https://socratic.org/questions/what-are-the-main-types-of-fronts

ADIABATIC LAPSE RATE

For condensation process, one of the pre requisite conditions is cooling of air. Cooling of air may occur in diabatic process and adiabatic process. In adiabatic process of cooling there is no addition or subtraction of heat is involved, only expansion of air due to decrease of pressure leads to cooling of air. Adiabatic change of temperature is of two types-

i. Dry adiabatic change

ii. Wet adiabatic change

DRY ADIABATIC CHANGE: When a parcel of air ascends, it passes through the successive lower pressure region. So the air parcel expands and cools. Conversely, descending air parcel heats up. This type of change of temperature of unsaturated air (dry air) is called dry adiabatic lapse rate. The average rate of this change of temperature is 10^0C per km.

WET ADIABATIC CHANGE: The rate of temperature change of ascending air beyond the condensation level (saturated air) is lowered down due to release of latent heat of condensation. This type of change of temperature of saturated air is called wet adiabatic lapse rate. It varies from 3^0 C to 6^0C per km depending on the temperature and moisture content of air.

UNIT-6: FORMATION OF PRECIPITATION; RAINMAKING

INTRODUCTION

Precipitation is any liquid or frozen water which forms in the atmosphere and falls back to the surface of the Earth. It comes in many forms, such as rain, sleet, and snow; however, it all comes from the same general process. Along with the processes of evaporation and condensation, precipitation is amongst the three major parts of the global water cycle that are important for the sustainability of life on Earth. In this article, we will learn about the precipitation meaning, what are the different forms of precipitation, what is the process of precipitation, classification of precipitation, and examples of precipitation in detail. Let us first have a look at what is precipitation.

The basic role in cloud formation and precipitation is played by the following dynamic factors: vertical motions, advective and turbulent inflows of heat and water vapour. With upward motion of air observed in the low-pressure areas (cyclones and troughs), the air temperature and mass fraction of water vapour at fixed levels are lowered, the relative humidity increases in time: water vapour comes nearer to a saturation condition. After achievement of saturation, the further rise of air is accompanied by cloud formation, increase in size of drops (crystals) under the influence of condensation (sublimation), and then also of coagulation. This process culminates in cloud precipitation - rain, snow, graupel, hail.

The type and intensity of precipitation significantly vary depending on what thermal stratification is observed near the condensation level when the saturation condition is achieved. With moist-steady stratification, the vertical velocity in the derived cloud remains the same as before its formation. In this case, a nimbostratus cloud forms producing steady rain characteristic of warm fronts.

Since with the upward motion the vertical gradient of temperature in the lower troposphere increases with time, it is also quite possible that moist-labile stratification could be registered in a cloud. Under its influence, mesoscale vertical motions originate in a cloud, their speed exceeding that of synoptic scale motions by one or two orders of magnitude. The mesoscale motions result in formations of cumulonimbus and showers. An essential role is played by the moist-adiabatic gradient. As the latter decreases with the temperature rise, with the conditions being the same (depth of a cyclone, initial gradient of temperature), the probability of formation of moist-labile stratification and showers is the greater, the higher the air temperature. This effect plays a relevant role in increase of probability for the formation of convective clouds and showers in the afternoon and summer, as contrasted to the night period and winter.

The essential influence on formation of precipitation is rendered by entrainment of the ambient air into a developing cloud. Despite its smaller duration, shower-type precipitation makes up the largest proportion (exceeding 70-80 percent) of its total quantity.

The role of the baroclinic factor - geostrophic advection of virtual temperature - is great in formation of synoptic vortexes (cyclones and anticyclones). With an advection of colder and drier air, cyclones would form; with an advection of warmer and moister air, anticyclones. The influence of this factor explains geographic features of precipitation distribution, as well as their seasonal fluctuations and connection with fields of temperature and pressure. The microphysical processes of precipitation formation are briefly reviewed: condensation and sublimation of water vapour, coagulation of cloud drops and ice crystals.

BASIC INFORMATION ABOUT PRECIPITATION

The formation of precipitation is intimately connected with the formation of field of clouds. Among those phenomena and values which are integrated by the concept of "weather", the clouds and precipitation possess a determining role.

Changing the thermal and radiation mode of the atmosphere, clouds and precipitation render essential influence on many sides of human activities, first of all, in the sphere of agricultural production, as well as on flora and fauna of the Earth. Visibility is sharply diminished by precipitation, clouds and fog. Therefore, different types of transport, primarily aircraft, are greatly affected by these phenomena.

Cloud is a visible set of drops of water and fragments of ice suspended in the atmosphere and located at some altitude1 above the earth's surface. If the lower boundary of this combination reaches the surface of the Earth, it is defined as fog (meteorological visibility less than 1 km) or mist (with visibility in the range from 1 to 10 km).

The drops of water and ice crystals falling out of the atmosphere on the earth's surface are called precipitation. The amount of precipitation is measured by the depth of liquid water (most often, in millimetres) that could be derived after falling on a horizontal impermeable surface: 1 mm corresponds to 1-kg water mass which has dropped out on 1 m2.

The amount of precipitation which has fallen in a unit time (for example, 1 hr or 1 min) is referred to as intensity of precipitation.

Generically, i.e., depending on physical conditions of formation, precipitation can be subdivided into the following types:

- widespread precipitation, prolonged and spreading over a large area, of moderate intensity (0.01-0.05 mm/min) falling from Nimbostratus clouds (Ns) as rain and snow, sometimes sleet;
- shower-type precipitation, falling from Cumulonimbus clouds (Cb) as rain, snow, graupel, hail; this precipitation suddenly begins and comes to an end; it is characterized by a sharp change in intensity, exceeding, as a rule, 0.05 mm/min; it is frequently accompanied by thunderstorms and squalls.

RELATIONSHIP BETWEEN PRECIPITATION AND CLOUDS

Though the duration of shower-type precipitation is much less than that of the widespread type, it makes the greatest contribution to the total amount of precipitation. So, in the European part of Russia the average share of shower precipitation makes up 73% in the warm half-year (April - October). At 11 stations (out of 26), this share exceeds 80 %, and at the foothills of the Caucasus it is more than 90 %.

Formation of precipitation and its intensity are related to microphysical structure and thickness of clouds. It follows from the data collected during 686 aircraft soundings of frontal clouds that the greatest recurrence (about 90 per cent) falls on the precipitation from clouds of totally mixed structure, throughout the entire thickness or in their considerable part. The precipitation falls from purely crystalline clouds and clouds consisting of separate strata in approximately 70 % cases; from purely water clouds, only in 9 % cases.

If the recurrence of precipitation is correlated with the total number of cases of cloud observation, the purely water clouds would produce 5 % of all cases with precipitation reaching the surface of the Earth; the clouds with separated strata, 55 %; and the clouds of wholly mixed structure throughout the entire thickness, 40 %.

There is a practically linear relationship between intensity of precipitation and thickness of clouds. The data obtained from 439 soundings indicate that drizzle falls from clouds with average thickness zcl equal to 850 m; rain with drizzle at zcl = 1400 m; rain, at zcl = 2150 m; snow, at zcl = 2300 m; snow with a rain, at zcl = 2690 m; and rain formed through the melting of ice particles, at zcl = 3150 m.

MECHANISM OF PRECIPITATION

BERGERON-FINDEISEN PROCESS: - The Bergeron-Findeisen process is based on two facts.

(a) Co-existence of water vapour, ice crystals and super cooled liquid water droplets, and

(b) Different values of saturation vapour pressure.

In cold clouds, where the ambient temperatures are below freezing point, the condensation products can be both liquid water and ice crystals. Below - 40°C all the products are ice crystals and all the cloud is said to be glaciated. Between 0°C and - 40°C water and ice coexist giving a mixed cloud.

The Bergeron-Findeisen Process is concerned with the growth of raindrops out of mixed clouds. In mixed clouds, the initial growth phase depends on the co-existence of ice and water.

The super cooled water droplets freeze when –

(i) they are disturbed or

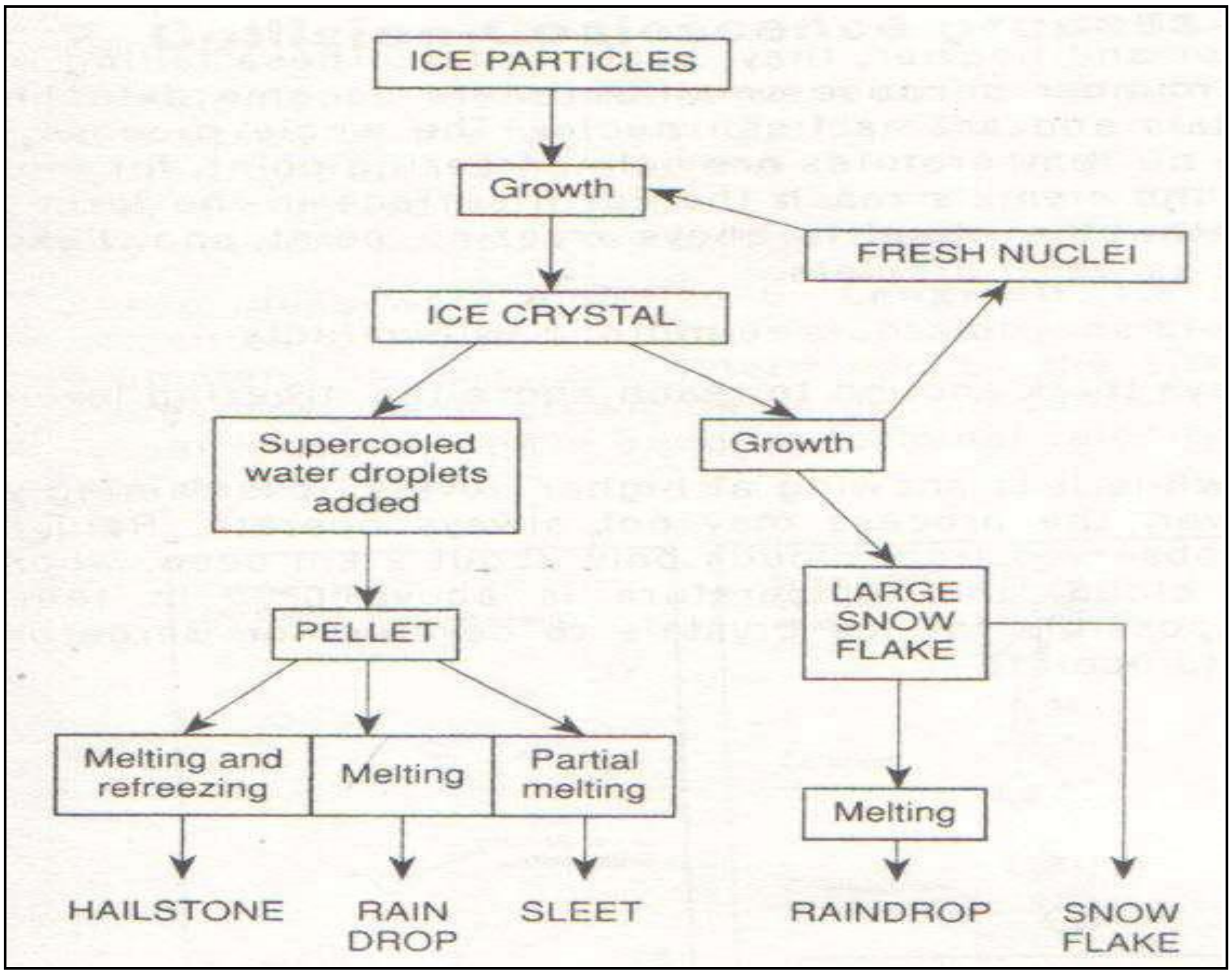
(ii) they come into contact with a freezing nuclei .

The nuclei that form water droplets far outnumber the nuclei that form ice crystals so at cloud temperatures above - 20°C, (the activation temperature for many deposition nuclei), super cooled water droplets are initially much more abundant than ice crystals. In fact, single ice crystals may be surrounded by thousands to hundreds of thousands of super cooled water droplets. Thus, when ascending air currents rise, the temperature in the ascending air currents drops well below - 20°C and the super cooled droplets freeze. This single ice crystal has the capability to change the whole of the cloud to an ice cloud. This is because the saturation vapour pressure over ice is lower than over water at the same temperature. This means that when air is saturated with respect to water, it is super saturated with respect to ice.

At a temperature of - 10°C this vapour pressure translates into a relative humidity of 100 per cent (saturation) for the air surrounding water droplets and a relative humidity of 110 per cent (super saturation) the air surrounding ice crystals. In response to super saturated conditions water vapour migrates towards ice crystals and deposition starts on ice crystals (sublimation nuclei) and the ice crystals grow and collect together to form snowflakes. Deposition removes water vapour from the clouds and thereby, lowers the relative humidity. The relative humidity of the air surrounding water droplets goes below 100 per cent and the droplets vaporise. Under these conditions snowflakes grow at the expense of super cooled water droplets. In addition, the growing ice crystals may come into contact with super cooled liquid water droplets, which will freeze onto or around the crystals immediately in contact. This type of growth process is known as riming. Riming is a primary mechanism of hail formation.

As the crystals grow larger and heavier, they begin to fall. These falling ice crystals also increase in number because small splinters become detached from large dendritic crystals and act as fresh nuclei. The whole process is thus repeated again. If the air temperatures are below freezing point, most of the way to the ground, the crystals reach the earth surface in the form of snowflakes. If the air below the cloud is above freezing point, snowflakes melt and reach the ground as water droplets. However, the process may not always operate. Rain in tropical regions is observed from clouds only about 2 km deep, where at the top of the cloud, the temperature is above 0°C. In these conditions, it is impossible for ice crystals to develop for Bergeron-Findeisen process to operate.

FIG. : BERGERON-FINDEISEN PROCESS



COLLISION-COALESCENCE PROCESS: - Collision coalescence ideas were put forward by George Simpson and Mason. It is based on the fact that clouds contain a variety of sizes of water droplets. Uniformly small droplets tend to move at the same speed in the cloud but if they are mixed with larger, slow moving droplets that have formed around hygroscopic nuclei, then this will encourage collisions and amalgamations. These ideas were modified by Langmuir. He pointed out that the terminal velocities of falling drops are directly related to their diameters. These diameters, in turn, are determined by the size of the condensation nuclei. Drops that have grown on large condensation nuclei become larger. The larger drops will have a higher terminal velocity than the smaller ones and so collide with them.

Raindrops or ice crystals often stick after colliding and thus, grow in finite steps, i.e., collide and coalesce. The larger drops then fall faster, as they overcome air resistance more easily. By falling faster, they are able to catch up even more rapidly with other droplets and crystals so that the larger they grow, the faster they grow. For collision to occur, several conditions are necessary.

(a) The smaller droplets must be close to the axis of fall of larger drop, otherwise it will follow the air currents around the falling drop and there will be no impact. As the larger droplets descend, they produce an air stream, a miniature air current that blows the small particles out of their path, i.e., smaller drops are swept away.

(b) Colliding ice crystals bounce off each other when they are plane in shape and dry and cold. In other words, colliding droplets may bounce away from each other as there is little surface tension.

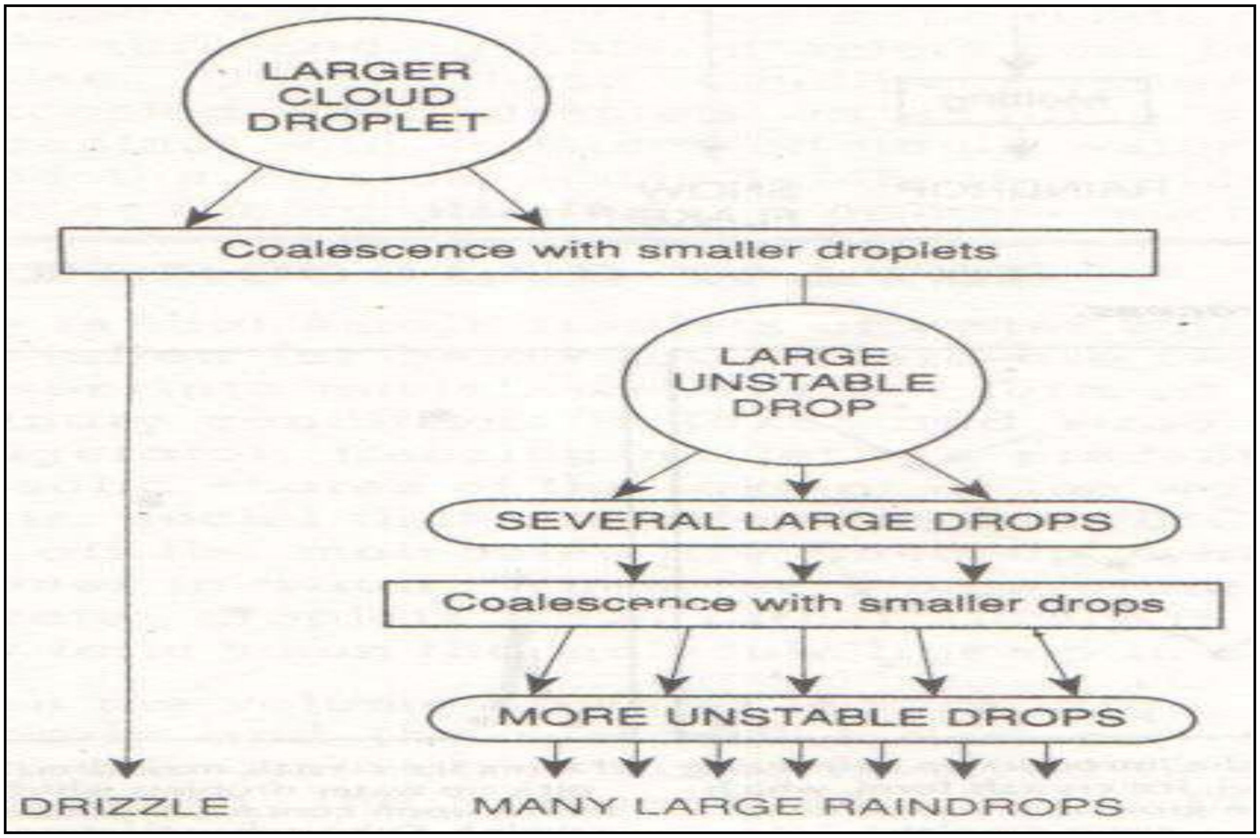
(c) Even with collision, growth will only occur if the two drops coalesce. This will occur most readily if

(i) The drops are of considerably different sizes, and

(ii) Atmospheric electricity is present to hold the droplets together. If a droplet with a negative charge should collide with a positively charged droplet, their electrical attraction will bind them together.

Continued collision, thus, leads to coalescence resulting in many large unstable drops which on further disruption produce several large drops and its continued coalescence and further disruption leads to many larger drops.

FIG. : COLLISION-COALESCENCE PROCESS



FORMS OF PRECIPITATION

Precipitated moisture falling on the ground takes various forms which depend on the following conditions: a. the temperature at which condensation takes place, b. the conditions encountered as the particles pass through the air, c. the type of clouds and their heights from the ground and d. the processes generating precipitation. All forms of precipitation regardless of appearance are collectively termed 'hydrometeors'. Hydrometeors have been classified into 50 specific types, of which only the more common types will be examined here.

The different forms of precipitation are as follows:

1. RAIN

Rain is a form of precipitation which is in the form of water drops of a size that is larger than 0.5mm. The maximum raindrop size is about 6mm. Drops of larger size break up into smaller drops as it falls down on the Earth's surface.

Rainfall is the predominant form of precipitation and therefore, the term precipitation is used synonymously with rainfall. The magnitude of the rainfall shows high temporal and spatial variation. This variation causes the occurrence of hydrologic extremes like floods and droughts.

2. SNOW

Snow consists of ice crystals in a flaky form, having an average density of 0.1g/cc. It is also an important form of precipitation that usually forms in colder climates and higher altitudes.

3. DRIZZLE

Drizzle is a fine sprinkle of tiny water droplets that have a size less than 0.5mm and intensity greater than 1mm/h. The tiny drops that form a drizzle appear floating in the air.

4. GLAZE OR FREEZING RAIN

The glaze is formed when rain or drizzle comes in direct contact with the cold ground at around 0 degrees Celsius. This water drops freeze to form an ice coating known as glaze.

5. SLEET

Sleet is frozen raindrops that are formed when rainfall passes through the air in the atmosphere at subfreezing temperatures.

6. HAIL

Hail is a kind of showery precipitation in the form of pellets or lumps that have a size greater than 8mm. Hail occurs during the violent thunderstorms.

These are the examples of precipitation. Let us now look at what are the different types of precipitation.

TYPES OF PRECIPITATION

Precipitation occurs when the moist air mass undergoes the process of condensation. This process occurs when the air is cooled and saturated with the same moisture amount. This process of cooling the air mass occurs only when the air mass moves up to the higher altitudes. The air mass can be lifted to higher altitudes primarily by three methods based on which there are three different types of precipitation which are as follows:

Cyclonic Precipitation

Convective Precipitation

Orographic Precipitation

1. **CYCLONIC PRECIPITATION**: - Cyclonic rainfall is occurs when air mass rise up due to pressure difference. When there is the formation of a low-pressure area, air from the other surrounding spaces flows to less pressure zone.

It is phenomenon forces warm and colder air to meet. As warm air is lighter in comparison to colder ones it rises above the colder air. Then the warmer air starts cooling beyond saturation point which results in heavy rain. Such rainfall is called Cyclonic Rainfall.

The cyclonic rainfall itself is categorized into two types,

- Frontal Rainfall
- Non-Frontal Rainfall

Frontal Precipitation: - The collision of two different air masses, due to differences in temperature and densities, results in condensation and precipitation at the contact surface. This contact surface is named as Frontal Zone, whereas the precipitation occurring in that zone is called Frontal Rainfall.

If cold air driving out warm air then the front is called Cold Front and if the Warm nature masses of air overlap the superior cold air mass, it can be known as 'Warm Front. Similarly Stationary Front is the result of both air masses are move towards an area of low pressure in a simultaneous way.

NON FRONTAL RAINFALL: - This phenomenon occurs when the moving mass of cold air meets the stationary warm nature of air mass. As a result of which the lighter warm air rises up and reaches the saturation point. The saturation point of lighter warm air causes precipitation known as Non-Frontal Precipitation. It is the type of Cyclonic Precipitation. This process differs from Frontal Precipitation in some sense hence categorized as the different type of Cyclonic Precipitation**.**

2. **CONVECTIVE RAINFALL**: - Convective Precipitation or Convectional Rainfall generally occurs in equatorial areas. The surface areas in these zones get heated frequently and constantly because of the sun's heat. Which in turn the air near the surface gets hated and spreads. Heating also makes the air lighter hence it tends to rise up. The air starts cooling as it moves up and reaches its saturation limit, resulting in precipitation. This precipitation is called Convention Precipitation, which occurs in the tropics that took on a hot day. Rising vertical velocities of air currents are dangerous to aircraft. Its intensity ranges from light showers to cloud bursts.

3. **OROGRAPHIC PRECIPITATION**: - Orographic Precipitation occurs when the moist mass of air, strikes natural barriers of the topography area. These barriers are like mountains, hills, etc which causes the air to rise up, condense and then precipitate. Hence mountains are the sites of higher precipitation than plain lands. The precipitation is concentrated on side of windward, the

side where striking occurs. Whereas, the side of the leeward gets a very little amount of rain, which results in the formation of a rain shadow region.

UNIT-7: EL NINO, SOUTHERN OSCILLATION AND LA NINA

EFFECTS OF EL NINO AND SOUTHERN OSCILLATION (OS)

During non El Nino period, high pressure in Peru coast and low pressure in Australian coast and wind flows from east to west (for details see 1.4.7) and moisture laden remnants air moves towards Indian ocean and energise Indian summer monsoon .

But, during El Nino period pressure cells transpose i.e. low pressure in Peru coast and high pressure in Australian coast leads to eastward movement of wind. As a result, Australian coast suffers from moisture supply and there is no supply of moisture to Indian subcontinent and Indian summer monsoon weakens. It is also the fact that during El Nino period, warm El Nino current in Peru coast move further southwards and crossing its normal limit 36^0S latitude, it merge with west wind drift and raises the temperature of southern ocean. Hence, increasing temperature weakens the high pressure formation of Indian Ocean as well as summer monsoon. This situation in winter energizes the low pressure of Indian Ocean and thereby north eastern monsoon gets momentum.

But, weakness and scanty rainfall of Indian summer monsoon not only for El Nino rather there are other factors also. This statesman comes from the instances that the monsoon reveals weakness and scanty rainfall in non El Nino periods also.

EL NINO, SOUTHERN OSCILLATION AND LA NINA

Both sea surface temperature and atmospheric circulation strongly influenced by each other and the result of this interaction is Walker circulation and Southern oscillation.

In normal situation, there is surface high pressure in Peruvian coast i.e. in equatorial eastern Pacific Ocean and low pressure in eastern Australian and Indonesia coast i.e. in equatorial western pacific coast and trade wind strongly blows from east to west and drive the surface warm water westwards and it help in upwelling of cold water along the equator in the eastern Pacific Ocean and piling up warm water in west pacific coast. Warm water heats the atmosphere, upward limb of air over western coast and subsidence limb over Peru coast. Such ocean atmospheric circulation is known as *Walker circulation* after the Sir Gilbert Walker.

But, when the walker circulation become weak that means transposition of high and low pressure between eastern and western pacific is created, a reversal situation is observed. In this situation, trade winds are weakened which allows the warmer western pacific waters to move east and reach the Peru coast and increases the temperature of the eastern pacific coast, sea level rises up, makes the thermocline (layer between surface water and deep water) deeper and convection cell shift towards eastern Pacific Ocean and a band of warm water flow southward which replaces the northward cold Peruvian current-such phenomena is called *El Nino.*

Such periodic transposition of pressure cell as well as the transposition of ocean and atmospheric circulation is known as southern oscillation (SO) as it brings turbulence in meteorological phenomena in southern ocean. Entire event of El Nino and related southern oscillation is known as *ENSO or El Nino Southern Oscillation.*

La Nina is the opposite phenomena of El Nino. In La Nina episode easterly trade winds again strengthen, convection cell shift far westward and upwelling of cold water increases in Peru coast and represent a below average sea surface temperature. Such reversal situation is known as *Non ENSO or Non El Nino Southern Oscillation***.**

UNIT-8: MONSOON: THEORIES OF ITS ORIGIN; RECENT TRENDS OF MONSOON IN INDIA

MONSOON: THEORIES OF ITS ORIGIN

Monsoon means (Arabic word 'mausim' or Malayan word 'monsin' means season) seasonal reversality of winds at least twice in a year. Monsoon may be true (directional reversality along at 180^0 axis) or pseudo monsoon (no definite 180^0 reversal direction). The origin of monsoon is explained in the light of different theories. Such as-

THERMAL THEORY

Origin of monsoon is explained by the unequal heating and cooling of land and water by Hailey in 1686. He explained that in summer during summer solstice there is high pressure over Indian Ocean and low pressure in Central Asia. As a result, moisture laden air blows from sea to land i.e. called south west monsoon or summer monsoon. But in winter (winter solstice), the situation become reverse as high pressure develop over Central Asia and low over the ocean and that's why, wind start to blow from land to sea and it is called north eastern or winter monsoon.

AIR MASS THEORY

H. Flohn, German meteorologist opined(1951) that monsoon is nothing but the modified trade winds and the modification is caused by the shifting of pressure and wind belts due to seasonal migration of the sun. Tropical convergence is formed due to convergence of southeast trade winds in the southern hemisphere and the northeast trade winds in the northern hemisphere. The meeting place of these winds is known as the Inter-Tropical Convergence Zone (ITCZ). The location of ITCZ shifts north and south of equator with the shift of pressure belts. In the summer season, the sun shines vertically over the Tropic of Cancer and the ITCZ shifts northwards. The southeast trade winds of the southern hemisphere cross the equator and start flowing from southwest to northeast direction under the influence of coriolis force. These displaced trade winds are called south-west monsoons when they blow over the Indian sub-continent. The front where the south-west monsoons meet the north-east trade winds is known as the Monsoon Front. Due to southward shifting of pressure and wind belts during winter, north east trades reestablished to its original position and become winter or north east mosoon (Cited in Singh, 2005).

Apart from the unequal heating and shifting winds, monsoon also influenced by other factors like shape of continent, orography, surface and upper air circulation. Therefore, other theories are put forwarded to explain the origin of monsoon. M.T. Yin (1949), P. Koteswaram (1952), Pierre Pedelaborde (1963) opined that monsoon has a relationship with Tibet plateau, upper air circulation, El Nino and southern oscillation etc.

EFFECTS OF JET STREAM ON MONSOON:

Jet stream is upper air, fast moving strong westerly wind, blows from west to east in middle latitude. In winter, the subtropical jet stream is bifurcated by the Himalayan ranges and Tibetan Plateau and southern branch positioned to the south of the Himalayan ranges and create an upper air anti-cyclonic situation. So, the surface high pressure (thermally originated in winter) and upper high pressure strongly obstructs the invasion of monsoon. But, in summer, the subtropical jet stream shifted to the north of Himalaya and creates vacuum at the upper atmosphere. Ground level low pressure develops due to thermal causes. These combined low pressures invite the landward invasion of moisture laden air mass as south west monsoon.

Besides, the subtropical westerly jet to the north of the Himalayas, there is the tropical easterly jet stream (June to September) over the peninsular India. This jet exert the Bay of Bengal branch of monsoon towards west and thereby monsoon forced to penetrate over the northern part of the sub-continent.

EFFECTS OF TIBET PLATEAU

Dr. P. Koteswaram, the director general of Indian Observatories when attended an international seminar on 'The Monsoon of the World' opined that summer time heating of Tibetan plateau is responsible for the origin of monsoon. Approximately 4000m high Tibetan plateau acts as a high level of heat source and physical barrier. During summer, sensible heat and latent heat released from remnants of pre-monsoon thunderstorms make the plateau into an elevated heat island and this heat transfer to the upper air. The rising air parcel diverges southwards and sinks in the equatorial part of Indian Ocean and energises the high pressure of Indian Ocean and onset of summer monsoon.

RECENT TRENDS OF MONSOON IN INDIA

INTRODUCTION

The monsoon system in the Indian sub-continent differs considerably from that of the rest of Asia. The centers of action, air masses involved and the mechanism of precipitation of the Indian monsoons are altogether different from other monsoon systems. Although pseudo-monsoons or monsoonal tendencies develop over other parts of the world, it is only around the Indian Ocean that monsoonal circulation in the true sense of the term is observed. Here the monsoons appear as truly massive interruptions and reversal of the normal global atmospheric circulation

According to Byers, Indian monsoon is the ideal monsoon when differential heating of land and ocean subjected to the annual latitudinal cycle of the sun at its zenith gives rise to immense seasonal wind regimes. However the mechanisms of Indian monsoons are not so simple as they are thought to be. G.B. Cressey has aptly remarked, it has been well said that although every school boy understands the Indian monsoon, the official meteorological department is still in doubt as regards its origin.

In the context of climate change, it is pertinent to ascertain whether the characteristics of Indian summer monsoon also are changing. The Indian summer monsoon (June to September) rainfall is very crucial for the economic development, disaster management, hydrological planning for the country. Long term trends of Indian monsoon rainfall for the country as a whole as well as for smaller regions have been studied by several researchers. The past performances of the monsoon rainfall may give an indication of the future scenario. But in order to do so we should also understand the climatology in a better way. The construction of a homogeneous rainfall data series (spatially as well as temporally) was the first step in this study. India Meteorological Department (IMD) has a good network of rain gauge stations. From the vast data set archived at the National Data Centre, IMD, Pune, a network of 1476 rain-gauge stations was selected which have only 10% or less missing years of data.

Very few studies of extreme temperature analysis have been done in India. Kothawale (2005) studied the temperature extremes in India by using the data of 40 stations well distributed over India for the period 1970–2002, and noted that heat wave conditions are relatively more frequent in May than in June, while very few heat waves occurred in the months of March and April. He also noted that the number of hot days is maximum over central part of India and minimum along the west coast of India during the pre-monsoon season. Recently, Rao et al (2005) studied trends in the frequency of occurrence of extreme temperatures over India during the pre-monsoon and winter seasons, using daily data at 103 stations for the period 1971–2000. They reported that, in Peninsular India during March– May, 80% of stations showed an increasing trend in the days with critical extreme maximum and minimum temperatures. In the northern part of India, 40% stations showed increasing trend in the days with critical extreme maximum temperature while about 80% of the stations showed increasing trend in the extremes in night temperatures.

In India, pre-monsoon is the warmest season of the year, also referred to as the hot-weather season, and anomalously high daily temperatures during this season severely affect human health and comfort. Identification of long-term changes in such conditions and their spatial extents is of critical importance to the development of appropriate risk management strategies. In view of this, the present paper examines the trends in the frequency of occurrence of temperature extremes in maximum and minimum temperatures and day-to-day fluctuations (intra-seasonal variability) of daily temperatures during the pre-monsoon season.

TRENDS IN ALL INDIA MONSOON RAINFALL

All India monsoon seasonal rainfall series as percentage departure from long period average. The series was subjected to a ‘low- pass filter’ in order to suppress the high frequency oscillations. The weights used were nine point Gaussian probability curves (0.01, 0.05, 0.12, 0.20, 0.24, 0.20, 0.12, 0.05, 0.01). It is clearly seen that no linear trend exists in this series. We have also used linear regression technique and the ‘Students t’ test for testing if there is any significance in the trend. All India summer monsoon rainfall as well the rainfall during the four monsoon months do not show any significant trend.

There was a need for development of a homogeneous (spatially and temporally) rainfall series for all the 36 meteorological subdivisions as well as for the country as a whole. The newly constructed rainfall series is uniformly distributed throughout the country and it represents all the existing districts. Though Indian monsoon rainfall as a whole does not show any significant trend, significant rainfall trends are observed over some specific areas. The alternating sequence of multi-decadal periods of thirty years having frequent droughts and flood years are observed in the all India monsoon rainfall data. The decades 1961-70, 1971-80 and 1981-90 were dry periods.

The first decade (1991- 2000) in the next 30 years period already experienced wet period. Therefore, there is a chance of wet period for the subsequent two decades viz. 2001-2010 and

2011-2020. Decadal variability is more for the June and September months while decadal variability of July rainfall is almost similar to that of monsoon rainfall. July rainfall is decreasing for most parts of central India while it is decreasing for the north eastern parts of the country. However June and August rainfall is increasing for the central and south western parts of the country. During the southwest monsoon season, three subdivisions viz. Jharkhand, Chattisgarh, Kerala show significant decreasing trend and eight subdivisions viz. Gangetic WB, West UP, Jammu & Kashmir, Konkan & Goa, Madhya Maharashtra, Rayalseema , Coastal A P and North Interior Karnataka show significant increasing trends. For the first time, we have also studied contribution of each of major rain producing month's (i.e. June, July, August and September) in annual rainfall and examine whether there is any significant change in their contribution.

June rainfall is getting importance as its contribution to annual rainfall is increasing in almost 19 sub-divisions while decreasing in the remaining 17 subdivisions. Contribution of July rainfall is decreasing in central and west peninsular India. But contribution of August rainfall is increasing in all these areas. Significant increasing trend is also observed in the annual rainfall for the sub-divisions Konkan & Goa, Madhya Maharashtra, North Interior Karnataka, Rayalseema, coastal Andhra Pradesh, Gangetic West Bengal, Assam & Meghalaya and Jammu & Kashmir.

VARIABILITY OF RAINFALL

The variability of annual rainfall is lower in high altitudes and eastern parts such as Munsyari (38%) and Mukteshwar (32%) as compared to other stations (Figure 3). The northern most and high altitude station, i.e. Joshimath also shows very low rainfall variability (26%). The Dehradun and Roorkee, located in the south-east part of the study area, are relatively low altitude stations. The observed variability is lowest in Dehradun (21%), while Roorkee has relatively high variability (50%). The analysis reveals that Tehri has highest rainfall variability (86%) among all the stations. Lower correlation coefficient value (−0.18) suggests weak relationships between variability of annual rainfall and elevation. The study further reveals that variability of annual rainfall is overall low as compared to inter-seasonal variability.

The monsoon rainfall has greater impact on the economy of the region due to its significant influence on water supply, agriculture, tourism and socio-cultural life. The study reveals that variability of monsoon rainfall is low in all the stations as compared to other seasons. The variability of monsoon rainfall is not significantly related with elevation as reflected by very low correlation coefficient (−0.07). The variability of monsoon rainfall largely decreases with increasing altitude. Maximum variability was noticed in lowest elevation observatory, i.e. Roorkee (42%). Highest elevation observatory (Mukteshwar) shows comparatively low variability (32%). The winter rainfall shows higher variability as compared to monsoon rainfall in all the stations. The Roorkee has highest winter rainfall variability (140%). There is a general trend that variability decreases towards north with one exception, i.e. Joshimath that records the variability of 47%. The lowest variability has been found in Dehradun, which is just north of Roorkee. The rainfall variability increases from monsoon to post-monsoon season, which gradually decreases to winter and pre-monsoon season. It is further observed that Roorkee has highest variability range among all the seasons followed by Mukteshwar. The high altitudinal areas viz. Joshimath and Munsyari have relatively lower variability range among all the seasons as compared to other observatories.

CONCLUSIONS

During the pre-monsoon season, for India as a whole, the frequency of hot days and nights has increased, however, statistically significant trend is observed for hot nights. The frequency of cold days and nights has decreased over the period 1970–2005, the trend in frequency of cold days is significant at 5% level. On a regional scale, frequency of hot days has significantly increased over the southern parts of India, but there is no significant change over the northern parts of India. Hot nights have increased over all the regions, with significant increase over NW, EC and WC India, whereas cold nights have decreased over all the regions with significant decrease over WH and NE.

The increase in frequency of hot days and nights and decrease in cold days and nights are consistent with the increasing trend in monthly/seasonal maximum and minimum temperatures, respectively. In general, trends in the frequency of the occurrence of extreme events as well as trends in 90^{th} and 10th percentiles of both daily maximum and minimum temperatures indicate that the increase in the frequency of hot weather events during the pre-monsoon season has also been associated with an increase in their intensity. Examination of changes in the day-to-day fluctuations in the temperatures indicates that the daily maximum and minimum temperatures are becoming less variable during the recent decades.

UNIT-9: HISTORY OF GLOBAL CLIMATE CHANGE

INTRODUCTION

Climate changes, with or without human influence, at every different timescale and over many orders of magnitude. The climate system sometimes appears overwhelmingly complex, with many possible interacting causes of change, thousands of feedbacks, and a maddeningly short instrumental record. How can we even hope to understand the system, let alone use that understanding to predict the future climate? And yet, we have, since the late 1960s, been able relatively successfully to simulate the climate in computer-based mathematical models and reconstruct the climate going back centuries, millennia, and even millions of years by using a wealth of proxy data. We have been able to identify both the necessary natural role and the potentially disastrous anthropogenic emissions of greenhouse gases, we have for the first time been able to predict and monitor an entire El Niño, and we can now glean annually from ice cores resolved information on the climate from the last ~110 000 years. All of these achievements were unimagined in the mid twentieth century and we become ever more aware that knowledge of the climate system is necessary for understanding our vulnerability and opportunities in the future. We only have to look at human history to be conscious of the important role that climate has played in our development and economy from exploration of the globe to the onset of agrarianism to the tourism boom.

Now we are faced with the daunting prospect of unwittingly embarking on a dangerous experiment with the climate. There are no historical analogues for the rate at which we are transferring carbon from its terrestrial reservoir to the atmosphere. It is going to need all our knowledge gained to this point and a great deal more to comprehend fully the role we are playing in altering the climate system. While we, in our heated, airconditioned homes and offices, feel untouchable by a change in the climate, the reality is that nearly all of our day-to-day existence depends on its relative stability. The harvesting of our food, our health, the availability of water, mobility, and subtler aspects of our economy, such as tourism, urban design, and insurance all depend on weather and climate. It is therefore imperative that we build on our achievements thus far in a continued effort to gain deeper insight into the complex atmosphere and its behavior in the future.

HISTORY OF GLOBAL CLIMATE CHANGE SCIENCE

A direct human influence on the environment is not new. Clear cutting for agriculture and changes in faunal composition due to hunting and domestication of animals are known to have occurred thousands of years ago. However, until recently, our influence on the environment has been relatively localized. With the industrial revolution and population increase, many environmental issues have become more widespread in nature with recent and future climate change or global warming, being an example of a truly global effect of human activity. Only in the second half of the twentieth century did the potential human impact on the climate system emerge, and the possible dimensions of the problem were not fully appreciated until the 1970s. Significant advances in the understanding of the different components of the complex climate system (the atmosphere, hydrosphere, biosphere, cryosphere, and to a certain degree the lithosphere) have been made since then.

These give us the tools, not only to investigate processes within subsystems and their complex interactions, but also to realize that their relative roles in changing the climate vary. Once the rapid increase of greenhouse gases was recognized, and data of atmospheric composition of previous centuries and millennia (including the ice ages) became available from ice cores, long known calculations of effects of increasing greenhouse gases were rediscovered. The "largescale geophysical experiment" was now recognized as potentially dangerous. Today, it is clear that the changes in atmospheric composition are human made. But to what degree have these changes already influenced the climate, and what will their effect be in the future? The challenge is to separate the natural variability from the humaninduced changes within the complex climate system. If the predictions are right, this problem become easier to resolve as time moves on. But can we afford to wait?

THE EARLY TIMES

The earliest interest in "climate" was of a rather pragmatic nature. The change of the seasons and with it changes in the food supply required adaptations of early hunters and gatherers, even more so after the agricultural revolution, when it became critical to farmers to know when to plant and when to harvest. Therefore, knowledge about the seasons and the annual cycle were and still are crucial in all of the cultural centers, from tribes to large civilizations. It is perhaps not surprising that very precise calendars were compiled in most of the early cultures. Hammurabi (reigned ~1792–1750 BCE) unified the different calendars of Babylon, the Romans improved the calendar to the Julian calendar, and the Maya reached the astonishingly precise determination of the year length with 365.242 days.

FROM THE MIDDLE AGES TO THE TWENTIETH CENTURY

Only after the development of instruments to measure meteorological parameters did the

Aristotelian view of the atmosphere slowly disappear. After receiving information of a first “thermoskop” built in the Netherlands (probably based on old descriptions of Philon of Byzantine (250 BCE)), Galileo Galilei (1564–1642) quickly built his own. This was around 1640, when the Florentine Academia del Cimento established the first network of stations through Tuscany, using the same type of thermometers at each location. The design of the first thermometers suffered from changes in air pressure, which led to the development of barometers by Berti (1600–1643) and Torricelli (1608–1647). The famous experiment initiated by Pascal and performed by Perrier on September 19, 1648, in which a barometer was carried onto Puy de Dôme and back down, finally confirmed the hypothesis that the air consists of a compressible gas with a measurable weight, began a new view of the atmosphere. Since the beginning of the balloon flight era in 1783, vertical cross sections of atmospheric properties without influence of the surface have been able to be measured. One of the most famous ascents took place in 1862, when Glaisher and Coxwell reached a record altitude of 8839 m.

In the mid nineteenth century, the use of the telegraph to transmit observations of climate and weather parameters rapidly showed the potential of weather forecasting. This led to the foundation of national weather services, and in 1873 to the formation of the International Meteorological Organization. Today, this organization, a part of the United Nations, is called the World Meteorological Organization (WMO) and is located in Geneva, Switzerland. Among many other efforts, the WMO is responsible for World Weather Watch, a program to monitor and analyze climate on a global basis.

ACCUMULATING EVIDENCE

The second period, from about 1945 to 1975, was a time of testing some of these hypotheses through a successful search for manifold new evidence. Governments granted money for such studies beginning with the International Geophysical Year of 1957-58. Roger Revelle and Hans Suess, attempting to establish the role of oceans in absorbing carbon dioxide, inspired Charles Keeling to establish remote stations on Mauna Loa in Hawai'i and in Antarctica to monitor the concentration of that gas in the atmosphere, resulting in the discovery of the "Keeling Curve," which showed that it is steadily increasing. Revelle and Suess made an ominous conclusion: "Human beings are now carrying out a large scale geophysical experiment of a kind that could not have happened in the past nor be reproduced in the future.

Satellites began to measure temperatures on the surface and in the atmosphere, and to provide detailed information on changes in solar irradiance, stratospheric chemistry including ozone depletion such as the Ozone Hole over Antarctica, temperature, precipitation, clouds and water vapor, wind velocity, ocean currents and surface temperatures including El Niño, sea level, vegetation and desertification, coastal configuration, volcanoes, snow cover, sea ice, glaciers and ice sheets, and human activities such as fires and the growth of cities.

Climate scientists braving difficult weather drilled ice cores in Greenland and Antarctica, revealing evidence of past concentrations of atmospheric gases, temperatures, dust, pollutants, and radioisotopes. 9 For example, it became possible to plot long-term temperature variations and changes in carbon dioxide concentration on the same time scale, and to note that they follow patterns that are remarkably similar. In a chart of these two curves, a relatively regular, approximately 100,000-year cycle is prominent, although it also appears that the climate has occasionally changed rapidly in very short time periods. The most recent decades have shown a divergence in which the level of carbon dioxide has risen more rapidly than temperature.

ATMOSPHERIC COMPOSITION AND THE GREENHOUSE EFFECT

Within about a century of the Puy de Dôme experiment and the conclusions by Descartes that the atmosphere was not uniform but compressible, the basic composition of the atmosphere was established. In 1657, Boyle and Hook found the "life substance," later called oxygen by Lavoisier. Black discovered carbon dioxide (CO_2) (a gas that killed animals but let plants grow) in 1753, and Rutherford found what was later called nitrogen in 1770. It was not until 1894 that Rayleigh and Ramsey found argon by using the technique of spectroscopy. This technique was based on Herschels detection around 1800 that light is a source of energy that is not restricted to the visible range, but also extends through the red spectrum into the infra-red. Maxwell then found that this energy is transmitted by electromagnetic waves, and that the "visible" is only a very small portion of the full spectrum including ultra-violet.

With this background, Fourrier (1768–1830) published a fundamental work about the temperature of the globe in 1824. He linked the temperature of the atmosphere to the atmospheric gases and their conservation of radiation, suggesting that the system acted like a large bell jar. After geologists, led by Agassiz, encountered evidence of ice ages in the geologic past, Tyndall postulated in 1861 that a reduction in CO_2 could lead to a new ice age. This was picked up by Arrhenius in 1896, when he computed the dependence of temperature on atmospheric composition in terms of CO_2. Although he was strongly underestimating the time period needed for humans to increase atmospheric CO_2 (he thought it would take about 3000 years to reach a doubling or tripling), his calculations of the temperature response are very similar to estimates of temperature response in present calculations. He referred to the climate for these high CO_2 concentrations as "hothouse" conditions, which Trewartha later (in 1937) termed the "greenhouse effect." While Arrhenius expected the increased temperatures during these hothouse conditions to be benevolent for humanity, Revelle and Suess realized that the changes were much more rapid and that observing responses to the changing atmospheric composition could offer insight into how the weather and climate works, in this way approximating a big "geophysical experiment." Callander and then especially the precise CO_2 measurements of Keeling in 1960 showed the rapid increase of CO_2 over the twentieth century (United States, Office of Science and Technology Policy, 1997), a time series recently expanded in much detail by ice cores (Indermuehle et al. 1999). Today we also include methane, nitrous oxide, chlorofluorocarbons (CFCs), and ozone (tropospheric) as additional greenhouse gases.

STATUS OF GLOBAL CLIMATE CHANGE

Any anthropogenic effects on future climate will be superimposed on a natural variability that may serve to mitigate or amplify these effects. The only method of assessing the role of natural versus anthropogenic forcing is through an understanding of past climate variations. To achieve this, annual or preferably seasonal information needs to be gained for as wide a geographic area as possible. However, instrumental records of weather and climate exist only for the last 100–150 years and prior to this it is necessary to rely on "proxy" information of climate. Great improvements have been made in recent decades to assemble a wealth of annually and seasonally resolved proxy climate data for the last 1000–2000 years and beyond so that we can address our understanding of climate processes with this perspective.

UNIT-10: CLIMATE CHANGE: CAUSES AND EVIDENCES

INTRODUCTION

Climate is an atmospheric condition of an area over periods ranging from at least three decades to thousand years. In other words, climate is average weather over a long time. Also, climate varies in different parts of the earth mostly because patterns of global air circulation and ocean currents distribute heat and precipitation unevenly from the tropics to other parts of the world.

Numerous research, including Goddard Institute for Space Studies, Britain‘s Hadley Centre for Climate Change, the Japan Meteorological Agency, NASA and NOAA‘s National Climatic Data Center have used raw measurement data to provide records of long-term global surface temperature change. These researches reveal that Earth‘s average surface temperature has increased by more than 1.4°F (0.8°C) over the past 100 years, most of this increase occurred over the past thirty five (35) years. A temperature change of 1.4°F may not be consider significant when dealing about daily or seasonal fluctuation but it must be consider significant when dealing with permanent increase globally.

In the same vein, scientists‘ research showed that for at least 2,000 years before the Industrial Revolution, atmospheric CO_2 concentrations were steady and then began to rise sharply during the late 1800s. Today, atmospheric CO_2 concentrations exceed 390 parts per million—nearly 40% higher than the preindustrial period (National Research Council of the National Academics, 2020). The study of some robust evidence of climate change in the twentieth-century in India was carried out by. Maximum and minimum temperature for the period of 1901-2003, rainfall for the period of 1871-2002 and sea surface temperature were analyzed in the study and increased trends of these parameters were noticed.

Greenhouse Gases such as carbon dioxide (CO_2) absorb heat (infrared radiation) emitted from Earth’s surface. Increases in the atmospheric concentrations of these gases cause Earth to warm by trapping more of this heat. Human activities—especially the burning of fossil fuels since the start of the Industrial Revolution—have increased atmospheric CO_2 concentrations by about 40%, with more than half the increase occurring since 1970. Since 1900, the global average surface temperature has increased by about 0.8 °C (1.4 °F). This has been accompanied by warming of the ocean, a rise in sea level, a strong decline in Arctic sea ice, and many other associated climate effects. Much of this warming has occurred in the last four decades. Detailed analyses have shown that the warming during this period is mainly a result of the increased concentrations of CO_2 and other greenhouse gases. Continued emissions of these gases will cause further climate change, including substantial increases in global average surface temperature and important changes in regional climate. The magnitude and timing of these changes will depend on many factors, and slowdowns and accelerations in warming lasting a decade or more will continue to occur. However, long-term climate change over many decades will depend mainly on the total amount of CO_2 and other greenhouse gases emitted as a result of human activities.

CLIMATE IS ALWAYS CHANGING

The largest global-scale climate variations in Earth's recent geological past are the ice age cycles, which are cold glacial periods followed by shorter warm periods. The last few of these natural cycles have recurred roughly every 100,000 years. They are mainly paced by slow changes in Earth's orbit which alter the way the Sun's energy is distributed with latitude and by season on Earth. These changes alone are not sufficient to cause the observed magnitude of change in temperature, or to act on the whole Earth. Instead they lead to changes in the extent of ice sheets and in the abundance of CO_2 and other greenhouse gases which amplify the initial temperature change and complete the global transition from warm to cold or vice versa.

Recent estimates of the increase in global average temperature since the end of the last ice age are 4 to 5°C (7 to 9 °F). That change occurred over a period of about 7,000 years, starting 18,000 years ago. CO_2 has risen by 40% in just the past 200 years, contributing to human alteration of the planet's energy budget that has so far warmed Earth by about 0.8 °C (1.4 °F). If the rise in CO_2 continues unchecked, warming of the same magnitude as the increase out of the ice age can be expected by the end of this century or soon after. This speed of warming is more than ten times that at the end of an ice age, the fastest known natural sustained change on a global scale.

CLIMATE CHANGE NOW

There is clear evidence to show that climate change is happening. Measurements show that the average temperature at the Earth's surface has risen by about 1°C since the pre-industrial period. 17 of the 18 warmest years on record have occurred in the 21st century1 and each of the last 3 decades has been hotter than the previous one. This change in temperature hasn't been the same everywhere; the increase has been greater over land than over the oceans and has been particularly fast in the Arctic. The UK is already affected by rising temperatures. The most recent decade (2008- 2017) has been on average 0.8 °C warmer than the 1961-1990 average. All ten of the warmest years in the UK have occurred since 1990 with the nine warmest occurring since 2002. Although it is clear that the climate is warming in the long-term, note that temperatures aren't expected to rise every single year. Natural fluctuations will still cause unusually cold years and seasons, but these events will become less likely.

Along with warming at the Earth's surface, many other changes in the climate are occurring:

• warming oceans

• melting polar ice and glaciers

• rising sea levels

• more extreme weather events

WARMING OCEANS

While the temperature rise at the Earth's surface may get the most headlines, the temperature of the oceans has been increasing too. This warming has been measured all the way down to 2 km deep.

The chemistry of the oceans is also changing as they absorb approximately a third of the excess carbon dioxide being emitted into the atmosphere. This is causing the oceans to become acidic more rapidly than perhaps any point in the last 300 million years.

MELTING POLAR ICE AND GLACIERS

As the Arctic warms, sea ice is decreasing rapidly. In the Antarctic, sea ice has slowly increased, driven by local changes in wind patterns and freshening sea water. However, in recent years Antarctic sea ice has stopped growing. Over the past few decades the ice sheets (the great masses of land ice at the poles) in Greenland and the Antarctic have shrunk, as have most glaciers around the world.

RISING SEA LEVELS

As land ice melts and the warming oceans expand, sea levels have risen. Global sea level has risen by around 20cm over the past century, likely faster than at any point in the last 2,000 years. The rate of sea level rise has increased substantially over the 20th Century and further rise this century is inevitable – how much depends on the amount of human greenhouse gas emissions.

MORE EXTREME WEATHER EVENTS

More damaging extreme weather events are being seen around the world. Heat waves have become more frequent and are lasting longer. The height of extreme sea levels caused by storms has increased. Warming is expected to cause more intense, heavy rainfall events. In North

America and Europe, where long-term rainfall measurements exist, this change has already been observed.

EVIDENCE AND ANALYSIS

EVIDENCE FROM PAST CLIMATE CHANGE

Ancient ice from the polar ice sheets reveal natural temperature changes over tens to hundreds of thousands of years. Air bubbles trapped in the ice show those levels of greenhouse gases in the atmosphere are closely linked to global temperatures. Rises in the temperature match closely with an increase in the amount of greenhouse gases.

These ice cores also show that, over the last 350 years, greenhouse gases have rapidly increased to levels not seen for at least 800,000 years and very probably longer. Modern humans, who evolved about 200,000 years ago, have never previously experienced such high levels of greenhouse gases.

NATURAL FLUCTUATIONS IN CLIMATE

Over the last million years or so the Earth's climate has had a natural cycle of cold glacial and warm interglacial periods. This cycle is mainly driven by gradual changes in the Earth's orbit over many thousands of years, but is amplified by changes in greenhouse gases and other influences. Climate change is always happening naturally, but greenhouse gases produced by human activity are altering this cycle. Volcanic eruptions and changes in solar activity also affect our climate, but they alone can't explain the changes in temperature seen over the last century.

Scientists have used sophisticated computer models to calculate how much human activity – as opposed to natural factors – is responsible for climate change. These models show a clear human 'fingerprint' on recent global warming. The latest Assessment Report from the IPCC said it was extremely likely that most of the observed increase in global temperature since the 1950s is due to human activity.

CLIMATE MODELS AND FUTURE GLOBAL WARMING

We can understand a lot about the possible future effects of a warming climate by looking at changes that have already happened on Earth. But we can get much more insight by using mathematical models of the climate. Climate models can range from a very simple set of mathematical equations (which could be solved using pen and paper) to the very complex, sophisticated models run on supercomputers (such as those at the Met Office). While these models cannot provide very specific forecasts of what the weather will be like on a Tuesday in

100 years time, they can forecast the big changes in global climate which we could expect to see in the future.

All these climate models tell us that under a scenario of ever-increasing greenhouse gas emissions the world could become up to 4.8°C warmer than the pre-industrial period by the end of this century . Note these are global averages and that temperatures in certain regions, such as the Arctic, would be even higher than this.

At the Paris climate conference (COP21) in December 2015, 195 countries adopted the first-ever universal global climate deal that is due to come into force in 2020. The agreement sets out a global action plan to put the world on track to avoid dangerous climate change by limiting global warming to well below 2°C above pre-industrial levels and pursue efforts towards limiting to 1.5°C. The country commitments we have seen so far represent a dramatic improvement on ‘business as usual’ emissions projections. But these commitments are predicted to give rise to global temperature increases of around 3°C. Further urgent action is needed therefore to put us on track to well below 2°C.

UNIT-11: GLOBAL WARMING: CONSEQUENCES AND ADAPTATIONS

Global warming has now become a matter of most serious concern for the humanity. Global warming is defined as the increase in the average temperature of earth's surface, atmosphere and oceans. The United Nations Framework convention on Climate Change defines global worming as "a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability observed over comparable time periods". Global warming is largely caused by the emission of carbon dioxide. According to Inter-governmental Panel on Climate Change (IPCC) climate change refers to 'any change of climate over time weather due to natural cause or as a result of human activity'. In 2013, the Inter-governmental Panel on Climate Change (IPCC) Fifth Assessment Report state that human influence is the dominant cause of warming since the mid 20^{th} century.

The major causes are burning of fossil fuels in industry and automobile, emission of methane from animal and agriculture, use of chemical fertilizers, volcanic eruptions, deforestation, ozone layer depletion etc. have been increasing the concentration of green house gases in atmosphere and increasing effects of green house effects. At present Global Warming potential (GWP) is used to measure the warming level of the globe due to increasing level of green house effects.

Major sources of green house gases

Name of the Gases	Major sources
Carbon Dioxide (CO2)	Burning of fossil fuel, deforestation, industrialization, auto mobiles, volcanoes, hot spring etc.
Methane(CH4)	Wet lands, volcanoes, paddy field and other organic substances etc.
Nitrous Oxide (N2O)	Blazing of fossil fuel, manufacturing centre, and agricultural fertilizer.
Dichlorodifluoromethane ChlorodifluoromethaneTetrafluoromethaneHexafluoroethane	Most of electronic equipments like refrigerator, air conditioner and various industrial sources

Water Vapour (H2O)	Water vapour directly not produced from human activity

Carbon dioxide is the most significant green house gas and minor gases like halogenated gases (chlorofluorocarbons), halons, methane, nitrous oxide and ozone etc. also have increase the greenhouse mechanism of atmosphere and increase the temperature of earth and atmosphere gradually. These gases are released from different sources both from human activities and natural sources. Some of the major sources of the green house gases are listed in the chart-

The increasing trend of temperature may cause reduction of ice covers in Arctic region, glacial retreatment, rise in sea level, breaching of coral reefs, crop failures, increases the frequency of severe storms like hurricanes, cyclones etc. and some of these effects are evident in the present earth.

According to Inter-governmental Panel on Climate Change (IPCC) climate change refers to 'any change of climate over time weather due to natural cause or as a result of human activity'. From past to present the evidences which are used to indicate climate change phenomena are-biological indicators: Floral and faunal remain in fossils, geological indicators: sedimentary deposits, cryogenic indicators: change of sea level, geomorphologic indicators, migration of man and animals due to drought and flood etc. Earth's elliptical orbit, variation of solar radiant, increasing green house effects, combustion of fossil fuels, industrialization, urbanization, deforestation, changes of consumer behavior etc, facilitate the change of climate. Global warming, heat island, ice melting, acid rain, desertification, changes in moisture budget, loss of biodiversity, increasing frequency of natural disasters etc. are some major effects of climate change. Efficient and wise use of resources and energy, proper landuse, green technology, afforestation, management, laws, policy making regarding resource use and conservation and human awareness and consciousness are necessary to combat, minimize, mitigate and prevent theharmful effects of climate change.

CONSEQUENCES

- The effects of global warming have taken its role on people, animals, birds and habitat. In fact no continent has been spared.

- At Antarctica, shrinking sea ice has reduced the population of the Adelie penguin by 33% in 25 years.
- In Canadian Arctic, the Peary caribou population has decreased due to heavy rainfall from 24,000 in 1961 to 1100 in 1997.
- Developing countries are twice as at risk to climate change as industrialized countries, and small islands states are thrice as at risk, according to a group of UN scientists. U.S is the fourth country in the world responsible for CO2 emission.
- Estimates drawn from reports by the Intergovernmental Panel on Climate Changes (IPCC) projects increase in average global temperatures ranging from 1.4 0C to 5.8 0C by the year 2100.
- A slight rise in temperature even by 1 degree C, can have adverse effect on the Sea levels. This sea level rise would threaten coastal cities (Calcutta, Mumbai, Chennai, etc.) and some 60-odd island nations such as Maldives, Bangladesh, etc.
- Global warming increases the desert. It increases temperature in North America, South Africa, Mexico, India and other countries. Changes of hurricanes, cyclones and floods will be more which will damage the lagoons, estuaries and coral reefs. Global warming may cause extinction of more than one million species of animals and plants by 2050 AD.
- Change in climate can also bring several other problems, such as Drying of surface water bodies, fall in ground water table, serious water shortage, desertification of vast areas which were hitherto fertile and productive lands, crop pattern change and reduced agricultural yields shortage of food, growth in micro organisms and spread of diseases, etc.
- In the United States, Chicago experienced one of the worst weather-related disasters in Illinois history when a heat wave resulted in 525 deaths during a 5-day period in July of 1995 (Kunkel *et al.*, 1996).
- The two American data sets have 2005 as their warmest year. While 2006 and 2007 were cooler than 2005 in all three data sets, two such cooler years is much too short a time to conclude that the clear warming trend over the second half of the 20th Century has stopped or reversed. Figure 2 shows many sets of three consecutive years with a short-lived cooling trend that is reversed soon afterwards.
- The British data set has 1998 as its warmest year. Is the ten years from 1998 to 2007 long enough to establish a cooling trend? We have noted that ten years is about the minimum

averaging time to remove the year-to year variations in these global temperature data sets, so ten years might be just enough to reveal any downturn in the underlying trend. However, there hasn't actually been a cooling over the decade 1998-2007. In all three data sets, the linear trend over 1998-2007 is upward

- (i.e., one of warming), even if the warming is weaker in the British data set than in the American data sets (Fawcett Robert and Jones David, 2008).

ADAPTATION: PAST, PRESENT AND FUTURE?

Given the potential risks associated with climate change, a serious effort on characterizing and understanding adaptation is therefore now underway. Analogues of adaptation in the past are complemented with policy and social science research on the present adaptive capacity of governments, civil society and markets to deal with climate perturbations. The economic costs of future adaptations are being derived by examining the differences between the economic losses associated with scenarios of technology uptake and diffusion. Among these approaches, a key issue is the identification of successful adaptations in the developing world where the greatest risk and physical vulnerability persists. But within examples of success, from indigenous strategies for resource management, to large-scale infrastructure and irrigation, there will still be winners and losers.

First of all it is necessary to distinguish adaptation by who is undertaking it and the interests of the diverse stakeholders involved. It is clear that individuals and societies will adapt and have been adapting to climate change over the course of human history – climate is part of the wider environmental landscape of human habitation. Thus individuals and societies are vulnerable to climate risks and other factors and this vulnerability can act as a driver for adaptive resource management. There are various geographic scales and social agents involved in adaptation. Some adaptation by individuals is undertaken in response to climate threats, often triggered by individual extreme events. Other adaptation is undertaken by governments on behalf of society, sometimes in anticipation of change, but, again, often in response to individual events.

But these levels of decision-making are not independent – they are embedded in social processes that reflect the relationship between individuals, their networks, capabilities and social capital, and the state. Sometimes a distinction is drawn between planned adaptation, assumed to be undertaken by governments on behalf of society, and autonomous adaptation by individuals. But this distinction obfuscates the role of the state in providing security, or in using security as a weapon of coercion when faced with an environmental risk. The nature of the relationship between individuals and agents of government in handling risk is a fraught but under-researched area. Political ecology approaches demonstrate that, for example, when faced with a flood risk, residents of marginalized but risky areas of Georgetown, Guyana, have only a limited set of adaptation options – and the state allows such risks to exist as part of the politized nature of urban planning and control.

So the role of collective action in facilitating adaptation is a key issue where lessons can be learned from political ecology, and other theoretical insights, for present day adaptation processes. From research on collective action it is clear that the size of the group undertaking the collective action, the boundaries of the resource at risk, the homogeneity of the decision-making group, the distribution of benefits of management and other factors are all important in determining the ultimate success of collective management. Research is required on how collective action is central to adaptive capacity at various scales of decision-making. At present, insights about responses to climate change as collective action are primarily used to examine national level cooperative action to reduce greenhouse gas emissions under the UN Framework Convention on Climate Change, rather than on how the process of adaptation evolves.

Analogues of past climate change contrast with scenarios derived from climate model experiments in the search for adaptation insights. The analogue approach involves taking detailed case studies of past responses to climate variability and extremes (temporal analogues) or present-day behaviour in regions with climate conditions similar to those that might possibly develop in the region of interest (spatial analogues). The aim is to establish how individuals and institutions anticipate or respond to reduce the risks of different types of climate variability and how policy has influenced these actions. Understanding the present-day effects and response to climate variability at all levels of social organization is a prerequisite for studying the effects and responses to future climate change and for identifying the key determinants of successful adaptation in the future.

High levels of inter annual rainfall variability and their effects on water resources in Africa can provide illustrative examples of climate–environment–society interactions. A commonly cited drawback to the analogue approach to climate change assessment is that the characteristics of future climate change are likely to be very different to past climate variability, particularly in terms of the rate and magnitude of change. Examples exist for Africa, however, where the observed rainfall variability is *greater* than changes suggested by climate models for the next 50–100 years.

CONCLUSION

Besides all the scientific data available with us showing the trend of global cooling, we must understand that the nature is unpredictable in its behavior and can come before us in a random manner violating the depicted trend, so the human being should be ever ready to tackle this problem or to face the consequences .Thus a new research trend should be initiated and promoted which can absorb the surrounding heat in a large amount to cause cooling. Further, the message of this study is a simple one: Global warming is a serious problem that will not solve itself. Countries should take cooperative steps to slow global warming. There is no case for delay. The most fruitful and effective approach is for countries to put a harmonized price, perhaps a steep price, on greenhouse gas emissions, primarily those of carbon dioxide resulting from the combustion of fossil fuels. While other measures might usefully support this policy, placing a near universal and harmonized price or tax on carbon is a necessary, and perhaps a sufficient condition, for reducing the future threat of global warming. But it does draw our attention to the fact that nature takes its own course. Somewhere in our arrogance and in our desire to play god, we might have forgotten that there is a universal force governing all life, including us.

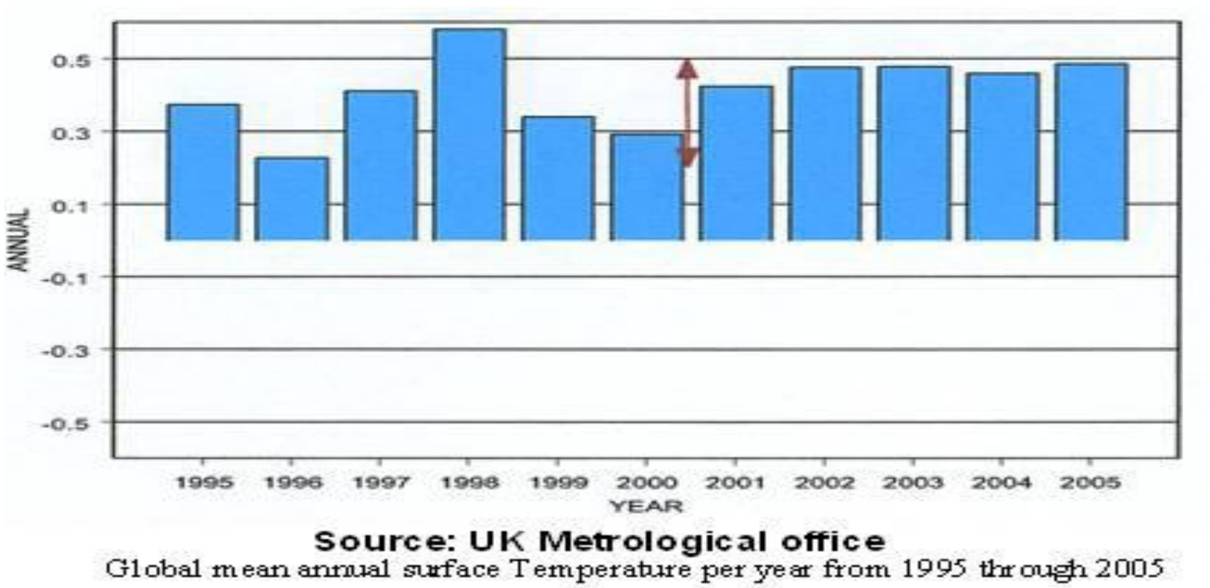


Fig: - Variations in the Temperature from 1995 to 2005

UNIT 12: - WEATHER FORECASTING: SHORT, MEDIUM AND LONG RANGE

DEFINITION OF WEATHER FORECASTING

Weather forecasting is very much important for aviation, navigation, economic activities specially agriculture even for socio-cultural activities and daily work. Simply, it refers to prediction of weather conditions i.e. temperature, humidity, cloudiness, wind circulation, precipitation and other atmospheric condition of different spatial scale (Local, regional etc.) in different temporal scale (hourly, daily, weekly, monthly, yearly). So, the atmospheric state in a given time and the physical laws that govern the changes of the state are two key concern of weather forecasting. However, three sequential procedures i.e. observation of weather conditions, collection, compilation and analysis of data and forecasting and different methods like synoptic, numerical and statistical as well as entomogenic methods are used to predict the weather.

TYPES OF WEATHER FORECASTING

On the basis of temporal range or periodicity forecasting may be short range, medium range and long range-

Weather forecasting	Duration (Validity)	Level of accuracy
Short range	Part or whole of a 24 hours period with a further outlook for the following 24 hours.	Satisfactory
Medium range	Period of two to five days	Helpful for general public specially farmers but accuracy level is relatively less than short range.
Long range	Period longer than five days, vary from a fortnight to a month, season or year.	Less than former two

On the basis of the purpose of the society, weather forecasting can be grouped as aviation forecasting, shipping forecasting, agricultural forecasting, local forecasting etc.

Aviation forecasting refers to the prediction of weather for different aspects of aviation like take off, climbing, flying, descent and landing of the aircrafts. Temperature, air pressure, winds, visibility, clouds, atmospheric storm, jet stream etc. are included in this prediction.

Shipping forecasting is for navigation and it includes both the atmospheric conditions as well as temperature of sea water, nature and ferocity of sea waves and issue of warning about sea condition.

Agricultural forecasting is mainly for farmers and includes information and warning about sky conditions, precipitation, frosts, fogs, dry spells etc. which are very useful for sowing, growing and harvesting of different crops.

Local forecasting is area-specific daily forecast of weather condition for general public and administrators.

METHODS OF WEATHER FORECASTING

Weather and climate have intense influence on physical, economic, social, cultural life and that's why weather forecasting is very much necessary in various purposes in our life. The forecaster predict future tendency of weather considering atmospheric status in a given time and the physical process that govern the state and its changes. Though weather forecasting is based on the theory of probability, but, for a more accurate and obvious forecasting different methods are adopted for different types of forecasting such as – Synoptic methods numerical methods, statistical methods, biological methods, folklore methods etc. The first here methods are mainly used by professional weather forecasters who observe and collect weather data through different tools and techniques and analyse to predict weather condition.

i. Synoptic methods are one of the most important weather forecasting methods. It means the observation of different weather elements in a specific time in a particular spatial unit i.e. the overall generalization of weather condition of an area a given time. Atkinson (1968) defined that 'synoptic forecasting entails the diagrammatical representation of weather system through time and the extrapolation of developments of such system into future'. These methods related to the preparation of synoptic chart of the present weather condition and on the basis of these chart future projections of the weather elements are prepared. It has two aspects viz. prognostic or prebartic chart which predict the surface weather conditions and it is for upper atmosphere is known as prontour chart.

Synoptic models, sometimes, are used to predict weather mainly cyclone and anticyclone. It is anideal synoptic situation and it is used keeping the facts that weather does not change in a complete random fashion rather follow a fairly ordered system. As it is a simplified ordered situation of weather, it may deviate from the actual.

Analogue method is another technique under synoptic methods. It depicts the future based on past occurrences. Though infinite variety of analogue pattern is possible for atmospheric circulation, based on the surface and upper surface atmospheric data and interactive situation future inference can be prostulated.

Synoptic process extrapolates the past and current occurrences but any future changes do not possible to include in the study. Besides, skill and experience of the forecaster influence the forecasting. However, these methods are widely used for short range weather forecasting and satellite imagery and computer techniques enhance the utility of the methods.

ii. Numerical methods are the application of mathematical equations, laws, models etc. in the prediction process of weather. It is based on the belief that atmosphere is a fluid and flow the physical laws and accordingly subsequent state develop from preceding ones. Historically, it stress back its origin into 1920s after V. Bjerkness and L.F. Richardson but got momentum after 1950s with the introduction of computer simulations that reduce the computation time to less than the forecast period itself.

In this method, manipulation and articulation of voluminous database and huge mathematical calculations are necessary and it requires most powerful supercomputer. Over the area of interest, grid are developed to feed the weather data and future condition of atmosphere are predicted using supercomputer and following physical laws that govern the motion of atmosphere and prepare the prognostic chart. This methods widely used in long range weather forecasting such as the preparation of simulated model of climate change and it also used in medium and short range forecasting like air quality modeling, ocean surface modeling, tropical cyclone forecasting etc.

Collection of prerequisite numerous data and billion of arithmetic operation for a model is very hectic process and changing of atmospheric conditions could not be included in ongoing calculations and so, when forecasting report comes to reality, large scale change within atmosphere may be occurred. Besides, physical assumptions and simplification of reality into model deviate from the reality.

iii. Statistical methods often supplement the numerical method and uses the past records of weather data based on the conception that future will be the repetition of past. In this method, data relating to temperature, air pressure, wind velocity, cloudiness, rainfall etc. are compiled and statistical analysis such correlation, regression, factor analysis etc. are used to depict the relationship, trends of the present weather condition. But, it is the drawback that it provides the average pictures of weather element.Though completely accurate forecasting is impossible and it is mainly follow the law of probability but now a day, due to the advancement of tools, techniques and database and ensemble methods, weather forecasting has become more accurate and reliable.

SUMMARY

- Climatology is the study of climatic phenomena over different spatial and temporal scales.
- Cooling or heating of air through the expansion or contraction of air is called adiabatic process.
- Retarded upward movement or tendency of downward sinking of air parcel is called stability while continuous vertical movement of air parcel is known as instability.
- Vertically and horizontally extensive body of air with homogeneous properties (temperature, moisture, lapse rate etc.) is called air mass and two contrasting air masses form front.
- Surface and upper surface flows form convective cells and three cells circulation exist in both hemesphere.
- Origin of monsoon can be explained in the light of thermal factor, air mass modification, jet stream etc.
- In southern Pacific Ocean, the normal circulation is walker circulation, warmer than average is El Nino and Colder than average is La Nina Situation.
- Weather forecasting is the future prediction of weather phenomena with the help of different surface and upper surface atmospheric data data and different tools and techniques.
- Gradual rise of earth and atmospheric temperature is called as global warming and alteration of climatic element due to natural or human causes is known as climate change.

SELF ASSESSMENT TESTS

Short type questions

- Define climatology.
- What is adiabatic lapse rate?
- What is short, medium and long range weather forecasting?
- Differentiate the stability from instability.

Descriptive type questions

- Discuss thescope of climatology
- Account for the classification of air masses on the basis of source region.
- Explain the theories relating to the origin of monsoon.
- Elucidate the atmospheric circulation with reference to tricellur model.

STUDY TIPS AND REFERENCES

Atkinson, B.W.(1968). *The Weather Business*, London: Aldus Books.

Berry, R.G. and Chorley, R.T.(1998). *Atmosphere, Weather and Climate* (7^{th} ed.), London: Routledge.

Blair, T. A. andFite, R. C. (1965).*Weather Elements: A Text in Elementary Meteorology*, New York: Prentice Hall.

Bryson, R.A. (1997). The paradigm of climatology: an essay,*Bulletin of the American Meteorological Society*, 78(3): 449–455.

Chritchfield, H. J. (1983). *General Climatology*, New Delhi: Prentice Hall

Das, P. K. (1995). *Monsoon*, (2nd ed.), New Delhi: National Book Trust

Lal, D.S. (1993). *Climatology*, New Delhi: Chaitanya Publishing House

Lamb, H. H.(1972). Climate: Past, Present, and Future,*Fundamentals and Climate Now*, Methuen and Co., 1: 613

Mathur, J. R. (1974).*Climatology: Fundamental & Application*, New York: Mc Grow Hill,

Morgan, J.M. and Morgan, M. D. (1997).*Meteorology: The Atmosphere and the Science of Weather* (5th ed.), Upper Saddle River, New Jersy: Prentice Hall.

Oliver, J.E. and Hidore, J.J. (1984).*Climatology*, Columbus:Merrill Publishing Company.

IPCC website ,www.ipcc.ch accessed on 18th April, 2017, 5:30 PM IST.

Siddhartha, K. (2001). *Atmosphere, Weather and Climate: A Textbbook on Climatology* (4th ed.), New Delhi: Kisalaya Publications.

Singh, S.(2005). *Climatology*, Allahabad: PrayagPustakBhawan.

Strahler, A.H. and. Strahler, A.N.(1976). *Elements of Physical Geography*, New York: John Wiley and Sons

Thornthwaite, C. W.(1948). An Approach towards a Regional Classification of climate, *Geographical Review*, 38: 55-94

Trewartha, G. T., and. Horn, L. H. (1980).*An Introduction to Climate* (5thed.) London: McGraw-Hill.

Disclaimer

This self-learning material is based on different books, journals and web-sources.

Post-Graduate Degree Programme (CBCS) in

GEOGRAPHY

Semester-I

Paper: GEO/CC/T-103
Soil and Biogeography

Self-Learning Material

DIRECTORATE OF OPEN AND DISTANCE LEARNING

UNIVERSITY OF KALYANI

Kalyani, Nadia West
Bengal, India

COURSE MATERIALS COMPILED BY			
Sl. No.	**Names**	**Designation and Affiliation**	**Group of the SLM**
1.	Dr. Tarun Kumar Mondal	Assistant Professor Department of Geography, University of KalyaniKalyani-741235, Nadia District, West Bengal	**GEO/CC/T-103: SOIL & BIOGEOGRAPHY**
2.	Bhaswati Pal	Ph.D. Research Scholar Department of Geography, University of KalyaniKalyani-741235, Nadia District, West Bengal	
3.	Dr. Abhay Sankar Sahu	Assistant Professor Department of Geography, University of KalyaniKalyani-741235, Nadia District, West Bengal	
4.	Subhankar Bera	Ph.D. Research Scholar Department of Geography, University of KalyaniKalyani-741235, Nadia District, West Bengal	
5.	Dr. Ayan Rudra	Assistant Professor (Contractual) Directorate of Open and Distance Learning University of Kalyani, West Bengal	

SYLLABUS

Paper–GEO/CC/T-103

SEMESTER-I						
Paper Code	**Paper**	**Theory/ Practical**	**Internal Assessment/ Evaluation**	**Examination/ Report/ Viva-Voce**	**Credit**	**Marks**
GEO/CC/T-103	Soil and Biogeography	Theory	10	40 (Semester-end Examination)	4	50

Unit-1: Soil forming processes with reference to podsolisation, laterisation and calcification
Unit-2: Soil system, soil taxonomy and world pattern of soils
Unit-3: Soil nutrients
Unit-4: Soil organisms
Unit-5: Soil degradation and soil conservation
Unit-6: Concept of integrated management of soil
Unit-7: Plant ecology: concept of adaptation, succession and climax
Unit-8: Dispersal and migration of animals: means and barriers
Unit-9: Ecological footprint
Unit-10: Biodiversity: issues and challenges
Unit-11: International Biological Programme; Man and Biosphere Programme
Unit-12: Wildlife conservation and management: sanctuaries, national parks and biosphere reserves with reference to India

Mode of Internal Evaluation: Class test

Content Page No

SEMESTER- I; PAPER – GEO/CC/T-103; SOIL and BIO- GEOGRAPHY

(Credit – 1; Marks - 50: Internal Evaluation – 10, Semester-end Examination - 40)

Introduction

Soil is an inevitable component of the environment that supports all the living organisms on the earth viz. plants, animals and micro-organisms. Agriculture, the dominant primary economic activity all over the world largely depends on the nature of soil. Soil plays an active role in the biogeochemical cycles. Therefore, an in-depth knowledge of soil is essential to maintain environmental sustainability.

Biogeography is a sub-field of the discipline of geography. Geography is a spatial science. Biogeography studies the spatial patterns on distributions and differentiations of various biological forms, starting from a single species to the entire biosphere, over the earth's surface. Thus, it studies both the physical components of the earth's surface, such as land, soil, water etc. as well as the organisms of an area in a holistic way. In this sense, it is a synthesis of the disciplines of ecology, botany, zoology, geography, geology, climatology etc. This self-learning material will help you to understand the nature, scope, content and significance of biogeography; different aspects related to plant ecology; impact of climate and soil on the distribution of plants, and also to understand the biodiversity and different measures taken to keep the balance between human beings and environment.

Learning objectives

1. To highlight soil taxonomy and world pattern of soils.
2. To learn the concept of soil nutrients and soil organisms.
3. To examine the relation of soil nutrients with soil organisms and micro-organisms.
4. To study the concept of soil pollution and soil degradation.
5. To understand the concept of integrated management of soil.
6. To study the intrinsic characteristics of biogeography through interdisciplinary study ofknowledge.
7. To study the organisms and their complex environmental relationships.
8. To know the spatial and temporal distribution of plants and animals kingdom and factors those influencing them.
9. To evaluate the anthropogenic impact on environment and interactions.
10. To study the present environmental degradation and survival questions of organisms.

Assessment of prior knowledge

General idea on soil profiles and horizons, chemical, physical and biological properties of soil and major soil forming processes will be considered to assess prior knowledge of the learners. The following questions will be asked to assess the prior knowledge in this field:

1. What is soil?
2. Define soil profile and soil horizon.
3. What is soil forming process?
4. What are the main components of soil?
5. What do you mean by properties of soil?
6. Is soil renewable or non-renewable resource?
7. What types of soil do you observe in your locality?

Knowledge of biogeography as a sub-discipline of geography is very much inter-linked with the knowledge of environment. Biogeography in many ways contains same elements of the environment. For the purpose of assessment of prior knowledge on biogeography of the students the following questions may be asked:

i. What do you mean by 'bio'?
ii. What is meant by 'biogeography'?
iii. What is 'biome'?
iv. How does climate control the distribution of plants over the earth surface?

Learning activities

Concept of soil taxonomy and the nature of world pattern of soils will be learned. Soil classifications in the Higher Categories (Thorp & Smith, USDA, 1949) and major soil groups and orders of India will be explained. Concept of soil nutrients and soil organisms will be discussed and their relationship will be leaned. The concept of soil pollution and soil degradation along with causes, consequences and controlling measures will be studied. The concept of integrated management of soil will be learned.

Knowledge of biogeography is based on distribution of plants and animals kingdom over the earth surface. To understand the theoretical backgrounds behind spatial distribution and differentiation of organisms, controlling factors, growth and development, threats and future conservation strategies, a practical documentation of the existing plants and animals on local as well as regional level may be done. Students may prepare a notebook mentioning existing names and numbers of the plants and animals considering a selected study area with physical descriptions like topography, climate, soil, landuse/

landcover, etc. of the study area. For better understanding of the present situation of biodiversity and its related threats along with management strategies, different governmental and non-governmental websites, films, reports, pictures, etc. are to be observed.

Feedback of learning activities

- Making chart on Soil Taxonomy
- Preparation of maps showing distribution of soil
- Preparation of model on soil conservation
- Class test
- For the purpose of feedback from students on their learning activities group discussions can be arranged. An internal assessment may be taken. During feedback, problem areas would be identified and necessary steps would be taken up to resolve them.

EXAMPLES AND ILLUSTRATIONS (SUBJECT CONTENT)

UNIT-1: SOIL FORMING PROCESSES WITH REFERENCE TO PODSOLISATION, LATERISATION AND CALCIFICATION

INTRODUCTION

The geological weathering produces weathered rock material i.e. the parent material and when the genetic factors set the stage for soil development. The pedogenic processes change the parent material in to soil with varying horizonations. Pedogenesis can be defined as the process of soil development. Late in the 19th century, scientists Hilgard in the United States and the Russian Dukuchaev both suggested independently that pedogenesis was principally controlled by climate and vegetation. This idea was based on the observation that comparable soils developed in spatially separate areas when their climate and vegetation were similar. In the 1940s, Hans Jenny extended these ideas based on the observations of many subsequent studies examining the processes involved in the formation of soils. Jenny believed that the kinds of soils that develop in a particular area are largely determined by five interrelated factors: climate*;* living organisms*;* parent material*;* topography*;* and time*.*

Climate plays a very important role in the genesis of a soil. On the global scale, there is an obvious correlation between major soil types and the Köppen climatic classification systems major climatic types. At regional and local scales, climate becomes less important in soil formation. Instead, pedogenesis is more influenced by factors like parent material, topography, vegetation, and time. The two most important climatic variables influencing soil formation are temperature and moisture. Temperature has a direct influence on the weathering of bedrock to produce mineral particles. Rates of bedrock weathering generally increase with higher temperatures. Temperature also influences the activity of soil microorganisms, the frequency and magnitude of soil chemical reactions, and the rate of plant growth. Moisture levels in most soils are primarily controlled by the addition of water via precipitation minus the losses due to evapotranspiration. If additions of water from precipitation surpass losses from evapotranspiration, moisture levels in a soil tend to be high. If the water loss due to evapotranspiration exceeds inputs from precipitation, moisture levels in a soil tend to be low. High moisture availability in a soil promotes the weathering of bedrock and sediments, chemical reactions, and plant growth. The availability of moisture also has an influence on soil pH and the decomposition of organic matter.

Living Organisms have a role in a number of processes involved in pedogenesis including organic matter accumulation, profile mixing, and biogeochemical nutrient cycling. Under equilibrium conditions, vegetation and soil are closely linked with each other through nutrient cycling. The cycling of nitrogen and carbon in soils is almost completely controlled by the presence of animals and plants. Through litterfall and the process of decomposition, organisms add humus and nutrients to the soil which influences soil structure and fertility. Surface vegetation also protects the upper layers of a soil from erosion by way of binding the soils surface and reducing the speed of moving wind and water across the ground surface.

Parent Material refers to the rock and mineral materials from which the soils develop. These materials can be derived from residual sediment due to the weathering of bedrock or from sediment transported into an area by way of the erosive forces of wind, water, or ice. Pedogenesis is often faster on transported sediments because the weathering of parent material usually takes a long period of time. The influence of parent material on pedogenesis is usually related to soil texture, soil chemistry, and nutrient cycling.

Topography generally modifies the development of soil on a local or regional scale. Pedogenesis is primarily influenced by topography's effect on microclimate and drainage. Soils developing on moderate to gentle slopes are often better drained than soils found at the bottom of valleys. Good drainage enhances an number of pedogenic processes of illuviation and eluviation that are responsible for the development of soil horizons. Under conditions of poor drainage, soils tend to be immature. Steep topographic gradients inhibit the development of soils because of erosion. Erosion can retard the development through the continued removal of surface sediments. Soil microclimate is also influenced by topography. In the Northern Hemisphere, south facing slopes tend to be warmer and drier than north facing slopes. This difference results in the soils of the two areas being different in terms of depth, texture, biological activity, and soil profile development.

Time influences the temporal consequences of all of the factors described above. Many soil processes become steady state overtime when a soil reaches maturity. Pedogenic processes in young soils are usually under active modification through negative and positive feedback mechanisms in attempt to achieve equilibrium.

LATERIZATION

The term laterite is derived from the word "later" means "brick" or "tile". In tropics, certain soils are massively impregnated with sesquioxides to the extent of 70 to 80% of the total mass, and forms a cemented horizon, which when dried becomes very hard like a brick. This soil forming process is called "laterization" or "Lotozation" Eg: Soils of Malabar hills of Kerala. In laterization, unlike podzolisation, silica is removed leaving sesquioxides to remain in solum. The favorable conditions are: -

- Warm and humid (tropical) climate with 2000 to 2500 mm rainfall and continuous high temperature (+25°C) throughout the year. Rapid dec omposition of parent material and organic matter, and intensive leaching are very likely in this climate.
- The rain forests of tropical areas are the suitable vegetation for this process. Under this vegetation organic additions are low but organic matter de composition is at very high rate.
- Basic parent materials, having sufficient ferro-mangnesion minerals (Pyroxenes, amphiboles, biotite and chlorite) are congenial for the development of laterites.

The iron released during weathering is oxidized to form FeO, Fe2O 3 and coats clay, silt or sand particles imparting characteristic red color to soils. The Al-oxides /hydroxides imparts grey coatings to the soil particles.

The high temperature, intense leaching and basic kind of parent material all favor the removal of silica (de-silication) and accumulation of sesquioxides. The soluble basic cations are quickly released during weathering, moves freely in the soil profile and shoots up the pH to neutrality. Under this basic environment silica liberated from parent material is solubilized and leached. The solubility of quartz and amorphous silica increases with increased temperature. The sesquioxides are left behind as these are more stable under these conditions. As the alkaline bases are removed from the seat o f their formation, the residual soil is acidic in reaction. Though considerable eluviation takes place, there is no marked horizonsiation as the eluviated materials are not re-deposited in the lower layers.

Laterite soils are non-plastic, non cohesive and have granular structure. They are low in cation exchange capacity and fertility. Phosphorus fixation is high in these soils.

PODZOLIZATION

It is a process of soil formation resulting in the formation of Podzols and Podzolic soils. In many respects, podzolization is the negative of calcification. The calcification process tends to concentrate calcium in the lower part of the B horizon, whereas podzolization leaches the entire solum of calcium carbonates.

Apart from calcium, the other bases are also removed and the whole soil becomes distinctly acidic. In fact, the process is essentially one of acid leaching. The process operates under favorable combination of the following environments.

i) Climate: A cold and humid climate is most favorable for podzolization
ii) Parent material: Siliceous (Sandy) material, having poor reserves of weather able minerals, favor the operation of podzolization as it helps in easy percolation of water.
iii) Vegetation: Acid producing vegetation such as coniferous pines is essential
iv) Leaching and Translocation of Sesquioxide: In the process of decomposition of organic matter various organic acids are produced. The organic acids thus formed act with Sesquioxide and the remaining clay minerals, forming organic- Sesquioxide and organic clay complexes, which are soluble and move with the percolating water to the lower horizons (Bh, Bs). Aluminium ions in a water solution hydrolyze and make the soil solution very acidic. As iron and aluminium move about, the A horizon gives a bleached grey or ashy appearance. The Russians used the term Podzols (pod means under, the Zola means ash like i.e. ash-like horizon appearing beneath the surface horizon) for such soils.

To conclude, the Podzolization is a soil forming process which prevails in a cold and humid climate where coniferous and acid forming vegetations dominate. The humus and Sesquioxide become mobile and leached out from the upper horizons and deposited in the lower horizon.

CALCIFICATION:

The calcification process results in. the redistribution of calcium carbonate or carbonate of lime in the soil profile without complete removal of it. Magnesium carbonate accumulates along with the carbonate of lime. The areas so affected arc normally those of restricted rainfall—varying from approximately 25 inches or less in the Temperate Zone to approximately 45 inches or less in the Tropics— and the dominant vegetation is grass or brush. Since the rainfall is low, the percolation of water through the profile is not sufficient to remove wholly the calcium carbonate that existed in the parent material or was produced by reaction between carbonic acid and the calcium hydrolyzed from silicate minerals. The usual result is the development of an accumulation of calcium and magnesium carbonates at some point in the profile below the surface, approximately the depths to which surface waters most frequently percolate. Marbut (240) called these soils Pedocals (soils with lime accumulation),lo While this is the normal formation, a calcium-carbonate deposit does not necessarily always occur, A secondary result of calcification is that the calcium tends to keep the colloid (fine clay) in a somewhat granular condition, and there is therefore relatively little downward movement of the colloid in the profile.

It has been possible to arrange a multitude of Pedocal soil series into a few great groups, using their color, organic-matter content, depth, and amount of lime accumulation as criteria for making combinations. The group names are largely of Russian origin, and their application to soils of the United States was first made by C. V. Marbut, who modified some of them. In the present classification further modification seems advisable. They are: (1) Chernozem (black earth); (2) Chestnut soils (dark-brown soils); (3) Reddish Chestnut soils; (4) I5rown soils; (5) Reddish Brown soils; (6) Sierozem soils; (7) Desert soils; (8) Red Desert soils.

Chernozems are very dark brown or black in the upper 2 to 4 feet and have slightly acid or slightly alkaline reaction and a nutlike structure. Organic matter is high, and the soils are naturally very fertile. A yellowish-brown or grayish-brown transitional horizon of a few inches separates the dark surface soil from the very calcareous horizon of lime accumulation. The lime is silty in character, includes more or less magnesium carbonate, and occurs as irregular soft masses and vertical streamers, extending to a dex)th of several feet in many places. Parent material below is less calcareous and in places contains no free lime. Accumulated lime comes partly from lime originally present in the materials and partly from the carbonation of calcium and magnesium silicates.

Chestnut, Brown, and Sierozem soils have prismatic, dark-brown, brown, and light brownish-gray upper horizons, respectively, and progressively less organic matter in the order listed. The thickness of the surface soils also becomes lens in the order listed, and the lime accumulation usually reaches its maximum development in the first two, except in cases of extremely old Sierozems and Red Desert soils. All typical members of these groups except tile

Sierozem and Red Desert soils have little or no free lime at the surface, but the brown horizons in places are calcareous in their lower portions. Calcareous dusts sometimes accumulate sufficiently to render calcareous the surface inch or two. Reddish Chestnut and Reddish Brown soils have a dull reddish tinge in surface horizons and reddish-browar or red heavier subsoils above the horizon of lime accumulation.

Chernozems occupy grasslands with sub humid climate, and Chestnut, Brown, and Sierozem soils occupy regions with progressively drier climate and scantier grass cover. Sierozems have little grass and some small shrubs and are in semi desert regions. Reddish Chestnut, Reddish Brown, and Red Desert soils support a greater percentage of shrubs and less grass than the corresponding groups farther north. True Desert soils have but a scanty growth of perennial shrubs and certain scattered drought-resistant grasses.

Desert and Red Desert soils are in the true deserts—the first in temperate regions and the latter in subtropical and tropical regions. They are usually calcareous to the surface but more so in substrata. Winds remove finer particles from surface horizons and leave a protective desert-pavement cover of small and large rock fragments. In many of the Desert soils the lime accumulation takes the form of hard concretions "or stony formations (caliche) and attains a thickness of many feet in places. Physiographic evidence indicates that some of these lime hardpans or crusts were formed in poorly drained depressions tens of thousands of years ago and have become well-drained during a geologically recent cycle of erosion.

According to Nikiforoff's hypothesis (281) ^ the reddish heavy subsoil of Red Desert soils is due to weathering of minerals in deeper horizons under the influence of moisture which percolates downward during brief rainy periods. This assumption might also be made in regard to the heavy'- subsoil horizons of the Reddish Chestnut and Reddish Brown soils. Black crusts of iron oxide form protective coatings on rocks and i)ebbles of the hotter deserts and are known as desert varnish. The calcification process not only connotes the accumulation of lime in the soil but also the adsorption of calcium ions by the colloids. Grasses and other plants requiring relatively large amounts of bases, particularly of calcium, bring these bases to the surface and, through decay, replenish the losses of leaching. For this reason the surface soils are seldom strongly acid—usually approximately neutral—and may be even faintly alkaline. When colloids are high in calcium (and to a less extent in magnesium) the reaction will usually range from slightly acid to slightly alkaline, and there will be an abundance of calcium available for crops. Brown Forest soils presumably owe their lack of cultivation and illuviation to this kind of calcification and are limited to areas with forest vegetation having a particularly high content of bases in its leaves. Rendzinas are also calcified but, in spite of calcareous parent material, do not always contain free carbonate of lime.

It is the process of precipitation and accumulation of calcium carbonate ($CaCO_3$) in some part of the profile. The accumulation of $CaCO_3$ may result in the development of a calcic horizon. Calcium is readily soluble in acid soil water and/or when CO_2 concentration is high in root zone as:

$$CO_2 + H_2O \rightarrow H_2CO_3$$

$$H_2CO_3 + Ca \rightarrow Ca(HCO_3)_2 \text{ (soluble) Temp.}$$

$$Ca(HCO_3)_2 \rightarrow CaCO_3 + H_2O + CO_2 \text{ (precipitates)}$$

$$CO_2$$

The process of precipitation after mobilization under these conditions is called calcification and the resulting illuviated horizon of carbonates is designated as Bk horizon (Bca).

UNIT -2: SOIL SYSTEM, SOIL TAXONOMY AND WORLD PATTERN OF SOILS

CONCEPT OF SOIL TAXONOMY (ST)

According to Brady (2002), soil taxonomy is 'the science of classification of soils; laws and principles governing the classifying of the soil' (p. 596).

Bockheim (2014) stated that 'Soil Taxonomy is a hierarchical system that is based primarily on diagnostic surface horizons (epipedons) and subsurface horizons that appear in the system at different taxonomic levels' (p. v).

Arnold and Eswaran (2002) mentioned, 'Soil Taxonomy is structured as a nested hierarchy, in which classes of the lower levels are an integral part of and confined by the defining properties of the classes at higher levels' (p.31).

According to Foth (1990), 'diagnostic surface and subsurface horizons, together with other diagnostic features, are used to classify soils in Soil Taxonomy' (p. 284). He also mentioned six categories of soil taxonomy; from highest to lowest category as 'order, suborder, great group, subgroup, family, and series' (Foth, 1990, p. 279).

Soil Survey Staff.(1999) identified the following two objectives of soil taxonomy:

- 'To establish hierarchies of classes that permit us to understand, as fully as possible, the relationship among soils and between soils and the factors responsible for their character' (p. 15).
- 'To provide a means of communication for the discipline of soil science' (p. 15).

The term soil taxonomy bears a narrow concept compared to the term soil classification. According to Soil Survey Staff.(1999), 'classification includes taxonomy, but it also includes the grouping of soils according to limitations that affect specific practical purposes' (p. 15), whereas,the term 'taxonomy is the part of classification' (p. 15).

Soil classification system

THE USDA (UNITED STATES DEPARTMENT OF AGRICULTURE) SOIL CLASSIFICATION SYSTEM

Curtis Fletcher Marbut, Director of the Soil Survey Division in USA presented a comprehensive soil classification scheme for the United States in 1935. Marbut's system was modified by Mark Baldwin, Charles Kellogg, and James Thorp in 1938. This soil system was further modified by James Thorp and Guy D. Smith in 1949. The modified system is widely known as USDA Soil Classification System 1949 (Table No. 2). In this USDA soil system (1949), soils are dividedinto three orders as follow (Table No.1):

Table No. 1: SOIL ORDERS (USDA SOIL CLASSIFICATION SYSTEM 1949)

Soil orders	Major characteristics
A. Zonal soils	i) Mature soils and develop over sufficient time. ii) Well defined soil profiles and distinct soil horizons. iii) Soils formation and development are largely controlled by thenature of climate, vegetation and organism. v) Soils develop over well drained surface.
B. Intrazonal soils	i) Thesesoils are found within zonal soils. ii) On an average, these soils reflect well developed and distinctfeatures. ii) Soil formation and development largely depend on local factors, viz. nature of terrain, parent materials, surface and subsurfacedrainage.
C. Azonal soils	i) Immature soils and develop over short period of time. ii) Lack of time, distinct nature of terrain and parent materials impedethe proper development of soil profile. ii) 'They do not have profile development and structure developed from the soil-forming processes (Fang and Chaney, 2017, p. 56).

Table No. 2: SOIL CLASSIFICATIONS IN THE HIGHER CATEGORIES (THORP & SMITH, USDA, 1949)

Order	Suborder	Great Soil Groups
	1. Soils of the cold zone	Tundra soils
	2. Light-colored soils of the aridregions	Desert soils Red desert soils Sierozem Brown soils Reddish-brown soils
	3. Dark-colored soils of	Chestnut soils Reddish Chestnut soils Chernozem soils

A. Zonal soils	semi-arid, sub-humid and humid grasslands	Prairie soils Reddish Prairie soils
	4. Soils of the forest-grasslandtransition	Degraded Chernozem Noncalcic Brown or Shantung Brown soils
	5. Light-colored podzolized soilsof the timbered regions	Podzol soils Gray wooded or Gray Podzolic soils Brown Podzolic soils Gray-Brown Podzolic soils Red-Yellow Podzolic soils
	6. Lateritic soils of forested warm- temperature and tropical regions	Reddish-Brown Lateritic soils Yellowish-Brown Lateritic soils Laterite soils
B. Intrazonal soils	1. Halomorphic (saline and alkali) soils of imperfectly drained arid regions and littoral deposits	Solonchak, or Saline soils Solonetz soils Soloth soils
	2. Hydromorphic soils of marshes, swamps, seep areas, and flats	Humic-Glei soils (includes Wie- senboden) Alpine Meadow soils Bog soils Half-Bog soils Low-Humic-Glei soilsPlanosols Ground-Water Podzol soils Ground-Water Laterite soils
	3. Calcimorphic soils	Brown Forest soils (Braunerde) Rendzina soils
C. Azonal soils		Lithosols Regosols (includes Dry Sands) Alluvial soils

Source: Thorp and Smith (1949, p. 118)

Biswas and Mukherjee (1994, pp. 379-380) presented soil groups in India, adapted from Soils of India by Raychaudhuri and GovindaRajan, 1971 and Studies on Soils of India by GovindaRajan and GopalaRao, 1978 (Table No. 3).

Table No. 3: SOIL GROUP AND SOIL ORDER OF INDIA (BISWAS AND MUKHERJEE, 1994)

Soil group	**Soil Order (Soil Taxonomy)**
1. Red loamy	Alfisols
2. Red sandy	Alfisols
3. Laterite	Ultisols, Oxisols
4. Red and yellow	Ultisols
5. Shallow black	Entisols, Inceptisols
6. Medium black	Vertisols
7. Deep black	Vertisols
8. Mixed red and black	Association of Alfisols and Vertisols
9. Coastal alluvium	Entisols
10. Coastal sands	Entisols
11. Deltaic alluvium	Entisols, Alfisols
12. Alluvial Khadar (recent) Bhangar (old)	Entisols Inceptisols Alfisols
13. Alluvial (Highly calcareous)	Entisols
14. Calcareous sierozemic	Aridisols
15. Grey brown	Aridisols
16. Desert - Regosolic	Aridisols Entisols
17. Desert - Lithosolic	Entisols
18. Tarai	Mollisols
19. Brown hill	Ultisols
20. Sub-montane (Podzolic)	Alfisols
21. Mountain meadow	Mollisols
22. Saline and alkali	Aridisols, Entisols, Vertisols
23. Peaty and saline peaty	Histosols
24. Skeletal	Entisols
25. Glaciers and eternal snow	

Source: Biswas&Mukherjee (1994, pp. 379-380)

SOIL ORDERS

Soil order is the largest category in Soil Taxonomy. According to Brady (2002), 'the order category is based largely on soil forming processes as indicated by the presence or absence of major diagnostic horizons. A given order includes soils whose properties suggest that they arenot too dissimilar in their genesis' (p. 63). In general, the following eleven soil orders have been identified in Soil Taxonomy (Brady, 2002).

- Alfisols: 'Alfisols are moist mineral soils having no mollicepipedon on oxic or spodic horizons. They have gray to brown surface horizons (commonly an ochricepipedon), medium- to high- base status, and contain an alluvial horizon in which silicate clays have accumulated' (Brady, 2002, p. 69).

- Andisols (Andosols): Andisols are found around the volcanoes and formed by the materials (e.g. tephra) ejected from volcanic eruptions. Andosols are characterized with high amount of 'active Al and Fe materials, and the lowest bulk density among mineral soils' (FAO & ITPS, 2015, p. 559).

- Aridisols: Aridisols are found mainly in arid regions. 'Aridisols are dry soils characterized by an ochricepipedon generally light in color and low in organic matter. They may have a horizon of accumulation of calcium carbonate (calcic), gypsum (gypsic), soluble salts (salic) or sodium (natric)' (Brady, 2002, p. 74).

- Entisols: These soils are mostly the 'weakly developed mineral soils without natural genetic (subsurface) horizons or with only the beginnings of such horizons' (Brady, 2002, p. 66). Most of the Entisols have 'an orchicepipedon and a few have man-made anthropicor agricepipedons' (Brady, 2002, p. 66).

- Histosols: Histosols soils contain high percentage of organic matter. These soils are generally referred as 'bogs, moors or peats and mucks' (Biswas& Mukherjee, 1994, p. 328). Four subdivisions of Histosols, viz. Folists, Fibrists, Hemists and Saprists are mentioned in USA Soil Taxonomy.
- Inceptisols: 'The horizons of Inceptisols are thought to form quickly and result mostly from the alteration of parent materials. Inceptisols have ochric or umbricepipedons

and/orcambic subsurface horizons’ (Brady, 2002, p. 67).

- Mollisols: Mollisols are the base-rich dark coloured soils observed in steppes regions (Biswas& Mukherjee, 1994). ‘Mollisols are characterized by a mollicepipedon or surface horizon which is thick, dark and dominated by base-forming cations’ (Brady, 2002, p. 67).
- Oxisols: These soils are infertile with ‘reddish, yellowish or grayish’ colour formed in tropical and subtropical climatic conditions (Biswas& Mukherjee, 1994). ‘The unique properties of Oxisols are extreme weathering of most minerals other than quartz to kaolinand free oxides; very low activity of the clay fraction and loamy or clayey texture’ (Biswas& Mukherjee, 1994, p. 329).
- Spodosols: ‘Spodosols are mineral soils that have a spodic horizon, a subsurface horizon with an accumulation of organic matter and of oxides of aluminum with or without iron oxides’ (Brady, 2002, p. 74).
- Ultisols: These soils are commonly found in warm and humid climatic regions with seasonal shortage of rainfall (Biswas& Mukherjee, 1994). Ultisols are characterized by ‘argillic horizon, low supply of bases, particularly in lower horizons and mean annual soiltemperature higher than 8^{o}C’ (Biswas& Mukherjee, 1994, p. 330).
- Vertisols: ‘The Vertisols order of mineral soils is characterized by a high content (>30%) of sticky or swelling and shrinking-type clays to a depth of 1 m, which in dry seasons causes the soils to develop deep, wide cracks’ (Brady, 2002, p. 73). These soils are mostly found in sub-humid to semiarid regions with average soil temperature more than 8^{o}C (Brady, 2002).

UNIT – 3: SOIL NUTRIENTS

CONCEPT OF SOIL NUTRIENTS

Soil nutrients largely determine soil fertility. Soil nutrients are essential for plant growth. Essential elements for plants nutrients are divided into two broad categories i.e.

A). Macronutrients

B) Micronutrients.

A.Macronutrients - 'A chemical element necessary in large amounts (usually 50mg/kg in plant) for the growth of plants' (Brady, 2002, p. 589). Macronutrients are classified into three categories i.e.

I. Structural nutrients e.g. Carbon (C), Hydrogen (H) and Oxygen (O);

II. Primary nutrients e.g. Nitrogen (N), Phosphorus (P) and Potassium (K);

III. Secondary nutrients e.g. Calcium (Ca), Magnesium (Mg) and Sulfur (S).

B.Micronutrients - 'A chemical element necessary in only extremely small amounts (<50mg/kg in the plant) for the growth of plants' (Brady, 2002, p. 589). Examples are Boron (B), Copper (Co), Iron (Fe), Manganese (Mn), Zinc (Zn), Molybdenum (Mo), Chlorine (Cl), Cobalt (Co) and Nickel (Ni).

FUNCTIONS OF SOIL NUTRIENTS

'As a precondition for growth, health and the production of nutritious food, plants require essential nutrients (macro and micronutrients) in sufficient quantities (IFA, WFO & GACSA, 2016, p.1). According to Biswas and Mukherjee (1994) 'the essential elements exist as structural components of a cell, maintain cellular organizations, functions in energy transformation and in enzyme reactions' (p. 224) (Table Nos. 4 and 5).

Table No. 4: FUNCTIONS OF MACRONUTRIENTS

Macronutrients	**Functions**
Carbon (C),	'Constituent of carbohydrates; necessary for photosynthesis' (Jones & Olson-Rutz, **2016, p. 2).**
Hydrogen (H)	'Maintains osmotic balance; important in numerous biochemical reactions; constituent of carbohydrates' (Jones & Olson-Rutz, **2016, p. 2).**

Oxygen (O)	'Constituent of carbohydrates, necessary for respiration' (Jones & Olson-Rutz, **2016, p. 2).**
Nitrogen (N)	'Nitrogen is needed by plants for the production of proteins, nucleic acids (DNA and RNA) and chlorophyll' (McCauley et al., 2011, p. 5).
Phosphorus (P)	'Constituent of many proteins, coenzymes, nucleic acids and metabolic substrates; important in energy' (Jones & Olson-Rutz, **2016, p. 2).**
Potassium (K)	'Potassium is utilized by plants in the activation of enzymes, photosynthesis, protein formation and sugar transport' (McCauley, et al., 2011, p. 9).
Calcium (Ca)	Activates plant enzymes, forms the constituent of cell wall and helps in maintaining cell membrane stability (Biswas& Mukherjee, 1994).
Magnesium (Mg)	'Enzyme activator, component of chlorophyll'(Jones & Olson-Rutz, **2016, p. 2).**
Sulphur (S)	'Essential constituent of certain amino acids and proteins, S deficiency results in the inhibition of protein and chlorophyll synthesis' (McCauleyet al., 2011, p. 10).

Sources: Biswas& Mukherjee (1994); McCauley et al. (2011); Jones & Olson-Rutz (**2016)**

Table No. 5: FUNCTIONS OF MICRONUTRIENTS

Micronutrients	**Functions**
Zinc (Zn)	'Involved with enzyme systems that regulate various metabolic activities (Jones & Olson-Rutz, 2016, p. 2).
Iron (Fe)	'Present in several peroxides, catalase, and cytochrome oxidase enzymes' (Brady, 2002, p. 383); takes part in oxidation-reduction reactions and plays an active role in the formation of chlorophyll (Brady, 2002).
Copper (Co)	'Copper is needed for chlorophyll production, respiration and protein synthesis' (McCauley et al. 2011, p. 12).

Manganese (Mn)	‘Activates decarboxylase, dehydrogenase and oxidase enzymes’ (Brady, 2002, p. 383); actively involves in the processes of photosynthesis as wellas metabolism and assimilation of nitrogen (Brady, 2002).
Boron (B)	‘Believed to be important in sugar translocation and carbohydrate metabolism’(Jones & Olson-Rutz, 2016, p. 2).
Molybdenum (Mo)	‘Involved with nitrogen fixation and transforming nitrate to ammonium’ (Jones & Olson-Rutz, 2016, p. 2).
Cobalt (Co)	Cobalt plays a significant role in nitrogen fixation (Brady, 2002).

Sources: Brady (2002); McCauley et al. (2011); Jones & Olson-Rutz (2016)

SOURCES OF SOIL NUTRIENTS

Major sources of soil nutrients (IFA, WFO & GACSA, 2016):

a) Weathering of rocks is the primary source of soil nutrients. This process releases nutrients in small amount into the soil.
b) Nutrients applied for crop production are not totally used. Residual nutrients are retained in the soil.
c) A significant amount of nutrients (e.g. nitrogen and sulphur) is added into the soil through atmospheric deposition process.
d) Crop residues are the important sources of soil nutrients. Mainly nitrogen (N), phosphorus (P) and sulphur (S) are mixed with the soil from decomposition of crop residues.
e) Compost is also a source of soil nutrients.
f) Livestock manure is a significant source of soil nutrients. Nitrogen, phosphorus, potassium, calcium, magnesium and sulphur are mainly added to soil from livestock manures.
g) Biosolids derived from urban waste water treatment are a source of soil nutrients.
h) By the process of Biological Nitrogen Fixation (BNF), mainly atmospheric nitrogen is converted to ammonia and then to nitrates. This is the main process of converting atmospheric nitrogen to soil nutrients.
i) Fertilizers are manufactured in industry and farmers apply fertilizers to the soil to raise crops.

UNIT – 4: SOIL ORGANISMS

INTRODUCTION

Soil organisms play an active role in the availability, transformation and supply of soil nutrients. According to Hesammi et al. (2014), 'soil biodiversity reflects the mix of living organisms in the soil. These organisms interact with one another and with plants and small animals forming a web of biological activity (p. 10328). He also stated that 'soil organisms contribute a wide range of essential services to the sustainable function of all ecosystems' (Hesammi et al., 2014, p. 10328).Classification of soil organisms (Brady, 2002, p.254; Biswas& Mukherjee, 1994, p. 46):

1. Soil flora (plants): Soil flora is divided into two broad categories i.e.

a) Soil macroflora: Roots of higher plants.

b) Soil microflora: (i) Bacteria (ii) Fungi (iii) Algae (iv) Actinomycetes

2. Soil fauna (animals): Soil fauna is also divided into two broad categories, i.e.

a) Soil macrofauna: (i) Largely herbivores and detritivores (ii) Largely predators

b) Soil microfauna: (i) Largelydetritivores, predators and parasites

ROLE OF SOIL ORGANISMS

The role and function of soil flora and fauna are indispensable in their influence on crop production 'through a number of processes such as decomposition, synthesis and transformation' (Biswas& Mukherjee, 1994, p. 70). 'During decomposition, soil organisms release essential plant nutrients in inorganic forms that can be absorbed by plant roots or leached from the soil (Brady, 2002, p. 276).

Hesammi et al. (2014) summarized the three key roles of soil organisms:

j) 'Decompose organic residues' ii) 'Re-cycle nutrients from organic residues' iii) 'Enhance soil structure' (p. 10331).

Soil organisms improve physical properties of the soil. Organisms play an active role for the formation of better soil structure. According to Hesammi et al. (2014), 'soil with good structure has many beneficial effects including enhanced water transmission into and through soil, lower bulk density and lower potential for soil erosion' (p. 10331).

Soil organisms degrade soil pollutants through biodegradation process. Micro-organisms help to degrade specific soil contaminants by bioremediation process.

Soil organisms decompose various organic residues and significant amount of nutrients are added to the soil in this way.

Organisms are the vital factor in soil formation. Organisms control the process of biomass formation and the process of biological weathering in creation of parent materials that are essential part of soil formation.

Aislabie and Deslippe (2013) mentioned that 'microbes contribute to soil formation through nutrient cycling and organic matter production. Microbial products are critical to soil aggregation, improved soil structure making soil more habitable for plants (p. 144).

Brady (2002) stated that 'organically bound forms of nitrogen, sulfur, and phosphorous are converted by the microbes into plant available forms' (p.275).

According to Aislabie and Deslippe (2013), microbes determine the process of mineralization of soil carbon and nutrients and consequently 'the carbon storage capacity of soils'. 'Denitrifying bacteria and fungi and methane producing and consuming bacteria regulate nitrous oxide (N2O) and methane (CH4) (Aislabie&Deslippe, 2013, p. 144).

MICRO-ORGANISMS AND THEIR RELATION WITH SOIL FERTILITY

According to Powlson et al. (2001), 'the soil micro-organisms (collectively the soil microbial biomass) are the agents of transformation of soil organic matter, nutrients and of most key soil processes. Their activities are much influenced by soil physico-chemical and ecological interactions (p. 41).

CLASSIFICATION OF MICROORGANISMS

Soil microorganisms are classified into two broad categories i.e. (a) Soil microflora and (b) Soil microfauna.

- Soil microflora (Brady, 2002, p.254; Biswas& Mukherjee, 1994, p.46): Algae (eg. Blue-green Algae, Yellow-green Algae and Diatoms), Fungi (eg. Mushrooms, Yeasts and Moulds), Actinomycetes (eg. Actinomycelaceae and Streptomycetaceae) and Bacteria (eg. Aerobic, Anaerobic, Autotropic and Heterotrophic).
- Soil microfauna (Protozoa, Nematodes and Rotifers).

ROLE OF MICRO-ORGANISMS IN SOIL FERTILITY

To know the role of micro-organism in soil fertility, it is essential to understand the concept of soil fertility. Foth (1990) defined soil fertility 'as the ability of a soil to supply essential elements for plant growth without a toxic concentration of any element' (p. 9-10). The term soil fertility is defined as 'the inherent capacity of soil to provide nutrients, in adequate amounts and in proper balance, for the growth of specified plants when other growth factors such as light, water and temperature, and the physical condition of the soil are favourable' (Biswas& Mukherjee, 1994, p.222).

Bollen (1959) mentioned two types of soil fertility: active and potential. According to him, 'active fertility is immediately available; potential fertility becomes available by chemical or microbial action on minerals and organic matter. The function of soil microorganisms is to render potential fertility available' (p.1).

Micro-organisms play a significant role in Biological Nitrogen Fixation (BNF). Plants cannot absorb elementary nitrogen from the atmosphere directly. 'The nodule organisms, especially those of legumes, free-fixing bacteria of several kinds, and some actinomycetes are most noted for their ability to fix nitrogen' (Brady, 2002, p.275). Hence, Biological Nitrogen Fixation is the most significant natural way to maintain soil fertility.

Biofertilizers are the important sources of soil nutrients. 'Biofertilizers are the cultures of micro- organisms used for inoculating seed or soil or both under ideal conditions to increase the availability of plant nutrients' (Biswas& Mukherjee, 1994, p. 84). Biofertilizers are broadly classified into two sub groups 'inoculants of specific organisms such as Rhizobia, Azotobacter, blue-green algae, phosphate solubilizers, cellulolytic microorganisms; and biomass producing organisms, e.g. Azolla' (Biswas& Mukherjee, 1994, p. 84).

Micro-organisms make the soil nutrients available to plants through various microbial activities. 'Organically bound forms of nitrogen, sulfur, and phosphorus are converted by the microbes intoplant-available forms' (Brady, 2002, p. 275).

Decomposition of organic matters and humus formation are the significant processes to enrich soil nutrients and to maintain nutrient cycle. Moreover in order to obtain energy 'soil organisms break down organic matter, aid in the production of humus, and leave behind compounds that areuseful to higher plants' (Brady, 2002, p. 275). Micro-organisms are actively involved in 'synthesis of microbial cell substance, oxidation and reduction of soil constituents' (Waksman & Starkey, 1931, p. 243) that enhance soil fertility.

UNIT -5: SOIL DEGRADATION AND SOIL CONSERVATION

SOIL DEGRADATION

'Soil degradation inherently reduces or eliminates soil functions and their ability to support ecosystem services essential for human well-being' (FAO & ITPS, 2015, p. 180).

According to Jie et al. (2002), 'soil degradation, defined as lowering and losing of soil functions, is becoming more and more serious worldwide in recent decades, and poses a threat to agricultural production and terrestrial ecosystem' (p. 1).

Young et al. (2015) stated that 'soil degradation occurs through the deterioration of the physical, chemical and biological properties of soil that results in soil compaction, salinisation, acidification and soil loss from wind and water erosion (p. 2).

TYPES OF SOIL DEGRADATION

Soil degradation takes place through wind and water erosion as well as physical and chemical degradation (Table No. 6).

Table No. 6: TYPES OF SOIL DEGRADATION FOR GLOBAL ASSESSMENT OF SOIL DETERIORATION (GLASOD), (AFTER OLDEMAN, 1992)

Type (code)	Subtype (code)
1. Water erosion (W)	i. Loss of Topsoil (Wt)
	ii.Terrain deformation (Wd).
2.Wind erosion (E)	i. Loss of topsoil by wind erosion (Et),
	ii. Terrain deformation (Ed)
	iii. Overblowing (Eo)
3. Chemical degradation (C)	i. Loss of nutrients and/or organic matter (Cn).
	ii. Salinization (Cs)
	iii. Acidification (Ca)
	iv. Pollution (Cp).
4. Physical degradation (P)	i. Compaction, crusting and sealing (Pc)
	ii. Waterlogging (Pw)
	iii. Subsidence of organic soils (Ps)

Source: Oldeman (1992, pp. 19-36).

EFFECTS OF SOIL DEGRADATION

The major effects of soil degradation include -

- Soil erosion leads to removal of top fertile soils and soil fertility is reduced.
- Contaminated soil with heavy metal and toxic elements destroys soil microorganisms which maintain the soil fertility and bio-geochemical cycles.
- Agricultural productivity decreases as result of the loss of soil nutrients.
- Soil degradation increases desertification that adversely affects the existing ecosystem.
- Soil degradation leads to increase in soil salinity and soil alkalinity.
- Soil degradation increases water-logging conditions and desertification.
- Soil degradation threatens the biodiversity.
- Soil degradation ultimately reduces the amount of arable lands. This situation accelerate food crisis.

STRATEGIES TO MITIGATE SOIL DEGRADATION

Effective strategies to mitigate soil degradation are as follow:

- Afforestation and reforestation are the most effective strategies to reduce soil degradation.
- Mulching, contour farming and terracing, strip farming, shelter belts and intercropping are to be introduced.
- Crop rotation is the most important farming technique to maintain soil fertility naturally.
- Integrated Pest Management and Integrated Soil Fertility Management have to be implemented.
- Diversified cropping and integrated farming systems have the potential to counter soil degradation.
- Effective measures should be taken to prevention all sorts of soil contamination and pollution.
- Specific measures have to be taken for reclaiming acidic and saline soils.
- Awareness generation about the causes, effects and preventive measures of soil pollution.

Soil conservation is the prevention of loss of the top most layer of the soil from erosion or prevention of reduced fertility caused by over usage, acidification, salinization or other chemical soil contamination.

Slash-and-burn and other unsustainable methods of subsistence farming are practiced in some lesser developed areas. A sequel to the deforestation is typically large scale erosion, loss of soil nutrients and sometimes total desertification. Techniques for improved soil conservation include crop rotation, cover crops, conservation tillage and planted windbreaks, affect both erosion and fertility. When plants die, they decay and become part of the soil. Code 330 defines standard methods recommended by the U.S. Natural Resources Conservation Service. Farmers have practiced soil conservation for millennia. In Europe, policies such as the Common Agricultural Policy are targeting the application of best management practices such as reduced tillage, winter cover crops,[1] plant residues and grass margins in order to better address the soil conservation. Political and economic action is further required to solve the erosion problem. A simple governance hurdle concerns how we value the land and this can be changed by cultural adaptation.[2] Soil carbon is a carbon sink, playing a role in climate change mitigation.

SOIL CONSERVATION MEASURES

Conservation and restoration of soil is necessary to protect our cultivated farms and expand available land for agriculture with a view to increasing food production for the future.

Conservation measures must therefore fulfill the following objectives:

(i) protection of the surface from the impact of raindrops,

(ii) increase in rainwater infiltration,

(iii) decrease in the volume and velocity of surface runoff,

(iv) enhancement in soil resistance to erosion by judicious modification of the physical and chemical properties of soil resource.

Before initiating soil conservation measures, some steps should be followed:

(i) extensive survey of affected areas,

(ii) classification of agricultural and forest lands on the basis of land capabilities,

(iii) identification of areas affected by low, moderate and severe soil erosion, and

(iv) enlisting the prime priorities of soil conservation and land reclamation.

Broadly, practical methods of soil conservation are: (a) Biological measures and (b) Mechanical measures.

BIOLOGICAL MEASURES

1. Improving the existing surface cover: This can be done by resorting to cover cropping by growing groundnut or berseem (a fodder crop) or through grasslands development by growing grasses like dub, kudzu and pans.

2. Crop rotation: This refers to growing of two or more different crops in sequence in a field maintaining the soil fertility. Continuous growing of clean-cultivated crops (e.g., tobacco) causes more erosion. A good rotation should include densely planted small grains, spreading legume which may check soil erosion.

3. Strip cropping: This practice consists of growing erosion-permitting crops (jowar, bajra, maize) in alternate strips with erosion checking close growing crops (grasses, pulses). The erosion checking strips check and hold the flowing water and soil.

4. Stubble- mulching: This means leaving crop residue or vegetative litter on the land as a surface protection against erosion and for conserving moisture by favouring infiltration and reducing evaporation.

5. Using organic manures: Organic manures, like cow dung, green manure, farmyard manure etc., improve the soil structure. Granular and crumby structures increase infiltration and permeability in the soil and conserve soil moisture. Other measures include checking overgrazing, reducing surplus cattle, stopping shifting cultivation and taking preventive measures against forest fires.

MECHANICAL MEASURES

1. Contour tillage: On sloping lands, all tillage operations should be done at right angles to the slope of the land.

This way, each furrow intercepts the flowing water and allows it to soak into the soil.

2. Contour bunding: The idea is to break the slope of the land into smaller, more level compartments by constructing mechanical structures of suitable size along contours. Each bund, thus, holds the rainwater within each compartment.

3. Terracing: On steeper slopes, terraces or flat platforms are constructed in steps in a series along the slope.

This way water is retained on each terrace which can be used to raise crops.

4. Basin listing: This refers to scooping out small basins at regular intervals on slopes which help in checking the run-off and in conservation of water.

5. Water harvesting: This refers to trapping or channeling of water into low-lying areas. This helps in checking the run-off and also acts as a flood control measure.

6. Scientific slope management: The cropping activity on slopes should be taken up as per the nature of slope. If the slope is between 1:4 and 1:7, proper farming can be done; if more, pastures should be developed; if still more, forestry operations can be undertaken; if it is still greater, then terracing is required before any cropping activity can be done.

UNIT – 6: CONCEPT OF INTEGRATED MANAGEMENT OF SOIL

INTRODUCTION

Soil management is the application of operations, practices, and treatments to protect soil and enhance its performance (such as soil fertility or soil mechanics). It includes soil conservation, soil amendment, and optimal soil health. In agriculture, some amount of soil management is needed both in nonorganic and organic types to prevent agricultural land from becoming poorly productive over decades. Organic farming in particular emphasizes optimal soil management, because it uses soil health as the exclusive or nearly exclusive source of its fertilization and pest control.

The term integrated management of soil is a set of integrated management approaches to protect soil from degradation, to maintain soil health and quality, to enhance agricultural productivity in particular and to sustain soil ecosystem in general. Integrated Soil Management (ISM) is a strategy that ensures sustainable agricultural production and to optimise social and economic benefits and to maintain environmental sustainability.

FRAMEWORK FOR INTEGRATED SOIL FERTILITY MANAGEMENT

1. Reduce soil disturbance: Intensive soil tillage is one of the greatest catalysts for soil degradation. Although plow tillage systems are a longstanding practice for land preparation in many places, they are the most detrimental practice to soil health in both the long and short term. Breaking up the soil and intensively working it with tillage equipment can lead to many problems. Tillage is highly destructive because it disrupts the habitats and populations of soil microorganisms that contribute significantly to maintaining and improving soil health. One of the major problems that modern intensive tillage has caused is the rapid loss of the soil organic matter, which is the food for microbes and a binding agent for soil aggregates. The more the soil is tilled, the higher the rate of organic matter oxidation (organic matter loss) in which soil organic carbon is converted to carbon dioxide and lost to the atmosphere. Therefore, an intensively tilled soil will likely have poor soil health because of the low organic matter content. This is true for many of the arable soils in New Mexico. Using reduced tillage or conservation tillage methods will preserve and improve the soil organic matter over time and will be less disruptive for the soil microorganisms. Many research studies are now showing that reduced tillage methods, such as strip tillage and no

tillage, can produce similar crop yields as the conventional plow-till method in arid and semi-arid farmlands (Darapuneni et al., 2019; Idowu et al., 2019). This means that by transitioning to reduced tillage practices, crop producers can remain profitable in terms of yield and at the same time conserve and improve the soil health of their fields. An added benefit to reduced tillage is fewer field passes for land preparation, leading to savings in fuel and tillage costs. Reducing tillage for crop production is not an easy task. There are multiple challenges that farmers must overcome before achieving effective reduced tillage benefits.

2. Crop rotation practices: Crop rotation is a method of farming in which farmers grow crops from different plant families in a well-planned sequence over multiple seasons in the same field. For example, a farmer may grow alfalfa for three years, after which chile peppers are grown for two years and then cotton for another two years before growing alfalfa again. In this case, the farmer has a seven-year rotation cycle. These three crops–alfalfa, chile, and cotton–belong to different crop families. Crop rotation can be of any length depending on farmers' preferences and other conditions on the farm, or local conservatory regulations. Cover crops also count within any rotation cycle. The key factor in rotation is that a variety of crops belonging to different families are being grown in the same soil.

3. Cover cropping practices: Cover crops are plants that are grown in between cash crop cycles to protect the land from erosion and to add root exudates and biomass for soil improvement. Cover crop biomass must be returned to the soil after the desired growing period for the soil health benefit to be fully realized. If the cover crop is harvested off the field as a forage crop, the soil health improvement derived from the cover crop may be undetectable. Soils of the arid and semi-arid Southwest have very low organic matter contents; therefore, growing a cover crop is one way to increase the soil organic matter through the addition of biomass to the soil. Some cover crops deliver other benefits to the soil. For example, legumes have the capacity to fix nitrogen in the soil by associating with a type of bacteria called rhizobium. Rhizobium can convert atmospheric nitrogen to plant-useable nitrogen. When leguminous cover crops are worked into the soil, the nitrogen that was fixed and translocated into the plant is made available for the subsequent crop after residue decomposition.

4. Diversify production systems: Introducing diversity into your production system will allow you to build a healthy soil. Diversifying your production system can be done in different ways. The key factor is to intentionally increase biodiversity on the farm or range either by increasing the number of crop species planted or by integrating livestock into the cropping system. Diversified farming systems include practices such as mixed varieties and mixed cropping, integrating livestock into crop production, cover cropping, crop rotation, fallowing fields, hedgerows, buffer strips, and many other practices (Kremen et al., 2012). All these practices introduce new species into the farming system that can improve soil health and support agricultural biodiversity. Other benefits of diversified farming systems include improvements in nutrient cycling, soil water management, pest control, and habitats for pollinators. A study conducted in eastern New Mexico, in which animal grazing was integrated into crop production systems, led to improvement in soil health through increases in microbial community size, soil organic matter, and total nitrogen compared to conventional cropland.

5. Add soil organic amendments: Organic amendments are any material of plant or animal origin that can be added to the soil to improve soil conditions and stimulate biodiversity. Examples of amendments include manure, compost, biochar, and similar materials. Long-term research studies have shown that regular addition of organic amendments to the soil can improve soil fertility, soil biological functions, and soil physical characteristics (Diacono and Montemurro, 2011). Adding organic amendments is useful in arid and semi-arid agricultural systems due to the scarcity of vegetation materials that replenish soil organic matter. In most cases, farmers apply organic amendments to supplement crop nutrition on the farm. An additional benefit of adding these amendments is increased microbial activity. Increased microbial activity after soil amendment application breaks down organic materials and releases nutrients.

6. Integrate livestock into cropping systems: Livestock are a big part of agriculture in New Mexico. In crop-range-livestock integrated systems, cattle, goats, and sheep are often used to graze crop residues and stubble during the winter season, while crops are grown from spring through the fall. Animals are moved to the rangeland when cash crops are growing in the crop production fields. Integrating livestock into the cropping system through winter grazing can benefit producers by increasing the efficiency of farm resource utilization and decreasing the need for synthetic fertilizers and pesticides, while improving soil health (Ghimire et al., 2013). Livestock-integrated systems support the growth of diverse microbial communities that promote soil organic matter accumulation and the development of healthy soils. Grazing speeds up the decomposition of leftover plant residues, thus enhancing the availability of nutrients like nitrogen, phosphorus, and sulfur. Hoof actions of animals can bring more residue in contact with the soil, while the saliva of the grazing animals helps the growth of microbial decomposers, leading to better soil organic matter formation and nutrient release. However, high-intensity grazing of limited amounts of residue (overgrazing) could negatively impact soil health due to excessive trampling and hoof-induced compaction of the soil surface. Overgrazing can also lead to aggregate breakdown and reduced water infiltration into the soil. Therefore, an optimal balance of crop-range-livestock is critical for maintaining or improving soil health in livestock-integrated systems.

7. Promote diverse plant species with different rooting depths: Introducing species with diverse rooting depths is essential for the long-term sustainability of rangeland soil health. Having a rich diversity of species has been shown to improve soil health due to increases in soil carbon sequestration and soil biodiversity (Chen et al., 2018). Diverse plants produce various soil enzymes that recycle nutrients and increase soil organic matter accumulation. Plant roots also release various compounds that bind or "glue" the soil particles together to form aggregates, which is important for infiltration, water retention, and erosion control. Introducing plant species with diverse rooting depths will also promote nutrient cycling within the rangeland by providing opportunities for plants with deeper roots to capture nutrients that would otherwise be lost to leaching. Increased plant diversity is also associated with better aggregate formation and stability. One of the major indicators used in rangeland health assessments is plant species loss or biotic integrity, which is defined as "the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region".

8. Sustainable animal grazing practices: Grazing strategies in rangeland systems can affect soil health positively or negatively. Many soil properties can be affected by improper stocking density of animals (number of animals per land unit area) that are grazing a given land. One of the major problems encountered due to overstocked grazing is the compaction of agricultural fields and the breakdown of the surface soil structure (Mulholland and Fullen, 1991). With overgrazing, the plant communities being grazed can be rapidly depleted and damaged, preventing their regrowth and thereby limiting the utility of the land for future grazing.

9. Be patient for the results: Soil health management is a long-term strategy. Just like human health, correcting soil problems cannot happen overnight. It takes consistent and diligent management to address soil health constraints. The following steps are necessary to build an effective long-term soil health management strategy for your land.

a. Education: Land managers need to be well educated on soil health management strategies and how to apply them to their land.

b. Thinking: There is a need to give careful thought to different options that are available to be able to know what will work and what will not work in your farm situation.

c. Planning: After identifying which soil health management strategy you want to use, you need to carefully plan when and how you will implement the practices. For example, if you want to use manure, your planning should include soil and manure testing, figuring out application rates and times, and planning for follow-up soil testing and field monitoring that needs to be done.

d. Discussion and reading: It may be important to speak to someone who has tried the specific soil health practice that you are planning to use. They may be able to share information that will enable you to succeed. You can also discuss options with your local NMSU Cooperative Extension Service agent or specialist to help you make the right decisions. Reading soil health books is a highly effective way to learn more. Some of the books and websites that can help with learning more about soil health are listed after the references section of this publication.

PRINCIPLES OF THE INTEGRATED SOIL MANAGEMENT APPROACHES

Integrated management of soil is based on the principles of the following integrated management approaches:

- Integrated Soil and Nutrient Management (ISNM): 'ISNM aims to optimize the condition of the soil, with regard to its physical, chemical, biological and hydrological properties, for the purpose of enhancing farm productivity, whilst minimizing land degradation' (FAO, 2000, p.5).
- Integrated Soil Fertility Management (ISFM): 'ISFM is defined as a set of soil fertility management practices that include the integrated use of mineral fertilizers, organic inputs

and improved germplasms' (Fairhurst, 2012, p.vii). These management practices are too 'combined with the knowledge on how to adapt these practices to local conditions which are aimed at optimizing efficient agronomic use of the applied nutrients and thereby improving crop productivity' (Fairhurst, 2012, p.vii).

- Integrated Pest Management (IPM): 'Integrated Pest Management (IPM) means the careful consideration of all available pest control techniques and subsequent integration of appropriate measures that discourage the development of pest populations and keep pesticides and other interventions to levels that are economically justified' (FAO & WHO, 2014, p.4). These managements also 'reduce or minimize risks to human and animal health and/or the environment' (FAO & WHO, 2014, p.4).
- Integrated Management of Soil Conservation (IMSC): Integrated Management of Soil Conservation (IMSC) primarily protects soil from all sorts of degradation, improves soil fertility and environmental stability.
- Integrated Farming System (IFS): Integrated Farming System (IFS) is the integration ofcrop, livestock and fish production and sustainable use of farm resources.

CONCLUSION

Soil health assessment and management will continue to play a prominent role in agricultural production systems of arid and semi-arid agroecosystems. Healthy soil is more resilient against fluctuations in growing conditions. With yearly weather uncertainties, the resiliency of the soil system needs to be enhanced to cope with these variations. Building and improving the soil health of New Mexico crop and rangelands will ensure continued productivity, enhance farmers' incomes, and promote food security. Building and maintaining a healthy soil is not an easy task in the arid and semi-arid Southwest, especially in the face of droughts and other production challenges. With the resolution of the farming and ranching community, it is possible to adopt soil health practices that will maintain and improve soil productivity.

Recent agricultural trends indicate that yields for many cereals are not rising as quickly as they did during the 1960s and 1970s. Part of the explanation for such a decline in yield growth is the mismanagement of nutrients and soil fertility. Future strategies will have to redress this poor management in order to create synergies with other yield increasing technologies. Boosting food supplies to meet projected demand by 2020 will require substantial increases in yields in Africa, Latin America, and Asia. The integrated management of nutrients and soil fertility, along with continuous technological change, farmer participation, technology transfer, and a conducive policy environment, are key components for attaining these increases. So long as agriculture remains a soil-based industry, there is no way that required yield increases of the major crops can be attained without ensuring that plants have an adequate and balanced supply of nutrients. The appropriate environment must exist for nutrients to be available to a particular crop in the right form, in the correct absolute and relative amounts, and at the right time for high yields to be realized in the short and long term.

In this regard it is important that governments encourage analysis of “nutrient cycles” to have a better basis for determining the flow of plant nutrients in and out of soils. Governments should establish adequate testing and monitoring systems to gather data on the nutrient cycle and nutrient balances in representative areas throughout their rural economies. Further, governments should support research for developing modern varieties and appropriate integrated nutrient systems for harsh climatic environments, such as those in Sub- Saharan Africa. Research should also be promoted on biological nitrogen-fixation as a low-cost “organic” approach to increasing nitrogen availability and organic matter content in soils. Government and extension services will initially need to stimulate the adoption of nitrogen-fixing species and innoculants by farmers.

UNIT-7: PLANT ECOLOGY: CONCEPT OF ADAPTATION, SUCCESSION AND CLIMAX

The word 'ecology' during mid of 19^{th} century evolved to interpret the fact of dynamic interaction between organisms and their surroundings. Literally, it is the study of organism at home. There are lots of controversies about who coined the term but it is consensus that German biologist Ernest Haeckel used and gave definition of this term. Ecology has four basic concepts:

a) Holism, b) ecosystem, c) succession, and d) conservation.

Plants play an important role as primary producer and provide food for all other living organisms in biosphere. Plant ecology is the study of plant species in their natural environment. It is considered as a sub-discipline of ecology. It studies the distribution and abundance of plants over the surface of the earth, the effects of environmental factors upon the abundance of plants, and the interactions among and between plants and other organisms. One of the early classic books on plant ecology was written by J.E. Weaver and F.E. Clements in 1938. It describes broadly about plant communities and particularly competition forces and processes like succession. Plant ecology is an important branch of geography and it is commonly known as plant geography. Basic contents of plant geography are origin and development, classification of plants, spatial distribution, dispersal, extinction and function. Spatial distribution of plant species has been discussed as plant community which is defined as association of population of different plant species in a given habitat. Plant community consists of more than one plant species. Plants grow in a particular habitat after adopting with regional environment.

There are four types of major habitats for plants. These are: i) terrestrial, ii) freshwater,

iii) estuarine, and iv) marine. Adaptation, succession and climax of plant community depend on habitat environment.

Plant ecology deals with the study of plant ecophysiology, plant population ecology, community ecology, ecosystem ecology, landscape ecology and biosphere ecology.

Carl Ludwig Willdenow noted that similar climates produced similar types of vegetation, even when they were located in different parts of the world. Willdenow's student Alexander von Humboldt described vegetation types and noted that the distribution of vegetation depends on environmental factors. Plant distributions are controlled by a combination of historical factors, ecophysiology and biotic interactions. A species must either have evolved in an area or dispersed through natural processes or through human agency. Physiological adaptation is very much significant for survival of a plant species. Plant communities are broadly distributed into biomes based on the present dominant plant species. Biomes are determined by regional climates (mostly temperature and precipitation) and they have general latitudinal trends.

CONCEPT OF ADAPTATION IN PLANT ECOLOGY

Adaptation of organisms is an important concept in ecology. Adaptation is the fit of organisms to their environment. For organisms it is the process of adopting better solutions of problems created by environment. Adaptation is the process by which an organism enables to exist under existing conditions of a habitat as well as environment. Adaptation process is ultimate outcome of evolution. Through adaptation organisms modified themselves according to alternation of environmental condition and make them as potential climax community. Life forms are very much divers and complex. This diversity is the also nothing but the adaptation of organisms to their respective habitat. Organisms when faced environmental problems and these problems need to solve then they guided by the complex mechanism of evolution through natural selection. Evolution process provides to organisms better solution for adaptation. So adaptation is relative rather than absolute. Slow or moderate change of environment is accompanied by adaptation of organisms but rapid changes are cause of extinction of organisms. Through the dynamic evolutionary process of adaptation, organisms enhance their fitness with slow but steady changing environment.

In general adaptation of organisms with their changing environment can be grouped into:

i) Structural adaptations,

ii) Protective adaptations,

iii) Animal association adaptations,

iv) Physiological adaptations, and

v) Biochemical adaptation.

Warming classified plants on the basis of soils into the following groups concerning theirgrowing (Shukla and Chandel, 1994):

i) Oxylophytes – plants of acidic soil,

ii) Halophytes – plants of saline soil,

iii) Psammophytes – plants on sands,

iv) Lithophytes – plants on surface of rocks, and

v) Chasmophytes – plants in rocks fractures.

Warming (1909) again classified plants on the basis of available water for growing (Shukla and Chandel, 1994). These are:

i) Hydrophytes - plants in or near water.

ii) Xerophytes – plants in very poor supply of available water.

iii) Mesophytes - .plants in an environment neither very dry nor very wet.

HYDROPHYTES

As described by Shukla and Chandel (1994), hydrophytes are plants which grow in wet places or in water. These are called aquatic plants. Plants remain submerged either partly or wholly under water. Hydrophytes are classified as: (a) submerged hydrophytes, (b) floating hydrophytes, and (c) amphibious hydrophytes. The controlling factors behind the growth of hydrophytes in the aquatic environment are: (a) temperature of water, (b) water quality (nature of dissolved chemicals), (c) surrounding environmental factors, etc. Hydrophytic adaptations characterised by morphological adaptations like roots, stem, leaves modification; anatomical modifications, physiological modifications, etc.

XEROPHYTES

As described by Shukla and Chandel (1994), xerophytes are plants which grow in dry environment. These types of plants are found in the desert areas. These plants are characterised by dry conditions, low humidity and high temperature. Xerophytes are grouped into several types based on their drought resisting power. There are like hydrophytes certain morphological adaptations like roots, stem, leaves modification; anatomical modifications, physiological modifications, etc.

MESOPHYTES

As described by Shukla and Chandel (1994), this community is intermediate community between hydrophytes and xerophytes. They cannot grow in waterlogged or in dry areas. These are classified into two main groups. These are: (a) grasses and herbs communities, and (b) woody plants communities. Mesophytic adaptation characterised by tropical rain forest biome, sub-tropical forests, deciduous forests, meadows, etc.

CONCEPT OF SUCCESSION IN PLANT ECOLOGY

The plant community developed in a habitat through series of stages with time. Some typical plant communities try to maintain equilibrium. But, generally one community is replaced by another concerning ecological complexity of an area. Communities are dynamic. They are changing over space and time. Colonisation of species as well as progress of community is a continuous process of ecosystem development. Environment always changes over a period of time and that influences on existing community. It is cause of modification of existing community or replaced by another community. When species composition changes then community structure and function continue to successive development one after another community until the final community reaches again in more or less stable for a specific period.

This relative occurrence of species over a specific area is known as succession or ecological succession. E. P. Odum opined that plant succession is an orderly process of community change in an area under consideration. According to Clements, succession is a natural process by which the same area becomes successively colonised by different groups or communities. Ecological succession is a sequential process of community progress and changes in species structure. Succession is the result of modification of physical environment. It ends in a stabilised ecosystem. Succession is characterised by continuous changes in animals and plants. It is also characterised by increasing diversity of species.

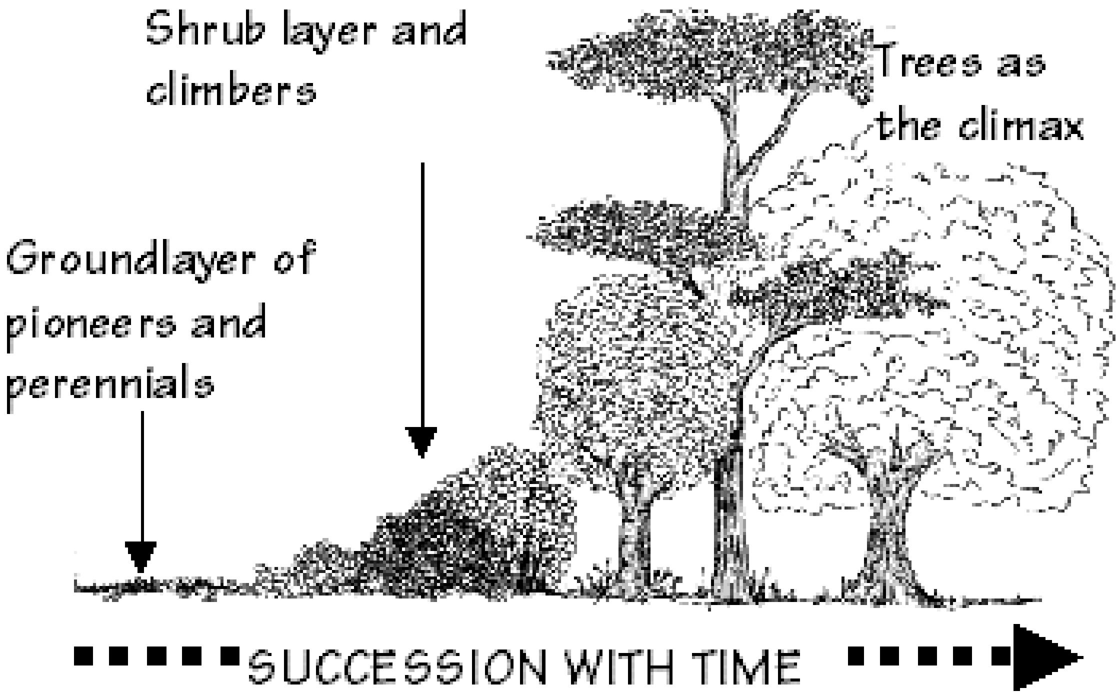
CAUSES OF SUCCESSION

Succession of community is a complex process and the causes are:

a) **Initiating causes**: Climatic causes, topographic causes and biotic causes.

b) **Continuing causes**: Migration, aggregation, competition, reaction etc.

c) **Stabilising causes**: Stabilisation of community.



(Source: http://www.self-willed-land.org.uk)

Figure 1: SUCCESSION OF PLANT WITH TIME

TYPES OF SUCCESSION

On the basis of different aspects basic types of succession are as follow:

d) **Primary succession -** starts from primitive substratum.

e) **Secondary succession -** new succession starts on previously built up substrata.

f) **Autogenic succession -** community modification with changing environmental conditions.

g) **Allogenic succession -** community modification due to external condition.

h) **Autotrophic succession -**characterised by early and continued dominance of autotrophic organisms.

i) **Heterotrophic succession -**characterised by early and continued dominance of heterotrophic organisms.

PROCESSES OF SUCCESSION

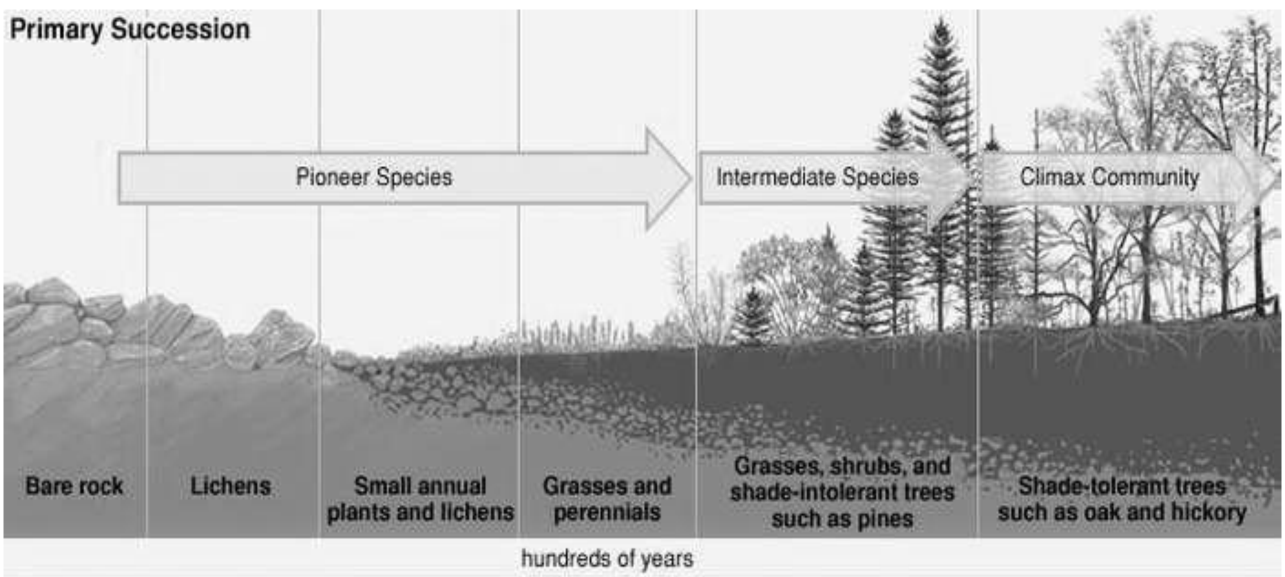
j) **Nudation –** first stage of succession where a bare land developed.

k) **Invation -** species successfully established in a bare land. This process completed through three successive stages: i) dispersal (migration), ii) establishment, and iii) aggregation.

l) **Competition and Coaction -** after invation species competition develops basically for nutrition and space.

m) **Reaction –** modification happened by the influence of living organisms and when modified environment unsuitable for existing community then it will be replaced by other community.

n) **Stabilisation -** the stage of climax of vegetation.



(Source: Encyclopedia Britannica, 2006)

Figure 2: PROCESS OF SUCCESSION AND CLIMAX COMMUNITY

CONCEPT OF CLIMAX IN PLANT ECOLOGY

The end point or final stage of succession is referred to as stabilisation. In this stage, final community becomes more or less stabilised for a longer period of time by process of equilibrium with the environment. This ultimate community is known as climax community and the stage is defined as climax stage. Here populations reach to an equilibrium condition which is stable and self-replicating community with the surrounding environment. A specific type of climax developed largely depending on the climatic condition of an area. Climax vegetation is characterized by dynamic equilibrium stage and self-sustaining. Small scale environmental changes are unable to alter the climax species and species are smoothly adopted with this modification. Sometimes a climax community is seen as more stable in condition for a considerable time period, called arrested climax. The climax concept in ecology has been discussed by different approaches. These are: i) Monoclimax theory, ii) Polyclimax theory, and iii) Climax pattern theory.

UNIT-8: DISPERSAL AND MIGRATION OF ANIMALS: MEANS AND BARRIERS

The geographical (distribution) study of animals is called zoo geography. The distribution of animals over the earth surface is the result of migration of animals over time into ecologically suitable region. In biogeography, there are several fundamental processes by which organisms respond to changes in the geographic template. These are:

(i) Evolution.

(ii) Speciation.

(iii) Extinction.

(iv) Dispersal.

Simply, dispersal is the movement of organisms away from their birthplace. The spreading of animals from the place of their origin to other areas is called animal dispersal. Dispersal is an ecological process that plays an adaptive role in the life history of the organisms which are involved. Dispersal and migration of animals or biological dispersal of animals refers to the movement of individuals or groups from their native habitat to another place due to different reasons. The fitness of the organism is increased in some way through the process of dispersal. In biogeography, it is an interesting topic since dispersal events related with species spatial movement and change their range by dispersing over long distances. Thus, it is the moving of organisms from their habitat to another region and it is also the consequences of individual fitness, population dynamics and genetics and species variation. To understand dispersal and its dynamic characteristics from evolutionary strategies at different level of ecosystem requires understanding on dispersal types and range of a given species and mechanisms involved in it.

The early dispersalists are Darwin, Alfred Russel Wallace and Asa Gray. According to them disjunctions (a situation in which two closely related populations are separated by a wide geographic distance) could be best explained as the result of long distance dispersal. The continental drift theory is considered as a means for dispersal. Some animals are percipient in dispersal throughout their lives and some others are limited phases of their life cycles. The strategies of animal's dispersal are predicted based on habitat nature and circumstances. Organisms can disperse either actively or passively. As for example: Monarch butterflies migrate great distances. They fly from southern Canada to the southern United States and central Mexico. Many organisms depend on other organisms for their long distance migration. Parasites are a good example. Many insects, spiders, and mites disperse through the atmosphere residing upon aerial planktons.

During dispersal as well as range expansion, organisms must be able to reach in a new area. They must survive against the harsh conditions occurring during the passage. They also survive and reproduce in the new area for the establishment of a new population.

TYPES OF DISPERSAL OF ANIMALS

(i) **Gradual dispersal**: It involves longer period of time and very slow rate of migration. It covers larger area and results in widespread distribution of animals over time.

(ii) **Rapid dispersal**: Here organisms are dispersed or migrated rapidly in the new areas from one area due to harsh environmental condition or human discourage.

(iii) **Seasonal dispersal**: Due to seasonal extreme weather conditions every year animals migrated to a favourable area. Here migratory animals again return to their native places.

(iv) **Forced dispersal**: Animals are forced to leave their native place or habitat due to sudden and catastrophic events like prolonged droughts, recurrent floods, forest fires, volcanic eruption, etc.

(v) **Dispersal for breeding place**: Animals migrated from their native place to find sufficient and suitable breeding places to avoid harsh condition.

(vi) **Anthropogenic dispersal**: In this type of dispersal, humans are largely involved in dispersal of animals through their deliberate efforts. It has caused several environmental and ecological problems.

Considering the process of dispersal, dispersal events broadly are grouped as: (i) Jump dispersal or sweepstakes, (ii) Diffusion, and (iii) Secular migration. Jump dispersal refers to the colonization over a long distance in a new area of animals, birds, etc. from their birth place. Bats are often through this process became island inhabitants. Diffusion denotes gradual spread of organisms outward from the boundaries of the range of their distribution. It is a slower form of dispersal. It involves not just individuals, but populations. As for example: Red fox in Australia, American muskrat in Europe, etc. Secular migration happens much more slowly. In fact, it is so slow that organisms can evolve during the process.

MEANS OF DISPERSAL AND MIGRATION OF ANIMALS

The dispersal routes allow movement of organisms from one region to another. It is characterised by allowing a balanced assemblage of plants and animals to cross them without any selective discrimination. At the two ends of a dispersal route as well as corridor, there are fairly similar assemblages of organisms. A 'filter' is a dispersal route that cannot allow all the passing organisms to cross it rather it selects to pass through it. The Arabian subcontinent acts as a filter. It only permits certain mammals, reptiles, etc. to disperse between northern Africa and central Asia. Lastly, 'sweepstakes' routes refer to the crossing of barriers by rare chance events. From south-east Asia different land snails dispersed to the eastward and southward.

Taking all natural and human factors together, in general, the common means of animal dispersal are as follow:

(i) **Corridor bridge**: It is like a corridor for dispersal and migration.

(ii) **Land bridge**: These are barrier to aquatic animals and serve as good passage for land animals and plant population.

(iii) **Filter bridge**: It is like a filter for dispersal and migration.

(iv) **Natural rafts and driftwood**: Animals take long journey upon the natural rafts of vegetation in marine waters. Driftwood acts as means for migration of animals in oceanic islands from main lands. Majority of animals are exterminated during such journey but a few may survive to reach other island.

(v) **Sweepstakes**: It is like a sweepstakes for dispersal and migration.

(vi) **Winds and storms**: Transportation by winds and storms is one of the important means of dispersal of animals and plants.

(vii) **Transportation by animals capable of flight**: Migratory birds transported small animals, insects, plant seeds, plants, etc. long distance to oceanic islands and others.

(viii) **Human agency**: Involving transportation and introducing a number of animals to places where these were not found.

BARRIERS OF ANIMALS

Barrier in dispersal is the dispersal range of a species. Barriers are that type of environment or features in nature which are hostile to the organisms those are under long distance dispersal. The effectiveness barriers to prevent dispersal depend on the nature of the barriers and also on the organisms dispersing. Mostly, barriers are species-specific phenomena. There are different types of barriers. These are:

(i) Physical barriers.

(ii) Climatic barriers.

(iii) Physiological barrier.

(iv) Seasonal barrier.

(v) Sedentary habit.

(vi) Home range or territoriality.

Physical barriers of dispersal are (a) topographic barriers: mountain ranges, deserts, etc.;

(b) large bodies of water and land masses: oceans, continents, etc.; (c) vegetative barrier: forests;

(d) lack of salinity in sea water. Here, the barriers in migration are physical factors which prevent organisms for crossing to a new region. Climatic barriers like temperature, moisture, amount of light, etc. All animals are not able to cross all types of barriers. Therefore, the dispersal and migration of animals also related with their psychology, body size, locomotive organs, physiological endurance, etc. These are physiological barriers. Physiological barriers are accomplished by environmental conditions to survive long enough for dispersal. Such barriers are salt (or fresh) water or unfavorable temperatures. In some areas, the nature of barriers may change with season. In temperate regions of North America, water bodies serve as barriers to the movement of many terrestrial species in summer, but during winter these waters may freeze and allow movement across them. Sedentary habit of animals makes them to aggregate in particular area only. Home range or territoriality involved with choice of animals for not migration to other areas from their native land.

UNIT-9: ECOLOGICAL FOOTPRINT

CONCEPT OF ECOLOGICAL FOOTPRINT

The ecological footprint measures human demand on nature, i.e., the quantity of nature it takes to support people or an economy. It tracks this demand through an ecological accounting system. The accounts contrast the biologically productive area people use for their consumption to the biologically productive area available within a region or the world (bio capacity - the productive area that can regenerate what people demand from nature). In short, it is a measure of human impact on Earth's ecosystem and reveals the dependence of the human economy on natural capital.

The ecological footprint is defined as the biologically productive area needed to provide for everything people use: fruits and vegetables, fish, wood, fibres, absorption of carbon dioxide from fossil fuel use, and space for buildings and roads.

Ecological footprint analysis is widely used around the Earth in support of sustainability assessments. It can be used to measure and manage the use of resources throughout the economy and explore the sustainability of individual lifestyles, goods and services, organizations, industry sectors, neighbourhoods, cities, regions and nations.

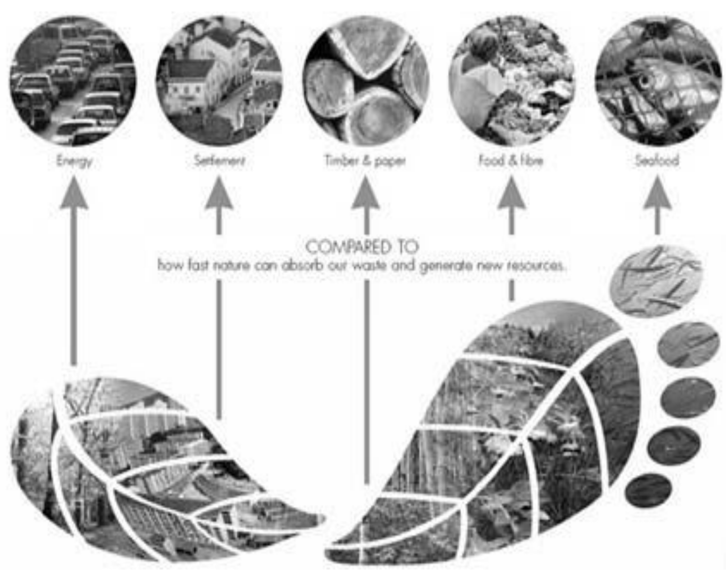


Figure 15 : Ecological Footprint

Source: Accessed from theretailineteligence.com

OVERVIEW

The first academic publication about the ecological footprint was by William Rees in 1992. The ecological footprint concept and calculation method was developed as the PhD dissertation of Mathis Wackernagel, under Rees' supervision at theUniversity of British Columbia in Vancouver, Canada, from 1990–1994. Originally, Wackernagel and Rees called the concept "appropriated carrying capacity". To make the idea more accessible, Rees came up with the term "ecological footprint," inspired by a computer technician who praised his new computer's "small footprint on the desk." In early 1996, Wackernagel and Rees published the book *Our Ecological Footprint: Reducing Human Impact on the Earth.*

Ecological footprint analysis compares human demand on nature with the biosphere's ability to regenerate resources and provide services. It does this by assessing the biologically productive land and marine area required to produce the resources a population consumes and absorb thecorresponding waste, using prevailing technology. Footprint values at the end of a survey are categorized for Carbon, Food, Housing, and Goods and Services as well as the total footprint number of Earths needed to sustain the world's population at that level of consumption. This approach can also be applied to an activity such as the manufacturing of a product or driving of a car. This resource accounting is similar to life cycle analysis wherein the consumption of energy, biomass (food, fiber), building material, water and other resources are converted into a normalized measure of land area called 'global hectares' (gha). Per capita ecological footprint (EF) is a means of comparing consumption and lifestyles, and checking this against nature's ability to provide for this consumption. The tool can inform policy by examining to what extent a nation uses more (or less) than is available within its territory or to what extent the nation's lifestyle would be replicable worldwide. The footprint can also be a useful tool to educate people about carrying capacity and over-consumption, with the aim of altering personal behavior. Ecological footprints may be used to argue that many current lifestyles are not sustainable. Such a global comparison also clearly shows the inequalities of resource use on this planet at the beginning of the twenty-first century.

In 2006, the average biologically productive area per person worldwide was approximately 1.8 global hectares (gha) per capita. The U.S. footprint per capita was 9.0 gha, and that of Switzerland was 5.6 gha per person, while China's was 1.8 gha per person. The WWF claims that the human footprint has exceeded the biocapacity (the available supply of natural resources) of the planet by 20%. Wackernagel and Rees originally estimated that the available biological capacity for the 6 billion people on Earth at that time was about 1.3 hectares per person, which is smaller than the 1.8 global hectares published for 2006, because the initial studies neither used global hectares nor included bioproductive marine areas.

Ecological footprint analysis is now widely used around the globe as an indicator of environmental sustainability. It can be used to measure and manage the use of resources throughout the economy. It can be used to explore the sustainability of individual lifestyles, goods and services, organizations, industry sectors, neighborhoods, cities, regions and nations. Since 2006, a first set of ecological footprint standards exist that detail both communication and calculation procedures. They are available at www.footprintstandards.org and were developed in a public process facilitated by Global Footprint Network and its partner organizations.

IMPLICATIONS

There are several advantages and limitations associated with the development of the ecological footprint concept. The major advantage of the ecological footprint concept over some other indicators like environmental space is that the former concept gives a clear, unambiguous message often in an easily digested form. The clarity of the message is an important function of any indicator for both policy makers and the general public. Next, the calculation upon which the ecological footprint is based is relatively easy to undertake and much of the data is available at different spatial scales. Third, more detailed calculations do include trade within the ecological footprint. If world trade were included then, under the assumption of all areas maintaining their inhabitants' standards of living, there would be some losers as well as winners. A glance at the Human Development Index gives some empirical support of the increasing numbers of poor within the Third World as well as pockets of poor and a growing underclass in rich Western democracies. Fourth, the measure is simply stated as a stock, for example, x units of land per capita. It is obvious that each areal unit can also supply a flow of goods, information, natural and manmade capital as well as pollution into and out of the region.

UNIT-10: BIODIVERSITY: ISSUES AND CHALLENGES

When we look at the geological history of evolution of both flora and fauna it is found that a gradual change of them occurred. With the emergence of different species some species extinct in parallel due to different causes. In this regard at present human activity is very important.

The term ‘biodiversity’ is a contraction of ‘biological diversity’. In 1985, Waler G. Rosen coined the term ‘Biodiversity’. But, long before, wildlife scientist Raymond F. Dasmann used ‘Biological Diversity’ in the year of 1968. Biodiversity denotes the variety of all life forms on earth and the essential interdependence of all living things. In the convention on Biological Diversity at Rio De Jenerio (Brazil) in 1992 by 154 countries, the term ‘biodiversity’ was defined as ‘the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic eco-systems and the ecological complexes of which the area part, this include diversity within species, between species and of ecosystem.’ According to IUCN (1998), biodiversity refers to ‘the variety and variability of species of their population, the variety of species of their life forms, the diversity of the complex association with species with their interaction and their ecological process which influences perform.’ Biodiversity has intrinsic value. Biodiversity has also utilitarian value. Biodiversity is dynamic.

TYPES OF BIODIVERSITY

Biodiversity is observed in terms of genes, species and ecosystem. There are three types of biodiversity. These are: (a) Genetic diversity, (b) Species diversity, and (c) Ecosystem diversity. Genetic diversity represents the heritable variation within and between populations of organisms. Species diversity refers to different kinds of organisms, relationships among species. It represents different types of plants, animals and other types of life forms within a region. ‘Species richness’ means the number of species in a site or habitat. Species diversity is an indicator to understand the biological richness/ stress of an ecosystem. Biodiversity is determined by species richness, species evenness, species composition, species interaction, etc. Ecosystem diversity represents the variety of different habitats/ ecosystems in a particular area (e.g. wetland, grassland, forest, etc.). The world ecosystems are maintained by their biodiversity.

SIGNIFICANCE AND BENEFITS OF BIODIVERSITY

Biodiversity has wider significance and value in all level to support the life on earth.

(i) Biodiversity is the sum of all life forms on earth and it includes genetic, species, and ecosystem diversity.

(ii) Biodiversity shows the health of the ecosystems on earth.

(iii) Biodiversity helps human life through oxygen, food, clean water, medicine, aesthetics, ideas, etc.

(iv) Biodiversity regulates climate, biogeochemical cycles and hydrological functions.

(v) Biodiversity helps soil protection.

(vi) Biodiversity helps crop pollination.

(vii) Biodiversity supports recreation and ecotourism.

(viii) Biodiversity supports ecological services.

(ix) Biodiversity gives traditionally free benefits to society or public goods.

(x) Biodiversity supports aesthetic and cultural benefits.

(xi) Biodiversity supports various local to global benefits involving commodity value, amenity value, ecological integrity value, ethical value and other values.

FACTORS INFLUENCING BIODIVERSITY LOSS

Biodiversity loss includes the worldwide extinction of different species, as well as the local reduction or loss of species in a certain habitat, resulting in a loss of biological diversity. The latter phenomenon can be temporary or permanent, depending on whether the environmental degradation that leads to the loss is reversible through ecological restoration/ecological resilience or effectively permanent (e.g. through land loss). The current global extinction (frequently called the sixth mass extinction or Anthropogenic extinction), has resulted in a biodiversity crisis being driven by human activities which push beyond the planetary boundaries and so far has proven irreversible

There are different factors responsible for variety of species over the earth surface. These are: (i) Spatial heterogeneity factor, (ii) Climatic factor, (iii) Time factor, (iv) Competition factor, (v) Predation factor, (vi) Productivity factor, and (vii) Environmental stability factor.

PRINCIPAL CAUSES FOR THE LOSS OF VARIETY OF LIFE ON EARTH

The numerous factors are responsible for the loss of Biodiversity such as pollution, habitat loss, and hunting, introduction of invasive species, overexploitation of preferred species, climate change, and natural disasters.

POLLUTION

Living organisms have developed over an extended period, however they strive to adapt to existence on a polluted planet. Air pollutants such as soot, dust, ammonia, or carbon dioxide can directly and indirectly influence biodiversity.

AIR POLLUTION

Air pollutions influence the respiratory apparatus of the animals and negatively impact their well-being including the egg laying capability and behavioral alterations. Air pollution is also known to influence the reproductive ability of the animals and hence success in breeding. The indirect influence of air pollution on animals is difficult to evaluate as it is difficult to examine in controlled environment. The tremendous discharge of greenhouse gases such as nitrous oxide, carbon dioxide, methane is swiftly modifying the climate of the earth. The animals and plants find it difficult to adjust and this influence the biodiversity. Acid rain is other pollutant, and it causes enhanced build-up of the mucus in fish gills leading to suffocation. Soil prone to acid rain has diminished activity of microbes. This influences the food chain and other forms of life too.

WATER POLLUTION

Water pollution had detrimental effect on biodiversity. Chemical fertilizers generally contain nitrogen and phosphorous and are added to soil to boost the crop productivity. The presence of these nutrients in the water bodies lead to eutrophication or excessive plant growth. Eutrophication causes the depletion in the oxygen level which is deleterious for biodiversity. Fish and other aquatic animals die because of lack of dissolved oxygen in water. Alike fertilizers, pesticides may also accumulate in water bodies. The pesticides negatively affect non-fl owing water bodies such as lakes and ponds given the fact that fertilizers are not washed away and animals in water bodies have difficulty in reproducing.

Various anthropogenic activities such as production of cement, cars; mining etc. lead to the introduction of heavy metals such as arsenic, cadmium, mercury into the water bodies. Heavy metals affect the behavior as well as the survival rates of aquatic animals specially fish. Further, events such as oil spills greatly impact the wildlife specially in the deeper oceans. The birds and the larger animals display the apparent hostile effects. Oil spills cause disruption of the animal senses, suffocation, impair the vital organs of the organisms, reduction in growth rates and induce the higher mortality of the larvae. Like the oil spills, plastic stays in the environment for longer time-period and hence influence the wildlife. Lager animals particularly turtles fail to regurgitate causing internal injuries leading to death. It has been noticed that the seabirds like Layson albatross upon consumption of plastic die prior to fl edging the nest. Micro plastics in the environment also impact the survival rate of larvae, diminished food consumption and gradually weight loss in aquatic animals.

SOIL POLLUTION

Soil pollution is another factor adversely affecting biodiversity. Soil contaminated with heavy metals greatly impacts the welfare of the microorganisms essential for the sustaining life of the living organisms. The excess of heavy metals present in the soil are not easily broken down and are accumulated by plants. The over-use of fertilizers, pesticides and antibiotics used in agriculture is also very deleterious for the biodiversity. These agricultural pollutants such as nitrogen from fertilizers alter the pH and the nutrient level of the soil. The enhanced presence of nutrients in the soil causes the vigorous growth of grass species, leading to stifle in the growth of wildflowers, essential for bees and other pollinating insects. This greatly impacts the biodiversity.

Light pollution such as the use artificial light too has negative impact on the biodiversity. The late-night streetlights, lights from the buildings, vehicles headlights etc. have specially affected the nocturnal animals such as bats. The feeding activity of bats has decreased and their emergences from roosts have been found to be delayed in artificial lights. The behavior of the moths is also largely affected by the light pollution. Moths are prey for other species and pollinators of many species.

Noise pollution has also been reported to have negative impact on biodiversity. The studies have reported that birds in noisier areas begin laying less egg as it masks the important territorial calls birds make. The machinery noise in Brazil at a mining site was found to impact wildlife. Species numbers declined at sites nearer the mine and were higher farther away.

HABITAT LOSS

The hefty menace to biodiversity and the vital discernment for species extinction is habitat loss. This is a grave issue both for wildlife as well as humans. Habitat and wildlife are causally related. For the active working of the ecosystem to reap the benefits, for which we confide in for our existence, we cannot risk neglecting the wildlife. The destruction, degradation and fragmentation of habitat are the three predominant categories of habitat loss.

The habitat destruction is the massive destruction of the natural habitat of the species that it becomes incapable of upholding the native ecosystems and the species. This ultimately results in species extinction i.e., biodiversity loss. The cutting the forests for preparing the fields for agricultural use, filling the wetlands and mowing fields for creating residential or commercial sites, harvest of the fossil fuels, etc. are all examples of habitat destruction.

Development of agricultural practices, reduced resources such as food, water, air quality, mining, pollution, logging, catastrophic fishing activities, activities related to urbanization and the interruption of processes related to ecosystem are the predominant elements of degradation of habitat. The habitat degradation affects both the species dwelling in the habitat and the humans both. The erosion, depletion of nutrients and desertification cause the further loss of the degraded land.

Habitat fragmentation is another gigantic issue arisen due to human developmental activities. Human beings for the purpose of development and to meet the never ending needs to expand even at the cost of extinction of their own species converts large wild areas into smaller fragments of land. These split up areas break down the habitats of the animal and plant species, isolate animal communities, compressing genetic diversity.

HUNTING

Hunting is the root cause of extinction of large numbers of animals holding position in food web. Due to this, the various species in the region are adversely affected as they face food scarcity or complete food unavailability compared with the normal situation. Hunting is extensive operator of loss of biodiversity. Hunting activities exert a considerable burden on wildlife, provoking immense downturn of wildlife and leading to disturbed and inefficient ecosystems. In most extrude situations; overhunting can lead to the elimination of large mammals in contrarily healthful unharmed habitat, compelling transition in forest structure.

OVEREXPLOITATION OF PREFERRED SPECIES

Overexploitation is the collection of species from their natural habitat at higher estimates than the native communities can retrieve. Presently, nearly one third of the earth’s vertebrates facing extinction are vulnerable due to overexploitation. Extreme fishing and hunting are the examples of overexploitation. Passenger pigeons which were once very commonly seen are the sufferers of overhunting. Similarly, various creatures both plants and animals are collected to be used as pets, trophies, or souvenirs. Such collections are illegal and usually called as poaching.

Overexploitations of plants which have medicinal values have eventuated in their loss from native habitats. Many plants such as *Drosera sp.*, *Gnetum sp.*, pitcher plants, *Psilotum sp.*, *Napenthes*

Khasiana etc. are relentlessly hunted and collected for academic tutoring and laboratory experiments. Similarly, various orchids have also been overexploited.

CLIMATE CHANGE

The biodiversity and climate change are strongly associated. Even though the climate has consistently altered during the whole of earth's history with ecological communities and species evolving and extinguishing, accelerated climate change disturbs ecological systems and species capability to acclimate and hence the loss in biodiversity enhances. The swift climate change, stimulating biodiversity loss jeopardize human interests and security for clean water, air, medicines, and additional natural resources we depend on, would be difficult to attain due to reduced or vanished flora and fauna they are obtained from. Climate change, beside other components such as habitat loss, land degradation, hunting, overexploitation of certain species etc., is turning up as a high threat to biodiversity on earth.

INVASIVE SPECIES

The introduction of invasive species is the tremendous threat to biodiversity crisis. The species, which is not native to the ecosystem, arrives or is introduced mostly via humans in the new ecosystem and start to pullulate, is called as invasive species. Such species are detrimental as they affect the ecosystem disproportionately compared to any other species. Most of the new species introduced in the ecosystem do not become invasive, but few of them turn into invasive species and adversely affect the ecosystem. The invasive species rack up the native ecosystem in many ways such as they modify the habitat, import pathogens, are herbivorous on plants in native ecosystem, lead to decline of genetic diversity by hybridizing with natives, for the resources they directly compete with and prey on the native species. The prevention of biodiversity loss is a huge developmental issue as the climate change and hence demands further ardent commitment and response from the communities.

NATURAL DISASTERS

Natural catastrophes, for instance volcanoes, wildfires, floods, hurricanes, draughts, epidemics, tsunamis etc. cause a heavy loss of biodiversity. In humid tropical areas such as central Africa, eastern and northern Australia, some areas of South America floods are common. The tropical areas harbor lot of vegetation and vast number of animals survives in the vegetation. Due to flooding, large amount of nutrients from the soil gets washed away. Drought too led to dry soil and decline in the level of water table. In this situation, both animals as well as plants suffer.

Similarly, wildfires in the thickly wooded forests and earthquakes significantly disrupt the life of the organisms and thus affecting biodiversity. Volcanoes frequently wreck animals and plants in the adjoining areas. Epidemics occasionally wipe out vast majority of the population. The occurrence of epidemics in nature is normally restricted to certain population of animal or plant since the pathogen is usually species specific.

BIODIVERSITY ISSUES AND MANAGEMENT

Biodiversity i.e. flora and fauna distributions depend on climatic elements, physiographic characteristics, availability of water, soil characteristics, presence of other organisms, human behavior and interaction, etc. In the tropical region over the earth surface maximum biodiversity is concentrated. Current worldwide landuse practices as well as activities related to economic development of countries lead to the degradation of biodiversity. Different reports, websites, papers, etc. show that by the end of the 21st century, two-thirds of the species on earth may extinct. Till date, about 50 percent of the natural habitats on earth have been cleared due to human uses like agricultural field generation, urbanisation, industrialisation, infrastructural development, etc., and another 0.5 to 1.5 percent of nature is lost each year. Nearly 75 percent of the genetic diversity of crops has been lost in the past century. Direct and indirect human activities have altered the terrestrial, freshwater and marine ecosystems throughout history. According to the principle, the rate of extinction should be one mammal and two bird species in every 400 years. But, recorded extinction of past nearly 400 years shows that 58 mammals and 115 bird species have disappeared. According to an estimation made by scientists that as many as three species per hour are going extinct and 20,000 extinctions occur each year. Extinction may be caused by natural reasons (e.g. droughts, fires, floods, etc.). Rich biodiversity regions are now destroying and the major causes of diversity loss are:

a) Catastrophic processes.

b) Habitat destruction and reduction.

c) Huge hunting of wildlife.

d) International trade of biotic resources.

e) Introduction of new species.

f) Pollution of habitats.

g) Other causes.

According to the Millennium Ecosystem Assessment, there five main causes of biodiversity loss are: (i) Habitat change, (ii) Climate change, (iii) Invasive alien species, (iv) Over exploitation, and (v) Pollution.

Habitat loss is the largest current threat to biodiversity. Based on relative threat of species extinction, IUCN categorised species into six threatened classes and incorporated in Red Data Book. This categorisation was made to restrict the extinction of threatened species. To protect the biodiversity, different strategies and measurements have been taken. Biodiversity has immense value to maintain the human and ecosystem balance. Considering this fact numerous conservation efforts were made to protect and restoring the loss biodiversity. These efforts at global level are: (i) on-site protection management system (in-situ), and (ii) off-site protection management system (off-situ). In-situ conservation, the conservation of species in their natural habitats, includes biosphere reserves, national parks, wildlife sanctuaries, biodiversity hotspots, etc. Ex-situ conservation is the preservation of components of biodiversity outside their natural habitats (e.g. gene bank, microbial culture collection, collection of living organisms in zoo, botanical garden, etc.). There are also biodiversity conservation programme at country level based on national priority task.

UNIT-11: INTERNATIONAL BIOLOGICAL PROGRAMME; MAN AND BIOSPHERE PROGRAMME

Human has immense impact upon nature. Human exploited natural resources unlimitedly. The International Biological Programme or IBP, an international level study on ecological and environmental issues, happened during 1964 to 1974. It was an attempt to promote and apply the methods of science to ecosystem and urgent environmental issues. It is a worldwide plan of research concerned with the biological basis of productivity and human welfare. Though, there was a continuous process of starting of this type of study from 1959 but finally it has been started in 1964. This idea was stimulated by the success of International Geophysical Year (IGY), 1957-1958. This programme was concluded in 1974. The main effort of IBP was to coordinate large scale ecosystem ecological and environmental studies. This study aimed to observe the impact of natural environmental changes on biological communities. The activities of man are bringing undesirable degradation of the environment. Its aim is to study the human impact and to fill the gaps in knowledge by means of sufficient research. The successful main result of this programme was studies of five biomes involving grassland biome, deciduous forest biome, etc. This programme enriched as well as intensified the study of ecosystem ecology. The IBP's research structure contained seven programme areas. These are as follow (http://www.nasonline.org):

1. Productivity of terrestrial communities.
2. Production processes.
3. Conservation of terrestrial communities.
4. Productivity of freshwater communities.
5. Productivity of marine communities.
6. Human adaptability.
7. Use and management of biological resources.

In criticism, many biologists, particularly molecular biologists and evolutionary ecologists, remarked that this programme was dramatic concerning its scale of funding than academic achievements. This programme also has no clear socially and scientifically pressing goal.

MAN AND BIOSPHERE PROGRAMME

Man and the Biosphere Programme (MAB) is an Intergovernmental Scientific Programme of UNESCO. It has been introduced in the year of 1971. This programme is to promote interdisciplinary approaches to research and education as well as management in ecosystem conservation and sustainable use of natural resources. MAB programme combines science and arts to improve human livelihood and to safe natural ecosystems. The fundamental aim of this type of scientific agenda is to improve human and environment relationship at global scale. The MAB programme also aimed to ecosystem specific networking development at regional, national and global level. In 1977, the World Network of Biosphere Reserves was introduced by the MAB. Towards implementation of the research outcomes and decisions taken by MAB on ground, MAB in maximum relies on this World Network of Biosphere Reserves. This World Network shared knowledge, experiences, innovations, etc. for sustainable development regional to global scale. The MAB programme addresses several challenging environmental problems in different sectors. It is aimed to improve quality life, equal sharing of resources and to manage ecosystem. This programme promotes the scientific approaches for economic development of a region and society without hampering environmental sustainability. It presents innovative approaches to economic development which is socially and culturally appropriate and environmentally sustainable. The MAB programme is a unique platform to share knowledge and research from different disciplines on human interference, sustainable development, environmental degradation, biodiversity loss, climate change, etc. There are several regional and sub-regional networks of MAB like AfriMAB in Africa, Southeast Asian Biosphere reserve Network, etc.

Framework of MAB programme within the World Network of Biosphere Reserves strives to (www.unesco.org):

i) Identification and assessment of the environmental changes due to human and natural activities and inverse impacts of these changes.

ii) Study and interpretation of dynamic interrelationship between ecosystem and socio-economic processes.

iii) To ensure the basic human welfare with a habitable environment.

iv) To promote the transfer of knowledge on environmental problems for sustainable development.

The main governing body of MAB programme is the International Coordinating Council of the Man and the Biosphere Programme, known as MAB Council or ICC. Role of MAB Council is:

i) To guide and supervise the MAB Programme.

ii) To review the progress of the MAB Programme.

iii) To recommend research projects to countries.

iv) To make proposals on regional or international cooperation.

v) To assess priorities among projects and MAB activities in general.

vi) To coordinate the Member States participating in the MAB Programme.

vii) To coordinate activities with other international scientific programmes.

viii) To consult with international nongovernmental organisations.

UNIT-12: WILDLIFE CONSERVATION AND MANAGEMENT: SANCTUARIES, NATIONAL PARKS AND BIOSPHERE RESERVES WITH REFERENCE TO INDIA

WILDLIFE SANCTUARY

Sanctuary is an area which is of adequate ecological, faunal, floral, geomorphological, natural or zoological significance. The Sanctuary is declared for the purpose of protecting, propagating or developing wildlife or its environment. Certain rights of people living inside the Sanctuary could be permitted. Further, during the settlement of claims, before finally notifying the Sanctuary, the Collector may, in consultation with the Chief Wildlife Warden, allow the continuation of any right of any person in or over any land within the limits of the Sanctuary.

Characteristics of Wildlife Sanctuary

1. It is natural area which is reserve by a governmental or private agency for the protection of particular species.
2. Area is designated for the protection of wild animals.
3. Only animals are conserved.0
4. It came under the category called "Protected Areas". The Protected Areas are declared under Wildlife (Protection) Act, 1972.
5. International Union for Conservation of Nature (IUCN) has defined its Category IV type of protected areas.

Example – Sajnekhali Bird Sanctuary; Manas Wildlife Santuary

WILDLIFE SANCTUARIES IN WEST BENGAL

I. Chapramari Wildlife Sanctuary, 1976
II. Haliday Island Wildlife Sanctuary, 1976
III. Lothian Island Wildlife Sanctuary, 1976
IV. Mahananda Wildlife Sanctuary, 1976
V. Sajnakhali Wildlife Sanctuary, 1976
VI. Senchal Wildlife Sanctuary, 1976
VII. Ballabhpur Wildlife Sanctuary, 1977
VIII. Bethuadahari Wildlife Sanctuary, 1980
IX. Bibhutibhushan Wildlife Sanctuary, 1980
X. Ramnabagan Wildlife Sanctuary, 1981
XI. Chintamoni Kar Bird Sanctuary, 1982
XII. Jore Pokhri Wildlife Sanctuary, 1985

XIII. Raiganj Wildlife Sanctuary, 1985

XIV. Buxa Tiger Reserve, 1986

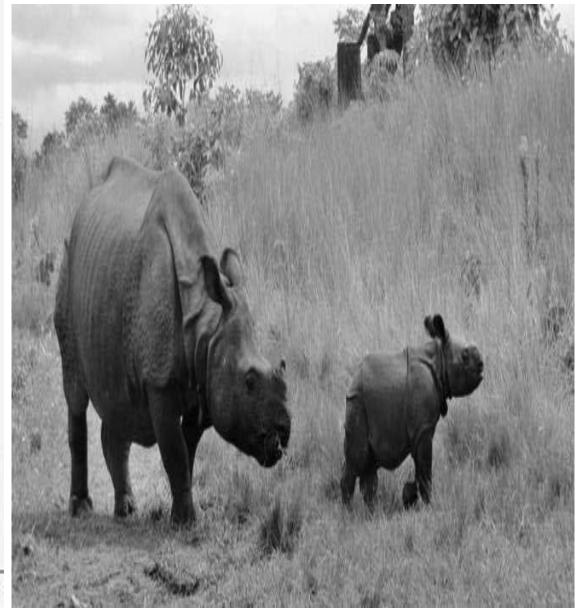
XV. Sundarbans West Wildlife Sanctuary, 2013

XVI. Sundarbans East Wildlife Sanctuary

A SHORT DESCRIPTION OF MANAS WILDLIFE SANCTUARY BRIEF SYNTHESIS

Manas Wildlife Sanctuary is located in the State of Assam in North-East India, a biodiversity hotspot. Covering an area of 39,100 hectares, it spans the Manas River and is bounded to the north by the forests of Bhutan. The Manas Wildlife Sanctuary is part of the core zone of the 283,700 hectares Manas Tiger Reserve, and lies alongside the shifting river channels of the Manas River. The site's scenic beauty includes a range of forested hills, alluvial grasslands and tropical evergreen forests. The site provides critical and viable habitats for rare and endangered species, including tiger, greater onehorned rhino, swamp deer, pygmy hog and Bengal florican. Manas has exceptional importance within the Indian sub-continent's protected areas, as one of the most significant remaining natural areas in the region, where sizeable populations of a large number of threatened species continue to survive.

Manas is recognized not only for its rich biodiversity but also for its spectacular scenery and natural landscape. Manas is located at the foothills of the Eastern Himalayas. The northern boundary of the park is contiguous to the international border of Bhutan manifested by the imposing Bhutan hills. It spans on either side of the majestic Manas River flanked in the east and the west by reserved forests. The tumultuous river swirling down the rugged mountains in the backdrop of forested hills coupled with the serenity of the alluvial grasslands and tropical evergreen forests offers a unique wilderness experience.



The Manas-Beki system is the major river system flowing through the property and joining the Brahmaputra River further downstream. These and other rivers carry an enormous amount of silt and rock debris from the foothills resulting from the heavy rainfall, fragile nature of the rock and steep gradients of the catchments. This leads to the formation of alluvial terraces, comprising deep layers of deposited rock and detritus overlain by sandy loam and a layer of humus represented by *Bhabar* tracts in the north. The *Terai* tract in the south consists of fine alluvial deposits with underlying pans where the water table lies near to the surface. The area contained by the Manas-Beki system gets inundated during the monsoons but flooding does not last long due to the sloping relief. The monsoon and river system form four principal geological habitats: Bhabar savannah, Terai tract, marshlands and riverine tracts.

The dynamic ecosystem processes support broadly three types of vegetation: semi-evergreen forests, mixed moist and dry deciduous forests and alluvial grasslands. The dry deciduous forests represent an early stage in succession that is constantly renewed by floods and is replaced by moist decidous forests away from water courses, which in turn are replaced by semi evergreen climax forests. The vegetation of Manas has tremendous regenerating and self-sustaining capabilities due to its high fertility and response to natural grazing by herbivorous animals.

The Manas Wildlife Sanctuary provides habitat for 22 of India’s most threatened species of mammals. In total, there are nearly 60 mammal species, 42 reptile species, 7 amphibians and 500 species of birds, of which 26 are globally threatened. Noteworthy among these are the elephant, tiger, greater one-horned rhino, clouded leopard, sloth bear, and other species. The wild buffalo population is probably the only pure strain of this species still found in India. It also harbours endemic species like pygmy hog, hispid hare and golden langur as well as the endangered Bengal florican. The range of habitats and vegetation also accounts for high plant diversity that includes 89 tree species, 49 shrubs, 37 undershrubs, 172 herbs and 36 climbers. Fifteen species of orchids, 18 species of fern and 43 species of grasses that provide vital forage to a range of ungulate species also occur here.

INTEGRITY

The property is a wildlife sanctuary with a focus on maintaining the integrity of the property as a natural area. It forms the core of a larger national park, the boundaries of which are clearly demarcated and supervised. Manas Wildlife Sanctuary is buffered on the north by the Royal Manas National Park of Bhutan and on the east and west less effectively by the Manas Tiger Reserve. Transboundary cooperation is therefore important to the effectiveness of its protection.

PROTECTION AND MANAGEMENT REQUIREMENTS

The property, which has six national and international designations (i.e. World Heritage Site, National Park, Tiger Reserve (core), Biosphere Reserve (national), Elephant Reserve (core) and Important Bird Area) has the highest legal protection and strong legislative framework under the provisions of Indian Wildlife (Protection) Act, 1972 and Indian Forest Act, 1927/Assam Forest Regulation 1891. The property benefits from government support at both national and regional levels as well as involvement of national and international conservation organizations.

The property is managed under the administration of the Assam Forest Department / Bodoland Territorial Council. A comprehensive and approved Management Plan is an essential requirement, together with effective patrolling and enforcement capacity to deal with the threats of encroachment, grazing and poaching. The provision of adequate infrastructure, skilled personnel and monitoring arrangements for the property are all essential requirements. Scientific research and monitoring for habitat and invasive species management and recovery of wildlife populations is a particular imperative for management to ascertain and maintain the Outstanding Universal Value of the property. The property is home to 400 varieties of wild rice, also making the management of its biodiversity values of high importance to food security.

Provision of effective tourism facilities, visitor information and interpretation is also a priority for the park management. A sustainable financing mechanism needs to be ensured to provide the necessary financial resources for the long term management of the property. The surrounding buffer zones are managed on a multiple use basis, and a balance is required between conservation and resource extraction in the management of these areas.

Involvement of local communities who live and make use of the areas adjacent to the reserve in protection efforts for the property is essential, and a key management objective is to enhance their engagement and awareness in the interest of the preservation of the property. There is potential to extend the property to coincide with the boundaries of the national park of which it forms the core. The establishment of a trans-boundary world heritage property across the Indian and Bhutanese Manas Tiger Conservation Landscape would enable greater coordination and cooperation in the management of habitat and wildlife populations and would strengthen protection as well.

NATIONAL PARK

National park is an area which is strictly reserved for the betterment of the wildlife & biodiversity, and where activities like developmental, forestry, poaching, hunting and grazing on cultivation are not permitted. In these parks, even private ownership rights are not allowed. Their boundaries are well marked and circumscribed. They are usually small reserves spreading in an area of 100 sq. km. to 500 sq. km. In national parks, the emphasis is on the preservation of a single floral or faunal species.

India's first national park was established in 1936 as Hailey National Park, now known as Jim Corbett National Park, Uttarakhand. By 1970, India only had five national parks. In 1972, India enacted the Wildlife Protection Act and Project Tiger to safeguard the habitats of conservation reliant species.

Further federal legislation strengthening protection for wildlife was introduced in the 1980s. As of July 2018, there were 110 national parks encompassing an area of 40,501 km2 (15,638 sq mi), under protected areas of India category II comprising 1.23% of India's total surface area.

Characteristics of National Park

1. Reserve area of land, owned by the government.
2. Area is protected from human exploitation, industrialization and pollution.
3. No cutting, Grazing allowed, Outside Species is allowed.
4. It came under the category called “Protected Areas”. The Protected Areas are declared under Wildlife (Protection) Act, 1972.
5. Conservation of 'wild nature' for posterity and as a symbol of national pride.

6. International Union for Conservation of Nature (IUCN), and its World Commission on Protected

Areas, has defined its Category II type of protected areas.

Example – Jaldapara National Park, West Bengal; Yellowstone National Park, USA etc.

NATIONAL PARKS (ACCORDING TO WILDLIFE PROTECTION ACT 1972)

Declaration of National Parks – (1) Whenever it appears to the State Government that an area, whether within a sanctuary or not, is, by reason of its ecological, faunal, floral, Geo-morphological, or zoological association or importance, needed to be constituted as a National Park for the purpose of protection & propagating or developing wildlife therein or its environment, it may, by notification, declare its intention to constitute such area as a National Park.

Provided that where any part of the territorial waters is proposed to be included in such National Park, the provisions of Sec.26A shall, as far as may be, apply in relation to the declaration of a National Park as they apply in relation to the declaration of a sanctuary.

- The notification referred to in sub-section (1) shall define the limits of the area which is intended to be declared as a National Park.
- Where any area is intended to be declared as a National Park, the provisions of Sec. [19 to 26- A (both inclusive except clause (c) of sub-section (2) of section 24)] shall, as far as may be, apply to the investigation and determination of claims and extinguishment of rights, in relation to any land in such area as they apply to the said matters in relation to any land in a sanctuary.
- When the following events have occurred, namely the period for preferring claims has elapsed, and all claims, if any, made in relation to any land in an area intended to be declared as a National Park, have been disposed of by the State Government, and
- all rights in respect of lands proposed to be included in the National Park have become vested in the State Government.
- The State Government shall publish a notification specifying the limits of the area which shall be comprised within the National Park and declare that the said area shall be a National Park on and from such date as may be specified in the notification.
- No alteration of the boundaries of a National Park shall be made except on a resolution passed by the Legislature of the State.

- No person shall, destroy, exploit, or remove any wildlife from a National Park or destroy or damage the habitat or any wild animal or deprive any wild animal or its habitat within such National Park except under and in accordance with a permit granted by the Chief Wildlife Warden and no such permit shall be granted unless the State Government, being satisfied that such destruction, exploitation, or removal of wildlife from the National Park is necessary for the improvement and better management of wildlife therein, authorizes the issue of such permit.
- No grazing of any [livestock] shall be permitted in a National Park and no livestock shall be allowed to enter except where such [livestock] is used as a vehicle by a person authorized to enter such National Park.
- The provisions of secs. 27 and 28, secs.30 to 32 (both inclusive), and CIS, (a), (b) and (c) of [Sec.33, 33A] and sec.34 shall, as far as may be, apply in relation to a National Park as they apply in relation to a sanctuary.

List of National Parks in West Bengal

Sl.No	**Name of National Park**	**Year of Notification**	**Total Area (km^2)**
	Buxa National Park	1992	117.10
	Gorumara National Park	1992	79.45
	Neora Valley National Park	1986	159.89
	Singalila National Park	1986	78.60
	Sunderban National Park	1984	1330.10
	Jaldapara National Park	2004	216.51

BIOSPHERE RESERVES

Biosphere reserves are sites established by countries and recognized under UNESCO's Man and the Biosphere (MAB) Programme to promote sustainable development based on local community efforts and sound science.The programme of Biosphere Reserve was initiated by UNESCO in 1971. The purpose of the formation of the biosphere reserve is to conserve in situ all forms of life, along with its support system, in its totality, so that it could serve as a referral system for monitoring and evaluating changes in natural ecosystems. The first biosphere reserve of the world was established in 1979, since then the network of biosphere reserves has increased to 631 in 119 countries across the world.

Biosphere reserves have three interrelated zones that aim to fulfil three complementary and mutually reinforcing functions:

I. **The core area**(s) comprises a strictly protected ecosystem that contributes to the conservation of landscapes, ecosystems, species and genetic variation.

II. **The buffer zone** surrounds or adjoins the core areas, and is used for activities compatible with sound ecological practices that can reinforce scientific research, monitoring, training and education.

III. **The transition area** is the part of the reserve where the greatest activity is allowed, fostering economic and human development that is socio-culturally and ecologically sustainable.

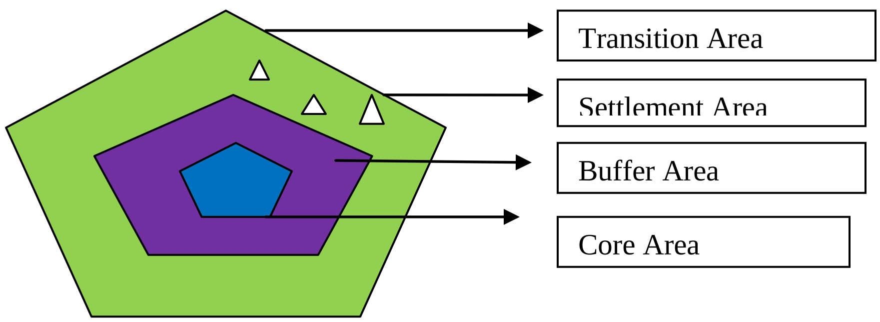


Figure 20 : Biosphere Reserve

SELF-ASSESSMENT TESTS

i. Define soil taxonomy?
ii. Discuss the major world patterns of soils with suitable diagrams.
iii. What is meant by soil nutrient?
iv. Specify the relationship between soil nutrients and soil organisms.
v. Analyse the role of micro-organisms in soil fertility with specific examples.
vi. Give the concept of soil pollution and soil degradation.
vii. Illustrate the concept of integrated management of soil.
viii. What is biogeography?
ix. What is the scope of biogeography as an interdisciplinary subject?
x. Briefly discuss about the significance of biogeography.
xi. What is plant ecology?
xii. Write a short note on physiological adaptation of plants.
xiii. Explain the concept of climax with suitable example.
xiv. What is plant succession?
xv. Describe the role of climate and soil on plant distribution.
xvi. Briefly discuss about means and barriers of dispersal of animals.
xvii. What is dispersal of animals?
xviii. What is biodiversity?
xix. Elucidate the controlling factors of biodiversity.
xx. What are the present issues of biodiversity?
xxi. What initiatives have been taken for biodiversity conservation?
xxii. What is IBP?
xxiii. Write a short note on Man and Biosphere Programme.

SUMMARIES

Biogeography is a sub-discipline of geography. It is the study of spatial patterns and temporal distributions and variations of all plants, animals and microorganisms. Plant ecology is the study of plant species in their natural environment. In biogeography, plant ecology studies the distribution and abundance of plants over the surface of the earth, the effects of environmental factors upon the abundance of plants, and the interactions among and between plants and other organisms. A plant species must either have evolved in an area or dispersed through natural processes or through human agency. Physiological adaptation is very much significant for survival of a plant species. Plant communities are broadly distributed into biomes based on the present dominant plant species. Adaptation is defined as the process by which an organism enables to exist under existing conditions of a habitat as well as environment. Environment always changes over a period of time and that influences on existing community, which is the cause of modification of existing community or replaced by another community. This relative occurrence of species over a specific area is known as succession or ecological succession. The end point or final stage of succession is referred to as stabilisation. In this stage, final community becomes less or more stabilised for a longer period of time by process of equilibrium with the environment. This ultimate community is known as climax community and the stage defined as climax stage. Environmental factors control plants growth and development, distribution and variation, etc. Climatic factors, among all the environmental conditions, dominantly control the life and development, distribution and variation of plants. Soil is another most dominating ecological factor called edaphic factor for plants growth and distribution. Impact of climate and soil on the distribution of plants is very much understood with changing biomes over the earth surface. The distribution of animals over the earth surface is the result of migration of animals over time into ecologically suitable region. Dispersal is the movement of organisms away from their birthplace. The spreading of animals from the place of their origin to other areas is called animal dispersal. Dispersal is an ecological process that plays an adaptive role in the life history of the organisms which are involved. In biogeography, there are several dispersal routes. Barriers are that type of environment or features in nature which are hostile to the organisms those are under long distance dispersal. Biodiversity denotes the variety of all life forms on earth and the essential interdependence of all living things. Biodiversity has wider significance and value in all level to support the life on earth. The International Biological Programme or IBP is an international level study on ecological and environmental issues during 1964 to 1974. It was an attempt to promote and apply the methods of science to ecosystem and urgent environmental issues. The fundamental aim of MAB programme is to improve human and environment relationship at global scale. The MAB programme also aimed to ecosystem specific networking development at regional, national and global level.

KEY POINTS

- Soil is the essential component of ecosystem and its characteristics vary across theworld.
- Soil Taxonomy helps to understand the variations of the soil characteristics over theearth.
- Plant growth depends on the availability of soil nutrients.
- Soil organisms are important component of soil. Soil organisms play crucial role in maintaining soil fertility.
- Soil degradation and soil pollution are two major concerns of soil geography.
- Soil degradation and soil pollution are caused by mainly anthropogenic factors.
- Soil degradation and soil pollution reduce soil fertility, decline agricultural productivity, deteriorate human health and pose threat on sustainability of the ecosystem.

- Introduction of integrated management of soil is utmost need of the hour.

REFERENCES

- Aislabie, J. &Deslippe, J.R. (2013). Soil microbes and their contribution to soil services. In J.R. Dymond (ed), *Ecosystem services in New Zealand – conditions and trends* (pp.143-161). Lincoln:ManaakiWhenua Press.
- Arnold, R.W. &Eswaran, H. (2002). Conceptual Basis for Soil Classification: Lessons from the Past. In H. Eswaran, T. Rice, R. Ahrens & B. A. Stewart (Eds.), *Soil classification: a global desk reference* (pp.27-42), New Work: CRC Press.
- Biswas, T. D. & Mukherjee, S. K. (1994). *Textbook of Soil Science*, Tata Mcgraw-Hill Publishing Company Limited, New Delhi.
- Bockheim, J.G. (2014). *Soil Geography of the USA: A Diagnostic-Horizon Approach*, Springer, New York.
- Bollen, W.B. (1959). *Microorganisms and Soil Fertility*, Oregon State Monographs, Number-1, Oregon State College, Corvallis, Oregon.
- Brady, N. C. (2002). *The Nature and Properties of Soils*, Prentice Hall of India Private Limited, New Delhi.
- Fairhurst, T. (ed.) (2012). *Handbook for Integrated Soil Fertility Management*, Africa Soil Health Consortium, Nairobi.
- Fang, H. Y. and Chaney, R. C. (2017): *Introduction to Environmental Geotechnology*, CRC Press, Taylor & Francis Group, New York.
- FAO & ITPS (2015). *Status of the World's Soil Resources (SWSR)* – Main Report, Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, Rome, Italy.
- FAO & WHO (2014). *The International Code of Conduct on Pesticide Management*, World Health Organization and Food and Agriculture Organization of the United Nations, Rome.
- Food and Agriculture Organization (FAO) of the United Nations. (2000). *Guidelines and Reference Material on Integrated Soil and Nutrient Management and Conservation for Farmer Field Schools*, AGL/MISC/27/2000, Rome.
- Foth, H.D. (1990). *Fundamentals of Soil Science*, John Wiley & Sons, New York.
- Hesammi, E., Farshidi, A., Sadatebrahimi, F, &Talebi A. B. (2014). The Role of Soil Organisms on Soil Stability; A Review, *International Journal of Current Life Sciences, 4*(11), 10328-10334.
- IFA, WFO & GACSA. (2016). *Nutrient Management Handbook*. International Fertilizer

Association, World Farmers' Organisation and GACSA Facilitation Unit, Food and Agriculture Organization of the United Nations.

- Jie, C., Jing-zhang, C., Man-zhi, T. &Zi-tong, G. (2002). Soil degradation: a global problem endangering sustainable development, *Journal of Geographical Sciences, 12*(2), 243-252.
- Jones, C. & Olson-Rutz, K. (2016). *Plant nutrition and soil fertility,* Nutrient Management Module No. 2, Montana State University Extension, 1-12.
- McCauley, A., Jones, C. & Jacobsen, J. (2011). *Plant nutrient functions and deficiency and toxicity symptoms. Nutrient Management Module, 9,*Montana State University Extension, 1-16.
- Oldeman, L. R. (1992). *Global Extent of Soil Degradation*, ISRIC Bi-Annual Report (1991-1992). The Netherlands, 19-36.
- Powlson, D. S., Hirsch, P. R., & Brookes, P. C. (2001). The role of soil microorganismsin soil organic matter conservation in the tropics. *Nutrient cycling in Agroecosystems, 61*(1-2), 41-51.
- Rodríguez-Eugenio, N., McLaughlin, M. &Pennock, D. (2018). *Soil Pollution: a hidden reality*, FAO, Rome.
- Soil Survey Staff. (1999). *Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys*, 2nd edition, Natural Resources Conservation Service, U.S. Department of Agriculture, Handbook 436.
- Thorp, J. & Smith, G. D. (1949). Higher categories of soil classification: orders, suborder, and great soil groups, *Soil Science*, 67(2), 117-126.
- Waksman, S. A., & Starkey, R. L. (1931). The Soil and the Microbe. In J.G. Lipman (ed.), *The Wiley Agricultural Series* (pp.1-260), New York: John Wiley and Sons.
- Young, R., Orsini, S. & Fitzpatrick, I. (2015). *Soil degradation: a major threat to humanity*, Sustainable Food Trust, Bristol, UK.

- Beeby, A. and Brennan, A.M. 1997 :*First Ecology*, Chapman and Hall, London. Chapman J.L. and Reiss, M.J. 1993: *Ecology: Principles and Applications*, Cambridge University Press, Cambridge.
- Dash, M.C. 1993: *Fundamentals of Ecology*, Tata McGraw-Hill, New Delhi.
- Dobson, A.P. 1996: *Conservation and Biodiversity*, Scientific American Library, New York. Galbraith, I. 1990: *Ecosystem and People: An Introduction to Biogeography*, Oxford University Press, Oxford.
- Jeffries, M.J. 1997: *Biodiversity and Conservation*, Routledge, London.

- Joseph, F.N.---: http://repository.kln.ac.lk
- Kormondy, E.J. 1984: *Concepts of Ecology*, 3rd edition, Prentice-Hall, India, New Delhi.
- Lincoln, R.J., Boxshall G.A. and Clark P.F. 1982: *Dictionary of Ecology, Evolution and Systematics*, Cambridge University Press.
- Myers, A.A. and Giller, P.S. (editors) 1988: *Analytical Biogeography: An Integrated Approachto the Study of Animal and Plant Distributions*, Chapman and Hall.
- Odum, E.P. 1997: Ecology: *A Bridge between Science and Society*, Sinaur Associates Inc. Publishers, Sunderland.
- Ricklets, R.E. 1993: *The Economy of Nature: A Textbook in Basic Ecology*, 3rd edition, W.H. Freeman and Co., New York.
- Singh, S. 1991: *Environmental Geography*, PrayagPustakBhawan, Allahabad.

  Singh, S. 2016: *Biogeography*, Pravalika Publications.
- Sharma, P.D. 1996: *Ecology and Environment*, Rastogi Publications, Mirat.
- Shukla, R.S. and Chandel, P.S. 1994.*Plant Ecology*, S. Chand & Company Ltd., New Delhi.
- Tivy, J. 1993: *Biogeography:A Study of Plants in the Ecosphere*, 3rd edition, Longman Scientific and Technical, Harlow.
- Weaver, J.E. and Clements, F.E. 1938: *Plant Ecology*, 2nd Edition, McGraw-Hill Book Co.
- World Wide Fund for Nature-India (Eastern Region) 1995: *Nature Conservation Handbook*, Calcutta.

Website Resources:

shodhganga.inflibnet.ac.in

www.wikipedia.org

www.unesco.org

www.nasonline.org

Disclaimer

This self-learning material is based on different books, journals and web-sources.

Self-Learning Material (SLM) for the Course of

M.A. / M.Sc. in Geography

Directorate of Open and Distance Learning (DODL)

University of Kalyani

Paper: GEO/CC/T-104

(Total Credit – 4; Total Marks – 50: Internal Evaluation – 10 + Semester-end Examination - 40)

University of Kalyani

Kalyani

Nadia – 741235

West Bengal

Course Materials Compiled by:	
Dr. Sayan Choudhary Assistant Professor (Contractual) Directorate of Open and Distance Learning University of Kalyani, West Bengal	**Group – GEO/CC/T-104: Geographical Thought**
Dr. Ayan Rudra Assistant Professor (Contractual) Directorate of Open and Distance Learning University of Kalyani, West Bengal	**Group – GEO/CC/T-104: Geographical Thought**

December 2021

Directorate of Open and Distance Learning, University of Kalyani.

Published by the Directorate of Open and Distance Learning,

University of Kalyani, Kalyani-741235, West Bengal.

Syllabus

Paper – GEO/CC/T-104: (Total Credit - 4, Total Marks – 50)

GEO/CC/T-104: Geographical Thought
(Marks - 50: Internal Evaluation – 10, Semester-end Examination - 40)

Unit. 01	Contributions of Greek, Roman and Indian scholars during the ancient period and Arab scholars during the medieval period
Unit. 02	Contributions of Humboldt and Ritter in Geography
Unit. 03	Social Darwinism and its importance in Geography; Morphology of cultural landscape (Carl O. Sauer)
Unit. 04	Major paradigms in Geography and their shift
Unit. 05	Dualism and Dichotomies in Geography: Physical and Human Geography, Regional and Systematic Geography, Ideographic and Nomothetic
Unit. 06	Positivism and Quantitative revolution in Geography
Unit. 07	System approach in Geography
Unit. 08	Critical revolution in Geography; Humanistic Geography; Radical Geography; Behavioural Geography
Unit. 09	Welfare Geography
Unit. 10	Feminism and Feminist Geography
Unit. 11	Postmodernism and Postmodern Geography
Unit. 12	Subaltern studies in Geography

Contents

PAPER – GEO/CC/T-104: (TOTAL CREDIT - 4, TOTAL MARKS – 50)

GROUP – GEO/CC/T-104:: GEOGRAPHICAL THOUGHT
(MARKS - 50: INTERNAL EVALUATION – 10, SEMESTER-END EXAMINATION - 40)

1.1 INTRODUCTION

Geography as a discipline evolved through different stages of historical development of mankind. During its inception geography was chiefly confined with descriptive studies of man and nature. The Greek, Roman and Indian scholars during ancient time made valuable contributions in the development of geography. Medieval period was chiefly a dark period for over all development of science. Modern geography was established by pioneers in Germany, France, Britain, Russia and the United States. During this time dualism and dichotomies in geography became prominent. Positivism and quantitative revolution in geography opened new vista in geographic research. Dissatisfaction with quantitative revolution gave way to critical revolution, welfare geography and feminist geography.

1.2 LEARNING OBJECTIVES

The present section aims to introduce the following topics -

- Contributions of Greek, Roman and Indian scholars during the ancient period and Arab scholars during the medieval period
- Contributions of Humboldt and Ritter in Geography
- Social Darwinism and its importance in Geography; Morphology of cultural landscape (Carl O. Sauer)
- Major paradigms in Geography and their shift
- Dualism and Dichotomies in Geography: Physical and Human Geography, Regional and Systematic Geography, Ideographic and Nomothetic approach
- Positivism and Quantitative revolution in Geography
- System approach in Geography
- Critical revolution in Geography; Humanistic Geography; Radical Geography; Behavioural Geography
- Welfare Geography
- Feminism and Feminist Geography
- Postmodernism and Postmodern Geography
- Subaltern studies in Geography

1.3 ASSESSMENT OF PRIOR KNOWLEDGE

Discussion about nature, scope and relevance of geography is necessary. Discussion about different branches of geography is necessary.

1.4 LEARNING ACTIVITIES

Preparation of short notes and essays on different topics of discussed.

1.5 FEEDBACK OF LEARNING ACTIVITIES

Debate and discussion on various topics discussed in the paper may be conducted.

Class seminar on various topics may be arranged

1.6 EXAMPLES AND ILLUSTRATIONS

UNIT - 1: CONTRIBUTIONS OF GREEK, ROMAN AND INDIAN SCHOLARS DURING THE ANCIENT PERIOD AND ARAB SCHOLARS DURING THE MEDIEVAL PERIOD

INTRODUCTION

In its wider connotation as a branch of knowledge concerned with the satisfaction of human curiosity about the lands and peoples away from one's home base, speculation regarding mysteries of the physical environment, and the role it plays in shaping the destiny of man upon the earth, geography is as old as human civilization. As such, each major cultural realm has had its own historiography of geography. Modern geography as practised over the entire world today represents, however, an outgrowth of the European geographic tradition so that the historiography of modem geography is essentially an account of the conceptual developments among Europeans regarding the nature of the earth and its environment and the way it influences man. Thus, the roots of modem geography are to be traced back to the thought of the ancient Greeks.

CONTRIBUTIONS OF THE GREEKS AND THE ROMANS

References to descriptive writings about lands and peoples in different parts of the earth's surface are found in the oral traditions of classical Greece and are reflected in the writings of Homer, whom the Greek geographers had themselves referred to as the father of geography. *Odyssey,* Homer's epic poem written sometimes in the ninth century B.C., presents geographical accounts of the lands and peoples located on the margins of the world then known to the Greeks. The poem records the wanderings of Odysseus to return to Ithaca after the fall of Troy, when he was blown off course by a storm, and it took him twenty years to reach home. The poem contains a geographical account of the distant places visited by the hero of the epic in course of his long journey. In it there are references to a land of continuous sunshine, and later of Odysseus's visit to an area of continuous darkness. Apparently a Greek poet could not have imagined these scenes. Somehow accounts about the nature of the earth in the far north of Europe during the long summer days and the continuous winter darkness had filtered back to Greece, and were woven with other geographical threads into an enchanting adventure story. As in the case of *Meghdoot* of Kalidas, many have tried to identify the many places referred to in the epic poem.

However, Thales of Miletus-a town located near the mouth of the Menderes river on the eastern side of the Aegean Sea (which was both a centre of learning and a flourishing centre of commerce)-who lived in the seventh and the sixth century B.C., is regarded as the first Greek to have devoted focused attention to the measurement and location of places on the surface of the earth. Thales himself was a very successful businessman. In the course of a business trip to Egypt, he had been greatly impressed by the geometrical traditions of the Egyptians in the measurement and computation of areas. He had introduced some of these ideas among the Greeks. Anaximander, a contemporary of Thales and a few years his junior, is credited with having first introduced the idea of the sundial consisting of a pole set vertically over a flat surface to measure the varying position of the sun by measuring the length and direction of the shadow cast by the pole. The shadow was shortest at noon and provided an exact north-south line for determining the correct longitude of the relevant place. Anaximander is also said to have produced a map of the known world with Greece as its centre. Thales and Anaximander have jointly been regarded as the originators of the mathematical tradition in geography in ancient Greece. The literary tradition in the writing of geography had also developed around the same time. Hecataeus, a resident of the town of Miletus, and born around 475 B.C.-about the time that Thales and Anaximander had passed away·-originated the literary tradition and his book *Periods Ges* (Description of the Earth) is regarded as the first known attempt to synthesize available knowledge about the world in a usable form. Hecataeus is also one of the earliest writers of prose in classical Greek literature.

The next great name in this context is that of Herodotus (circa 485- 425 B.C.) who is widely known as the father of history, but is also generally regarded as one of the founders of geography. His history of the Greek struggle with the "barbarians" included (as digressions) descriptions of various places visited by the author. Herodotus firmly believed that all history must be treated geographically and also that all geography must be studied historically. For him geography provided the stage, or the setting that gives meaning to historical events. Herodotus had travelled a great deal. Throughout his travels he had retained a keen interest in the nature of the landscape so that he not only described geographical phenomena but also tried to explain them. Examples include his attempt to explain the annual fluctuations in the flow of the Nile, and the processes involved in the origins of deltas occurring at the mouth of the Meander (Menderes) river at Miletus. Herodotus had no interest in the mathematical tradition and showed no interest in problems like measurement of the earth's circumference. He accepted the Homeric view of the earth as a flat disc over which the sun was believed to travel in an arc from east to west.

Plato (428-348 B.C.) also made an important contribution to the development of geographical ideas. Plato was a great proponent of deductive reasoning. He insisted that the observable phenomena on the earth's surface represent poor copies of ideas from which these observable phenomena had degenerated. By way of illustration he referred to the case of Attica (the ancient territory of which Athens was the capital). According to Plato, the area was originally very fertile and capable of supporting a large population of men and animals. He wrote that compared to its original state, the Attica of his time was "like the skeleton of a sick man, all the flesh and soft earth having been wasted away, and only bare framework of the land being left" (cited in Glacken, 1967, p. 121). Contemporary philosophers in Greece generally accepted the idea that symmetry of fotn, is one of the essential alttibutes of perfection, and that the most completely symmetrical for11, was a sphere. It was argued that since the earth had been created to serve as the home of man, it must have a perfect form, and therefore it must be a sphere. Plato is regarded as the first scholar who put forward the concept of a spherical earth located in the centre of the universe, and the sun and all the other celestial bodies moving around it. Plato offered no argument or evidence as proof that the earth is round. Providing the proof for the spherical shape of the earth was left to Aristotle (384-322 B.C.), who was a student and a member of Plato's academy for twenty years.

Aristotle is regarded as the pioneer of inductive reasoning and the inductive approach to acquiring knowledge. He was convinced that the best method of building a reliable theory was to start with the observation of facts. This required reasoning from the particular to the general, in contrast to Plato's deductive approach which required the student to proceed from the general to the particular. Aristotle laid the foundation of what has been regarded as the world's first paradigm to guide research procedures. He laid down four fundamental principles of scientific explanation: First, it is necessary to establish the necessary characteristics (i.e., the nature) of the phenomenon being investigated; Second, it is necessary to identify the substance of which it is composed; Third, it is necessary to identify the process through which the phenomenon has attained its present form; and lastly, it is necessary to identify the purpose that the phenomenon concerned fulfils in the overall scheme of nature. This last principle makes Aristotle stand out as the first teleologist in that he believed that everything was changing in accordance with a preexisting plan.

Aristotle argued his propositions so convincingly that his research methodology appeared irrefutable at the time it was presented. His intellectual status in the contemporary world of scholarship was so high that his ideas were accepted without question for a long time. (Some of his ideas were patently false, however. One such was the idea that habitability on the earth surface is a function of distance from the equator, and that areas around the equator are too hot for human survival.)

Although Plato and Aristotle gave intellectual leads that contributed to the development of knowledge about the earth as human habitat, neither of the two could be identified as a geographer. As contrasted to this, Erastosthenes (276-194 B.C.) is often referred to as the father of geography as a branch of knowledge. He is said to have coined the word geography. The term is derived from *ge,* meaning the earth, and *graphe,* meaning description. Thus was born geography as a field of study which specialized in presenting reasoned description of the Greek ecumene, and speculated about the nature of peoples and places beyond the range of knowledge in contemporary Greece. He wrote the first formal text on geography entitled *Geographica.* His estimate of the earth's circumference was remarkably accurate, and, therefore had proved a major step forward in the development of knowledge about the earth. Erastosthenes was the chief librarian at the famous museum at Alexandria-a post that he occupied for about forty years until his death in 194 B.C. Under his leadership, the museum had developed into a major centre of astronomical research, a field of knowledge that was at that time viewed as closely associated with geography. Erastosthenes identified five climatic zones, one torrid zone, two temperate zones, and two frigid zones. He also improved upon the Aristotelian idea on this subject by giving latitudinal boundaries to the five climatic zones. The *torrid zone* extended 24° north and south of the equator, and the *frigid zones* extended to 24° from either pole. The areas in between were the two *temperate zones.*

After the death of Erastosthenes the post of chief librarian of the museum at Alexandria went to Hipparchus who was the first to divide the circle into 360 degrees. He also defined a grid of latitudes and longitudes for the earth, and identified the equator as a great circle that divides our spherical planet into two equal parts. Hipparchus pointed out that since the earth makes one complete revolution in 24 hours, it covers a journey of 360 degrees in a day and so covers fifteen degrees of longitude in one hour. He also made a significant contribution to the development of map projections by suggesting ways for overcoming the difficulty of representing the spherical earth on a flat sheet of paper.

The cartographical cosmographical traditions set by Erastosthenes and Hipparchus were further advanced by the succeeding generation of students at the museum. The cumulative knowledge gained through these exertions culminated in Ptolemy's (9-168 A.D.) eight volume work entitled *Guide to Geography.* Ptolemy was himself a great astronomer of his time and was the author of the famous text on classical astronomy entitled *Almagest* which had for long remained the most standard reference on the movement of celestial bodies. His *Guide to Geography* was also of related interest. By adopting the system of latitudes and longitudes based on the division of the circle into 360 degrees, he attempted to give precise location for all the known places in precise mathematical te1111s. Six out of the eight volumes of his *Guide to Geography* consisted of tables of latitudes and longitudes. The first volume was devoted to a discussion on map projections, and the eighth volume contained maps of different parts of the world showing all the places that had been included in volumes two to seven. It is true that from the perspective of the present, Ptolemy's book would appear as a monumental collection of errors. It was, however, a piece of great scholarship at the time when it was originally presented. Ptolemy's calculations of latitudes and longitudes are found to be wrong since these calculations had been based on estimated lengths of journeys between places; and these could never be accurate. Another major source of error was that Ptolemy had rejected Erastosthenes' almost correct estimate of the earth's circumference in favour of Posidonius's (which gave a figure that fell short of the actual by a little over one-fourth). (Erastosthenes had estimated the earth's circumference at 252,000 stadii, and Posidonius at 180,000 stadii (one stadium being equal to 157.5 metres.)

Strabo (64 B.C. to 20 A.D.), born a century-and-a-half before Ptolemy, had carried forward the tradition of topographical work of Greek geography as started in the works of Herodotus. His seventeen-volume work named *Geography* was largely an encyclopaedic description of the world known to the Greeks. Unlike the works of most other Greek scholars, Strabo's book was found almost intact. The first two volumes of his book contain a review of the work of other geographers since the time of Homer. They give a fair idea of the nature of geographical writing in ancient Greece. The next eight volumes were devoted to Europe, six to Asia, and one to Africa. Strabo's book was written to cater to the needs of a specific group of readers, namely the officers of the administration, statesmen, and commanders of the Roman empire. The purpose was to provide a handbook of information about places and people to help the imperial officers in the better appreciation and accomplishment of their task.

Strabo's book had laid down a clear foundation for chorological writing in geography. Explaining the method of writing geography, Strabo wrote:

... just as the man who measures the earth gets his principles from the astronomer and the astronomer his from the physicist, so too, the geographer must in the same way take his own point of departure from the man who has measured the earth as a whole, having confidence in him and in those in whom he, in his tum, had confidence, and then explain in the first instance, our inhabited world, its size, shape and character, and its relation to earth as a whole; for this is the particular task of the geographer. Then, secondly, he must discuss in a fitting manner the several parts of the inhabited world, both land and sea, noting in passing where the subject has been treated inadequately by those of our predecessors whom we have believed to be the best authorities on the matters (Strabo, trans., Jones, 1917, pp. 429-431; cited in James, 1972, p. 47).

Both Strabo and Ptolemy had lived at a time when the Roman empire was at its zenith. It was the largest centralized empire in history till that time. The state needed to have exact description of its territories as well as the other territories it interacted with. This knowledge was necessary both for effective administration and trade as also for the training of the younger generation from among whom the future crop of administrators was to be recruited. The work of the two Greek scholars, besides extending the frontiers of knowledge, was designed to meet a definite need of society. Geography was flourishing because it served a useful purpose.

GEOGRAPHY IN THE MIDDLE AGES

During the fifth century A.D., the Roman empire suffered demise. The central administration had greatly weakened and consequently the constituent territories gradually became independent. As trade and commerce declined, the geographic horizons of the people rapidly narrowed down so that, with the passage of time, the geographical horizon of most people in Christiandom became confined to their immediate surroundings. Given the extremely narrow world-view of contemporary European societies, it was natural that religious orthodoxy should increase. Before long, scriptures had begun to be regarded as the ultimate repository of knowledge of every kind so that an impression was created that there was no need to learn anything outside the Holy books. Anything that did not conform with the "truth" of the scriptures was regarded as the product of a perverse mind and had, accordingly, to be rejected. Under these conditions, science (and, therefore, geography) could not develop and the Middle Ages represented the Dark Age in the history of scientific knowledge in Europe. During this long period, scientific concepts developed by the ancient Greeks were reshaped with a view to make them conform with the "truth" preached by the Church. For example, the idea of a spherical earth was abandoned in favour of the old concept of the earth as a flat disc, with Jerusalem as its centre. This dismal state of affairs continued almost until the end of the twelfth century A.D.

By the end of the eleventh century A.D., overland travel of Christian pilgrims to Jerusalem across Turkey and Syria had been made very difficult on account of Muslim domination over these territories. This aroused the religious sentiments of Christian Europe. A series of military campaigns were organized with a view to rescue the Holy Land of Jerusalem from the control of Muslims. Between 1096 and 1270 A.D., eight different crusades were organized for the purpose. These crusades (religious wars) played a major role in broadening the geographical horizon of Christian Europe. Men from different parts of Europe had come together to participate in them. These participants went back to their homes with new knowledge and information about the landscapes and customs of many areas beyond the range of the familiar. This stimulated interest in, and the urge to gain knowledge about, unfamiliar places. The religious wars, therefore, had led to a new beginning-a revival of interest in geography as a branch of knowledge. Expeditions began to be organized to distant places. The most famous of such expeditions was Marco Polo's voyage to China, the Far East, and the Indian Ocean undertaken between 1271 and 1295.

The Crusades proved a stimulant to the revival of interest in the study of peoples and places in far-off lands in another way also. Owing to the "religious" wars the Muslims had closed the overland routes to India and beyond to European merchants who had until then participated in the highly profitable spice trade between India and Europe. Attempts were, therefore, directed to finding an alternative route to the Indies. Two such attempts led to the glorious discoveries of Columbus and Vasco de Gama.

GEOGRAPHY IN THE ARAB LANDS

The fall of the Roman empire, and the decline of scientific learning in Europe was followed by a period of great ascendancy in the Muslim world which, under the influence of Prophet Mohammed, had been transformed from a multitude of tribes divided by intertribal feuds into integral components of a larger all-inclusive identity based on adherence to a common set of religious beliefs and practices. The followers of Islam soon embarked upon a course of conquest of the world outside Arabia with a view to spreading the new religious ideology to the farthest corners of the world. Persia and Egypt were conquered in 641-642 A.D., and by A.D. 732 the whole of the West Asian desert region was under their control. They soon overran the Iberian peninsula and Spain and Portugal remained under Muslim rule for almost nine hundred years. Muslim influence also extended eastward into India and parts of south-east Asia. The act of holding on to such a huge politico-cultural empire had, in itself, become a major stimulant to the rise of interest in geographical learning. The Arabs held a monopoly over the spice trade between India and Europe. This trade required a great deal of travel over land and sea. Travels between places spread over such a large expanse of territory became the source of considerable extension of knowledge about geographical environment in tropical regions.

Following the widening of its geographical horizon, the Arab world became fired with a new zeal for scientific learning. Baghdad (founded in 726 A.D.) became a major centre of learning. Its rulers (the Caliphs) employed learned men of different faiths to make authentic translations of the major scientific works of their respective languages. Included in these works were books on astronomy and geography. Scholars were also employed to calculate the circumference of the earth, and to fix latitudes in the plain of the Euphrates. The method employed was the one used by Erastosthenes about a thousand years earlier. Available texts on geography written by the ancient Greeks (including Ptolemy) were translated into Arabic, and new texts were got written after duly incorporating the new knowledge derived from the records of observations made by Arab merchants and explorers.

Thus, as a result of Ibn-Hakul's voyage to the south of the equator (made between 943 to 973 A.D.), the wrong notion regarding the inhabitability of the torrid zone (as perpetuated by Aristotle) was abandoned. Around the same time, in course of his travels down the east coast of Africa upto the Mozambique, Al-Masudi had reported the phenomenon of monsoonal winds. Another contemporary named Al-Maqdisi had established (in 985) the general truth that the climate of any place is a function not only of its latitude but also of its position on the east or west side of a landmass. He is also credited with the knowledge that most of the earth's landmass lies north of the equator. Al-Idrisi (1099-1180 A.D.) made many corrections to Ptolemy's book. On the basis of the new information collected by Arab explorers, and some collected from other sources, he produced his own book on geography (in 1154) in which the Greek idea about the Indian Ocean as a closed sea was corrected. Also corrected were the positions of many rivers including the Danube and the Niger.

Another great Arab explorer was Ibn Batuta (1304-1368 A.D.). He extensively explored regions of North Africa and West Asia, sailed along the Red Sea, and travelled south along the east coast of Africa to Kilwa, a point about 10 degrees south of the equator. He had reported that an Arab trading post was located in the latitude of 20 degrees south thereby confirming the habitability of the torrid zone. Ibn Batu ta had also travelled overland from Mecca to Persia, Bukhara and Samarkand, and from there across Afghanistan to Delhi. I-le had visited several islands, including Ceylon (Sri Lanka), Sumatra, and the Maldives. He also visited China and returned to Fez, the capital of Morocco, in 1350 A.D. From there he travelled across the Sahara to Timbuktu on the Niger, gathering valuable information about Black Muslims on the way. He settled down in Fez in 1353 A.D. after travelling an estimated 75,000 miles, a world record for his time. On the request of the king of Morocco, he put down a detailed account of his travels for posterity. The last great Muslim scholar who contributed significantly to the development of geographical knowledge was Ibn-Khaldun (1342-1405 A.D.) who wrote (in 1377) a detailed introduction to world history published under the title *Muqaddimah.* In his introduction to the book he identified two sets of influences on man's progress (i.e., history): One, the physical environment, and two, the social environment derived from culture and belief rather than the natural environment. This distinction between the two sets of environmental influences on man was a remarkable intellectual achievement for his time so that Kimble (1938) was prompted to remark that Ibn-Khaldun had "discovered ... the true scope and nature of geographical inquiry".

lbn-Khaldun had concentrated on the study of the tribe and the city the two most important elements in the political organization of the desert society in the Arab world. He identified the tribe and the city as two distinct stages in the evolution of social organization in a desert environment. While the nomads represented an earlier (primitive) stage of social organization, the city dwellers represented the last stage in the development of social life, almost the point where decay sets into the social organization owing to the sedentary lifestyle of the urban community. Many credit Ibn Khaldun with having presented in this way one of the earliest concepts of the life cycle of the states. Surprisingly, however, Ibn-Khaldun had clung to the Aristotelian idea about the inhabitability of the equatorial regions. To the great credit of lbn-Khaldun is the fact that he was the first great scholar to direct attention specifically to the study of the man-environment relationships.

The significance of Arab contribution to the historiography of modern geography lies in that the development of geographical knowledge in the Arab world represented, in some ways, a further development over the original base provided by the geographers of ancient Greece, whose works had been translated into Arabic, and widely used by Arab scholars. Thus, while Europe itself had forgotten the Greek heritage in geography, the Arabs had held the banner aloft and it was largely through contact with the Arab world, and the translation of their books (including retranslation of Greek works from the Arabic translations) that geography got revived as a living science in fifteenth century Europe. Indeed, the countries having the closest contacts with the Arab world, such as Spain and Portugal, spearheaded the series of exploratory voyages that ultimately opened the way to the revival of interest in geographical learning. Large parts of Spain and Portugal had remained under Muslim rule since the eighth century A.D. While Portugal had become free in the middle of the thirteenth century, the Spaniards pushed out the Muslims gradually from the peninsula through a series of efforts lasting over a century from 1391 to 1492.

Both the Portuguese and the Spaniards had mastered the art of shipbuilding and navigation, and had launched ambitious programmes of voyages of exploration with a view to promoting trade and commerce with the outside world, particularly the spice trade with the Indies, and trade in gems and precious stones with parts of Africa south of the Sahara. Since the overland routes in each case were then under the control of the Arabs, it was necessary to find alternative sea routes, which contemporary science had shown to be well within the pale of possibility in view of the round shape of the earth, and the continuity of the oceans. The importance of the spice trade for contemporary Europe lay in that owing to the inadequate supply of sugar, spices were required to make food palatable. Besides, in the absence of refrigeration, meat was to be stored in dried and salted form.

Such meat required spices in order to be made reasonably palatable. What was more, Genoa and Venice which had earlier been flourishing centres of trade in spices from India and beyond, were now deserted as the Arabs had blocked direct contact between Europe and the regions of supply further east in Asia.

THE AGE OF EXPLORATION

Portugal's Prince Henry "The Navigator" who had (in 1415) succeeded in capturing the Muslim base at Ceuta on the southern side of the Strait of Gibraltar, took the first initiative toward wider exploration across the high seas. From his Muslim prisoners, the Prince had learnt that many of the most valuable items of merchandise traded in European markets by the Arabs were brought from areas in Africa to the south of the Sahara. This inspired him to sponsor sea voyages of exploration along the western coast of Africa. Around this time-in 1410-two important publications had appeared in geography. One was the Latin translation of Ptolemy's *Geography* made from a copy preserved in Byzantium (Istanbul), and the other was a book called *Imagine Mundi* authored by Pierre d' Ally in which he presented a summary of various geographical writings then existing in the countries of Christian Europe. The two were very influential in promoting interest in geographical knowledge and created a favourable climate for the launching of voyages of exploration, and for developing better techniques of cartography and map design.

Prince Henry was responsible for establishing a broad-based institute of geographic research at Sagre near the port of Lagos where a rich library of all the available literature in geography, cartography, astronomy and related subjects was stocked, and scholars (including linguists) from all parts of Europe were invited to teach Portuguese students the art of navigation, and to inform them about the existing knowledge regarding the earth and its environment in different parts of the world so that suitable preparations could be made to meet the contingent situations likely to be encountered in the process of exploratory voyages then being planned for the exploration of the western coast of Africa, and to find an alternative route to the spice islands beyond India. These explorations initiated by the Prince laid firm foundations for the larger ventures by subsequent explorers under the patronage of the royal house of Portugal, culminating in the great voyage of Christopher Columbus (who discovered the new World in 1492) and the discovery of an alternative sea route to India by Vasco da Gama in 1498.

Columbus had studied at Sagre, and he had been greatly influenced by Pierre d'Ally's *Imagine Mundi* which had suggested that since the earth was round, a route to China and India could be found by sailing west from the Canary Islands. It is a different matter though that in the process of finding an alternative route to the Indies, Columbus landed in America rather than Asia, his intended destination. Columbus died in 1506 still believing that he had discovered a part of Asia. The task of finding an alternative route to Asia by sailing west and then north along the coast of South America was accomplished by another great Portuguese explorer three decades later in October 1520. This voyage was performed by Magellan whose name the (Magellan) Strait connecting the Atlantic with the Pacific now bears**.**

Voyages across the sea required maps and charts to guide the sailors in course of their travels. Ptolemy's map was used in the beginning. It was the task of the royal cartographers to correct the old map in the light of new information then available. Thus, Venice and Genoa soon emerged as great centres of cartographic learning. European sailors and merchants departed from either of the two port towns for their journeys to the eastern Mediterranean to pick up the cargo brought by Arab ships from the East. The first globe showing the earth as a sphere was produced by Martin Behain in 1490, and map projections tackling the problem of representing the round earth on a flat sheet of paper began to receive attention of scholars soon after. In 1530, Peter Apian produced a heart-shaped map of the earth in which both latitudes as well as longitudes were shown as curved lines. Neither distance nor direction was represented correctly, and the map showed only one hemisphere. Apian's student, Gerhard Kramer (who later adopted the name Gerhard Mercator) made a world map in 1538 by joining two heart-shaped projections:-one for each hemisphere. Mercator earned celebrity in 1569 when he succeeded in designing a projection that showed the whole of the earth surface on a single network of latitudes and longitudes.

This was the famous Mercator Projection-the orthomorphic cylindrical projection. As we know, even though theoretically an orthomorphic projection, it greatly distorted the shapes of continents, but its great advantage lay in that on it compass bearings could be shown by straight lines so that navigators could plot their course without being required to draw cumbersome curves. The projection could not be easily used until English geographer Edward Wright (1558-1615) produced the trigonometric table to reproduce the projection. This improvement made the Mercator projection universally accepted for maps on which to base navigators' charts. Focus on improvements in cartographic techniques continued through the sixteenth and seventeenth centuries. New projections were devised and old map projections improved upon. Map makers remained busy revising old maps in the light of new information obtained from travellers and explorers. From the time of Magellan (who explored the outlines of South America between 1518 to 1521), and James Cook (1728-1779) (who through his three different voyages, performed between 1768 to 1779, drew the outlines of the Pacific Ocean and eliminated the possibility of the existence of Ptolemy's Southland), scholars were directly addressing the task of drawing correct outlines of landmasses and water bodies. They were also busy devising techniques of surveying and cartography to be able to present true-to-scale reality of the earth's surface on their maps and charts. This task was almost complete by the time of James Cook's death in 1779 and a good deal of new information about world climates, wind regimes, distribution of flora and fauna, and patterns of human civilization over the earth surface had been obtained. Incorporating the ever increasing information and data with a view to presenting a correct and meaningful description of the earth surface had become a formidable task. The challenge posed by the problem attracted a number of leading scientists to the study of geography.

While explorers were busy fixing the outlines of continents and oceans, and cartographers remained busy in drawing more accurate representations of the earth surface on maps, the world was experiencing a great revolution in knowledge about the nature of the universe and the earth's position in it. The old-time concept of the earth as the centre of the universe was abandoned in favour of the concept of a heliocentric universe first put forward by the Polish scholar Nicolaus Copernicus in 1543. The concept was further refined by Kepler (1571-1630) in 1618 and Galileo (1564-1642) in 1623. Galileo further revolutionized scientific thinking by formulating the concept of mathematical order in the universe i.e., an order in which relationships between phenomena could be described in terms of mathematical laws rather than verbal logic. A further scientific advance came in the form of Newton's law of gravitation in 1686. Thus, in the course of a century and a half, seeds of scientific revolution had been sown. These heralded the beginning of the rise of specialized branches of knowledge, each focusing on some particular theme, object, or relationships between phenomena. The rise of specialized systematic sciences, each focusing on a particular category of facts or relationships signalled the demise of the era of universal scholarship and of cosmographies in which scholars had attempted to bring together all that was known about the earth and its parts in single volume works in the style set by Strabo.

In view of the rapid flow of new information derived from the increasing stream of explorations and scientific research, compilation and synthesis of knowledge in a meaningful manner became an increasingly challenging task that required a high degree of scholarship. The cosmographers of that period were, therefore, far from mere popularizers. The first great cosmographer of the age of exploration was the German scholar Sebastian Munster (1489-1552) who had been engaged, with the help of 120 other authors and artists, in writing a broad-based cosmography incorporating the latest information on every important aspect for over eighteen years. The outcome was a six-volume work entitled *Cosmographie Universal is* published in 1544. Written in the tradition of Strabo's *Geography,* the book earned its author the popular title of "the German Strabo". The first volume of Munster's cosmography presented a general picture of the earth on the lines of Ptolemy's *Geography,* while the remaining five volumes were devoted to descriptive accounts of the major divisions of the earth's surface.

Munster's work was a combination of tradition (imaginative stories about people and places which were part of popular belief) and science (incorporating new information derived from explorations and scientific investigations). Thus, his account of America and Africa included stories of men with heads on their chests, and having a combined animal and human form. Such beliefs were part of contemporary scholarship and consequently, Munster's volumes were avidly read necessitating several editions between 1544 and 1550; and the book remained a popular reference for about a century thereafter. Another leading cosmographer was the German scholar Cluverius (1580-1622) who had published a six-volume compendium on universal geography (following the general plan of Munster's work but better informed) in 1624. The first universal geography to appear in the English language was written by Nathanael Carpenter (1589-1628) a scholar at Oxford, who had benefited from his association with Cluverius during the latter's frequent visits to Oxford. Carpenter's book had appeared in 1625, a year after the publication of Cluverius's book.

FROM COSMOGRAPHY TO SCIENTIFIC GEOGRAPHY: CONTRIBUTION OF BERNARD VARENIUS

In course of time, the tradition of writing cosmographies got concretized into a coherent body of knowledge that came to be described as "general geography". *Geographia Generalis* (1650) of Bernard Varenius (1622-1650) (a Dutch scholar) was an outgrowth from the cosmographic tradition even though it is rightly regarded as a major step forward toward laying the foundation of scientific geography. Varenius' s book was, according to Dickinson (1969), the first work "which sought to combine general, mathematical, and physical geography and chorology". Varenius set forth clearly the distinction between two forms of geographical scholarship-the one concerned with the description of particular places (i.e., regional description), and the other concerned with developing general laws and hypotheses of wider applicability. He termed the first Special or Particular Geography (i.e., geography of particular places etc.) and the second as General Geography.

Varenius was writing at a time when voyages of exploration were pouring in a flood of new information and data so that one of the major problems facing contemporary scholars was how to relate specific pieces of information to general principles. Among geographers, Varenius was the first to focus attention on this problem; and the solution that he offered through his *Geographia Generalis* was to become the basic tenet of geography as a branch of knowledge which has ever since retained a twofold division into Regional and Systematic (or General) geography-the former focused on the study of particular places, and the latter was devoted to the study of the nature, and pattern of spatial distribution of particular items of geographical interest over the earth surface and its parts.

The most creditable part of Varenius's contribution lay in that he underlined the relationship between the two streams of geographical scholarship: Special geography provided the results of in-depth study of particular places and regions which became the raw material (the data) on the basis of which General geography could pursue its task of depicting spatial patterns of distribution, and inferring therefrom general hypotheses and laws explaining why they occur where they do, and thereby providing valuable inputs for better work in the area of Special (i.e., Regional) studies. Varenius pointed out that while Special geography was of great practical value in the pursuit of government and commerce, General geography provided information on the principles governing the distribution of particular phenomena on the earth surface so that the administrator and the businessman may be suitably informed about the nature of the environment they are likely to encounter in particular parts of earth's surface. To Bernard Varenius, therefore, General and Special geography did not suggest a dichotomy or a separation of ways, and division of objectives. To him, the two represented mutually interdependent parts of geography as a unified field of scientific learning. In this vision of geography, Varenius was far ahead of his peers. This explains why he had so greatly influenced the concept and scope of geography in Europe for well over a century.

In the foreword to his book, Varenius had set out a plan for Special geography, according to which the description of particular places should be based on celestial conditions, including climate; terrestrial conditions, including relief, vegetation and animal life; and human conditions including trade, settlements and forms of government in each country being studied. It is true though, that Varenius was none too enthusiastic about human geography since its subject matter could not be put to exact mathematical analysis for purposes of generating laws of behaviour (Gettfried Lange, 1961, paraphrased in Holt-Jensen, 1980, p. 14).

Like most other great works of scholarship, Varenius's book had been inspired by the demands of his time. In 1647, Varenius had accepted the position of a private tutor in a family in Amsterdam, then the commercial hub of the Netherlands. Here he came in contact with merchants engaged in international trade. Many of the merchants needed information about Japan where the Dutch had established a trading post in Nagasaki. This is what had inspired his first book entitled *Regional Description of Japan and Siam* published in 1649. The experience gained in writing a regional geography of Japan gave Varenius the idea that descriptions of particular places "could have no standing as contribution to science so long as these are not related to a coherent body of general concepts". His *Geographia Generalis* was written with a view to promoting the search for and the building of this much-needed conceptual coherence in geographical scholarship. His book went through several editions in Latin-two of these (published in 1672 and 1681) edited by no less a person than Sir Isaac Newton. An English edition was published in 1693 (Baker, 1955a and 1955b).

Varenius passed away in 1650 at the tender age of 28 so that the world of scholarship was deprived of many more conceptual leads. Underlining the methodology of Special vis-a-vis General geography, Varenius pointed out that while in General geography (dealing mostly with phenomena of physical origin), most things can be proved by mathematical laws, in the case of Special geography, with the exception of celestial features (i.e., climate), things must be proved by experience that is, by direct observation through the senses (James, 1972, p. 226).

THE IMPACT OF DISCOVERIES

NEW ANSWERS TO QUESTIONS ABOUT THE ORIGIN OF THE EARTH AND ITS SURFACE FEATURES, AND MAN'S PLACE IN NATURE

Speculation about the origin of the earth, and man's place in the web of nature, had for long remained constrained by theocratic domination of thought in mediaeval Europe. Intellectual thinking had continued to be conditioned by traditions inherited from ancient Greece as well as from biblical accounts. All this began to change during the seventeenth century when steps were initiated to cut the thought process loose from the strangle hold of biblical beliefs, and to start experimenting with rational methods, so that geographical exploration had "immense significance in the history of science and of thought" (Parry, 1981, p. 3).

By the end of the seventeenth century, a good deal of speculation on the origin of the earth had led to the belief that the earth is a physical phenomenon that has acquired its present form though natural processes of change spread over millions of years, and that it was wrong to regard it as a divine creation. Inspired by the theory of comets given by Edmond Halley (1656-1742) in 1682, William Whiston (1667-1752) developed the theory that the earth was made from the debris of a comet, and that the gravitational pull of a second approaching comet had caused the elliptical orbit of the earth around the sun, and had led-through the tidal waves caused by its gravitational pull-to the creation of the continents and ocean basins. While the crests of the waves were occupied by landmasses (continents), the troughs became the ocean basins. Later the German scholar Abraham Gottlob Werner (1749-1817) developed the theory that the great flood that had been caused by the cooling of the earth's atmosphere, had led to dissolution of materials of the earth's crust. The dissolved materials of the crust were deposited on the surface of the earth in the form of a series of layers so that large parts of the earth surface are covered with sedimentary strata.

Simultaneously, a good deal of speculation had begun on the origin of landforms. In 1719, John Strachey (1671-1743) showed that landforms reflected the rock structure lying underneath them. Subsequently, in 1777 Simon Pallas (1741-1811) published geological maps to show that the cores of most mountain ranges are made of granite. Alongside, ideas about the mechanics of river flows and valley development were being developed at a rapid pace. The French scholar Louis Gabriel Comte Du Buat (1734-1809) mathematically explained (in 1786) how the flowing water of a river can establish equilibrium between velocity and the load of sediment being transported by it. This had ledto the idea of "graded river profiles". During the 18th century James Hutton (1726-1797) popularized the concept of uniformitarianism, according to which the processes that shape the earth surface indicate a perpetual process of change "with no vestige of a beginning and no prospect of an end".

New methods of scientific classification of plant and animal life were also influential in shaping geographic thought and practice. The most influential figure in this field was Swedish botanist Carolus Linnaeus (1707-1778). He developed a system of classification based on classes, orders, genera, and species. French naturalist Lamarck (1744-1829) drew attention to the need for a system of classification of plants and animals in accordance with their natural characteristics. He challenged the widely believed notion that plants and animals were created in their present form. Thus, he presented the rudiment of a theory of evolution that was later advanced and refined in a big way by Darwin and others who laboured to explain the mechanism through which the process of evolution of life forms had taken place. Coupled with the idea of uniformitarianism of Hutton, the theory of biological evolution had greatly impressed geographers about the role of time in the evolution of landforms.

The early part of the eighteenth century also witnessed the first beginnings of scientific study of man. The German scholar J.P. Sussmilch (1707-1767), in a book published in 1741, had demonstrated the existence of statistical regularities in population data. His research showed that the ratios between the sexes had remained nearly balanced, and that birth and death rates could be predicted on the basis of past trends. (However, the idea that numerical information about individuals tends to group around averages in accordance with the theory of probability) was put forward by Lambert Quetelet (1796-1874) only much later in 1848.)

As knowledge about the lands and peoples over the earth's surface increased, so did speculation about the role of the environment in shaping human behaviour patterns. French philosopher Jean Bodin (1529-1596) was one of the first to present a major work on this theme in 1566. Placing belief in the Greek concept of climatic zones, Bodin formulated the theory that people in the southern parts of the world (being under the influence of Saturn) are religious by nature; those living in the northern regions (living under the influence of Mars) were endowed with martial characteristics; and only people living in the middle regions (owing to the influence of Jupiter) were able to evolve a civilized way of life and live under the rule of law. English geographer Nathanael Carpenter in his *Cosmography* (1625) further advanced Bodin's idea regarding climatic zones and their influence in shaping human behaviour.

From these early beginnings of what may appear to us today as unscientific speculation about the man- nature relationships, progressively evolved more rational scientific analyses based on detailed observations and comparative case studies. In a piece published in 1719, the French scholar Abbe de Bos established a definite relationship between the weather and suicide rates in the cities of Paris and Rome. His analysis showed that in Paris, suicides were most common in the period before the onset of winter and just after the end of winter. In Rome on the other hand, most suicides had occurred in the two hottest months in summer (Glacken, 1967, pp. 556-558). Until the 19th century the most influential scholar who worked on this theme was Charles Louis Montesquieu (1689-1755). In line with the scientific ideas current in his time, Montesquieu wrote that warm climates favoured growth of despotism and slavery, whereas colder climates encouraged democracy and freedom so that, according to him, democracy tended to increase in direct proportion to increase in distance from the equator. Despite these crude observations on the relationship between man and the environment, Montesquieu was far from a crude determinist (Kriesel, 1968); he had given due allowance to human initiative and technology in reducing environmental constraints to human progress.

The progress in scientific thought through the newly acquired habit of questioning everything in sight represented a new tradition in scholarship. As James (1972) wrote, all these efforts were "new" in the sense that they offered new hypotheses, new methods of classification, and new ways of making use of mathematical principles of explanation. In the development of this new way of thinking, the ground breaking work was performed by the French scholar Count Buffon (George Louis Leclerc, Comte de Buffon, 1707-1788), who was director of the Jardin du Roi botanical garden in Paris from 1739 to 1788. By virtue of the position that he held, Buffon had access to a large collection of specimens of plants and animals, and to descriptions written by travellers and explorers.

His forty-four volume work on *Histoire Naturelle, Generale et Particuliere* (1749-1804) (written in active collaboration with a large number of scholars) "represents one of the first works resulting from the reports of voyages of discovery in which attention was turned from oddities and marvels to a search for regularities and for the laws governing processes of change. His approach was nonmathematical and ... strictly inductive ... aimed at finding some kind of order in the flood of new inforrnation" obtained from the explorations and discoveries (James, 1972, p. 136).

While Buffon subscribed to the idea of a divinely created earth, he rejected the theory that the final plan of creation was in the mind of the creator and as such there was no need to look for causes of earth phenomena. Buffon was the first to focus attention on man as an agent of geographic change. He developed the idea that the earth has been cooling gradually, and that part of the warmth on the earth surface was derived from its hot interior. Buffon subscribed to the theory of climatic determinism inherited from the ancient Greeks but he was positive that man was not a passive agent, and that he was capable of adjusting to any climate through his technology and culture.

Inclusion of panels of trained scientists in the voyages of discovery, beginning around the last quarter of the seventeenth century, had greatly promoted scientific knowledge about the earth. The first such scientific traveller during 1698-1708 was the English astronomer Edmund Halley (1656-1742), the great scientific genius at deriving order out of complex data. He was the originator of the mortality tables in 1693, as also of many graphic methods for showing geographical distribution of physical features of the earth. His maps and discussions of the trade winds of the Atlantic (1686) provided the first illustration of wind directions and wind shifts. He also prepared the first map of magnetic variations using isogonic lines in 1701.

The father-son team of Johann (1729-1798) and George (1754-1794) Forster had accompanied Captain James Cook on his second voyage to the Indian and the Pacific oceans. In the course of this voyage, the two made botanical observations. It was in the course of this voyage that George Forster found out that the patterns of temperature on the eastern and western margins of landmasses are very different so that there was similarity between the climates of Western Europe and the western coast of North America. George Forster later played the pivotal role in attracting Alexander van Humboldt to geography. Another great scientific traveller of this period was Major James Rennel (1742-1830). He was one of the founders of the science of oceanography, and had served as the Surveyor General of India during 1767-1777. His *Atlas of Bengal* (1779) had gone through several editions, and it had remained a standard work of reference until around 1850.

The growing spirit of inquiry had, by the last quarter of the eighteenth century, egged on many scholars to seek scientific answers to the age old questions regarding man and his life upon the earth. A most prominent name in this regard was that of Thomas Robert Malthus (1760-1834) who published his famous essay on population in 1798 in which he set out his theory about the interdependent relationship between increase in population and food supply. He noted that population increases in geometrical progression whereas food supply grows only in arithmetical progression. As a result, population keeps on increasing until subsistence level is reached, so that its further increase is checked by famine and epidemics. At one place in his essay Malthus had used the phrase "struggle for survival", a term which, several decades later, was to inspire Charles Darwin (1809- 1882) in his explorations toward the theory of evolution of species through the process of natural selection. In his studies Malthus had showed that increase in agricultural production could not cope with the natural increase in population, irrespective of technological inputs. He was also the first to formulate the economic law of diminishing returns from increased employment of capital and labour (Dikshit 2018).

UNIT - 2: CONTRIBUTIONS OF HUMBOLDT AND RITTER IN GEOGRAPHY

INTRODUCTION

After the great age of Discovery, two leading German scholars, viz., Alexander von Humboldt and Carl Ritter made valuable contributions to the fields of basic sciences, humanities and arts. The foundation of geography as a modern science was primarily laid by German scholars during the period from 1750 to 1850. The second half of this period, the time of Humboldt and Ritter, is known as the "classical period of geography". They lived in the same time in the same country – for over 30 years in the same city.

The discovery of each new law of nature leads to the development of some more general laws. This can be aptly said for Humboldt and Ritter who in the late nineteenth century established most of the theoretical concepts of geography and thus have rightly been called as the "*Founding Fathers of Modern Geography*". Based on the available information both of these scholars came up with large number of syntheses. The geography demonstrated by them has been defined as '*classical geography*' by Hartshorne (1976); this is because they dominated the foundation period of the discipline through their uniform and simple methods leading to the end of the classical period.

ALEXANDER VON HUMBOLDT

Alexander von Humboldt (1769-1859) led the way in the expansion of the geography outside and outside Germany, he travelled about 40,000 miles with modern equipment and the people travelled to the Andes south to traversing, explorers reached Lima, Peru-Guyana Bird pieces were seen on the banks of the river, which are very famous prices. In addition, Peru's current cold water was also recorded and recorded for the first time. The temperature and velocity of this current were measured. In March 1803, a campaign was started from Guayaquil to Mexico Port, that is, Acapulco travelled to different parts of Mexico and saw the effects of landscapes on the cultural landscape. The ongoing campaign for a short period in Philadelphia and Washington returned to France in 1804.

The adventurous nature of Humboldt made him travel to Vesuvius Volcano (Italy) in 1806. After completion, he wrote his experiences and comments in thirty volumes in the French language, which was later translated into many foreign languages. It encouraged many young scientists to examine the geography of the unemployed areas of the world. In his writings, Humboldt explained the reason for the prosperity of Mexican residents to better use of land resources. The idea of digging a canal across Isthmus was also put forward by him.

In 1829, Humboldt was invited by the Russian Tsar to the city of Petersburg (Leningrad) and was tasked with finding the virgin land of Siberia across the Ural Mountains. During the Siberian campaign, a regular record of temperature, and pressure was kept. Based on these comments a result was estimated that the temperatures on the same latitude are inward from the coast. It was on his advice that many weather stations were established by Tsar in different parts of Russia. Humboldt also founded the concept of continent, apart from that; the noun of Permafrost was fabricated to explain the frozen characteristics of Siberian soil. After the undertaking that the term '*climatology*' appeared in geographical literature, it is related to all variations of atmosphere, temperature, humidity, barometric pressure, winds, atmospheric purity and degree of visibility.

In 1845, the great work of Humboldt was published, which was accepted all over the world and translated into many foreign languages. Kosmos is a comprehensive account of Humboldt's journey and expeditions; the first volume was published in 1845 and the fifth volume was published in 1862 after his death. In the first volume a general picture of the whole universe is presented. The second volume has a portrayal of nature through the ages since the time of ancient Egyptians; based on the paintings of landscape painters and even poets. In the third volume Humboldt discusses the laws of celestial space. Man has been discussed in the fourth volume; he believed that all the races of man had a common origin and that no race was inferior or superior to other.

With respect to the subject of geography, Humboldt pronounced a word 'cosmography' and divided it into urography and geography. In his opinion, urography is descriptive astronomy, which is related to celestial bodies. Geography on the other hand was limited to geographical geography, which is related to the terrestrial part. Geography, according to him, is the description of the earth (*Erdbeischreibung*) which is related to the interconnection of events in one area together.

Humboldt believed in the inductive method and emphasized the importance of the experimental method of research. He studied comparative studies of various geographical areas, especially steps and deserts. They emphasize the importance of the geographical representation of the data on the usefulness of maps and maps for geographical studies. He believed in the unity of nature and used to accept the idea of underlying casualties (accidental relationship). Humboldt believed that there was a common origin among all the peoples of human race and no race was superior or low for other people. Apart from this, he emphasized the need for a careful measurement of nature's observational observations and observations in the field. This theory was an approach towards creation and modeling. In a nutshell, Humboldt asked for a great variety of specific questions. For example, they attempted to develop a general picture of the distribution of average temperatures in the world in relation to the distribution of continents and oceans. He also tried to heighten height in the tropical areas of plants, animals and human life.

Humboldt concentrated largely, though not exclusively, on physical features, climate and vegetation. About the use of natural he was the follower of Kant. The word ‘natural’ in its broadest sense included all the phenomena observed outside the observer’s mind or the objective reality. It is because of these contributions that he is considered not only the founder of plant geography and climatology but also modern geography. Humboldt thus was the last master of universal science.

CARL RITTER

One of the contemporary scholar of Alexander von Humboldt and distinguished scholar, Carl Ritter (1779-1859) is also known as one of the founders of modern geographical ideas. He has influenced more on the development and growth of geography than Humboldt. His initial work in geography includes two volumes on Europe written between 1804 and 1807. In 1810, he wrote a complete systematic geography of the world entitled "*Handbuch der Physischen Geographie"* but was never published. In this work Ritter expressed hid interest towards inductive method.

For Ritter, geography was an empirical and descriptive science. He declared Geography as the Earth Science, which is related to local conditions and embraces the properties of the place in relation to the temporal, formal and physical characteristics. The first feature was topographical, that is, it is related to the natural division of the Earth's surface. The second included water distribution, sea and atmosphere -the bases of human life. The third conditions were described as the geographical aspect of natural history; it covered the distribution of minerals, plants and animals. Thus, according to Ritter, Geography, it is the branch of science that brings together all its features, events and relations with the world as an independent entity and reflects the relationship of man and man with this unified 'whole' is. He claimed that the central theory of geography is 'the relationship of all events and the nature of nature to humankind'. He makes rhetoric of geography claiming that the Earth goes beyond the real purpose in the form of science, that is, the earth is described as a human home. It is in Erdkunde (published in 1817) that he provides the above explanation about geography; making his geography anthropocentric. In the words of Tatham (1967) Ritter's man-oriented geography clearly states "*the earth and its inhabitants stand in the closest reciprocal relation and one cannot truly present in all its relationships without the other. History and geography must always remain inseparable. Land affects the inhabitants and the inhabitants the land*".

Ritter was also the founder of comparative method in regional geography. Thus according to Ritter, geography is that branch of science which deals with the globe in all its features, phenomena and relations as an independent unit and shows the connection of this unified 'whole' with man and with man's creator. He claimed that the central principle of geography is "the relation of all phenomena and forms of nature to the human race." He makes the theoretical claim that geography as the science of earth (in the sense of natural philosophy of cosmology) reaches far beyond the real objectives, namely, the description of the earth as the home of man. Ritter was the first great opponent of what may properly be called "armchair geography".

The principle of unity in diversity was the fundamental principle developed by Ritter. According to him, there is a fundamental unity in the biological and uninterrupted constituents of the habitat in which man gives sculpture to his cultural environment. In this approach, all the physical and cultural components of the environment are kept in mind and their interconnection is understood to understand the geography of an interstate unit. It is a regional view that unity in diversity means that every naturally bound area is unity in relation to climate, production, culture, population and history.

Ritter was a *teleologist*; his geography seemed to be an attempt to interpret philosophically that which science could not explain. He even sought to construct 'a law of the arrangement' of the parts of each continent to the whole. In short, Ritter gave a systematic frame to geography which remained valid for the future scholars for number of centuries. With the death of Humboldt and Ritter in 1859 the classical period of geography came to an end.

The major geographical concepts of Ritter may be summed up as follows:

1. Ritter conceived geography as an empirical science rather than one based on deduction from rational principles or apriority theory.
2. There is coherence in the spatial arrangement of terrestrial phenomena. Areal phenomena are so interrelated as to give rise to the uniqueness of the areas as individual units.
3. Boundary-lines whether wet or dry (such as rivers or mountains), were instruments for understanding the real purpose of geography which is understanding the content of areas.
4. According to Ritter, geography was concerned with objects on the earth as they exist together in an area. He studied areas synthetically, i.e., in their totality.
5. Ritter holds a holistic view with respect to the content and purpose of geographic study, and the whole study was focused on and culminated in man.
6. He believed that the earth was an organism made, even in its smallest details, with divine intent, to fit the needs of man to perfection. He was a teleologist in his approach.

DIE ERDKUNDES

Ritter's monumental work is entitled as Erdkunde. Erdkunde is a comprehensive German word which stands for science of the earth in relation to nature and history. Ritter remarks that "the earth and its inhabitants stand in the closest reciprocal relations and one cannot be truly presented in all its relationships without the other. Hence History and geography must always remain inseparable." Land influences the inhabitants and in turn the inhabitants transform the landscape. In Europe for example, only in the east (Russia), there was uniformity of history. But in the West, there was variety of environment and history, and in the diverse south (Europe) too history was rich, studded with the efforts and achievements of Egyptians, Carthageans, Greeks, Romans, Gauls and Iberians. In Erdkunde he advanced the theory of the north-west movement of civilization in Europe.

The first two volumes of the Erdkunde were intended to be followed by a study of history. Between 1817 and 1859, he completed 19 volumes of Erdkunde but these volumes cover only Africa and parts of Asia. In spite of the fact that he lived long, he was not able to finish his work of Europe.
Even during the period of Humboldt and Ritter geography was still not related to a specific discipline. In fact geography remained an umbrella concept for a variety of expeditions and other activities within the natural and social sciences, to a large extent supported by geographical societies.

The contributions of Humboldt and Ritter to modern geography are therefore great. They molded the substance of geography into scientific form. Their scientific organization of knowledge was done in two stages. The first stage was the assembly of accurate facts- by observation and measurement. The second stage consisted of giving the material coherence and making it intelligible by considering it under a number of laws to show the relationship of cause and effect to be found in the phenomena in a simple and concise way. Thus both Humboldt and Ritter emphasized the importance of the empirical method of research. In a letter to Blumenbach in 1795, Humboldt is reported to have stressed the importance of the empirical method of science as one best grounded and most likely to succeed. For geography to be considered a science it must establish laws and this is- what Humboldt and Ritter strove to do in their major works. Ritter put the position clearer when he said that he had demonstrated that geography had a right to be considered a sharply defined science, of kindred dignity with the others. Geography must go on to know the causes of things, 'rerum cognoscere causes".

Ritter and especially Humboldt showed the value of the comparative met.hod. Humboldt clearly showed this in the handling of his multitudinous observations. His essay' on the steppes and deserts is full of comparisons between steppes and oceans and between all the steppes of the world. The idea was revealed in his studies of the heaths of central Europe and Llanos of South America. By these comparisons he sought to reveal the characteristics and physiognomy of each study as determined by diversity' of soil, climate and altitude, He compared new observations to previous ones of similar kind and recorded the differences and similarities. Ritter observed that causal relationships were to be sought by the comparative method, the geography that went beyond mere description was comparative- This can be seen in the Erdkunde where the method was applied in the study of regions to distinguish natural regions. Thus Humboldt and Ritter gave geography a strong foothold among the natural sciences, offering more to fire the imagination, to attract the spirit of scientific inquiry.

UNIT – 3: SOCIAL DARWINISM AND ITS IMPORTANCE IN GEOGRAPHY

A DARWINIAN BEGINNING

Whoever says "Social Darwinism" says "Survival of the fittest, in society as in nature." There are good reasons to regret that the term ever acquired this meaning, and good historical studies to learn from in understanding how and why it nevertheless became the entrenched meaning. But there is also an oft-mentioned bad reason for regret, linked to bad history. Social Darwinism, it is sometimes said, has nothing whatever to do with Charles Darwin (1809-1882) or with the body of time-tested facts, theories and practices deriving from Darwin's work. On this view, Darwinism is science, and social Darwinism a pseudo-science masquerading as an application of Darwinism.

To see why such a separation is untenable, we need only consider an exchange of letters from Darwinism's beginning. In October 1859, while Darwin was awaiting publication of On the Origin of Species by Means of Natural Selection, he corresponded about it with the geologist Charles Lyell, who had read the book in proof. Lyell's verdict meant a great deal to Darwin. There was no one whose judgment Darwin esteemed more highly. But there was also no one more fully committed to an idea that the Origin aimed to discredit. For Lyell, species were independent creations; each brought into being by the "Author of Nature" (Lyell's phrase) according to a divine plan, and designed to thrive for a particular period in a designated region of the Earth. By contrast, Darwin in the Origin explained the origin of species by appeal only to ordinary, continuously operating, law governed natural processes, which anyone could see in action now – chiefly, processes of variation (when individuals of a species are born a little different from the others), inheritance (when those individual differences are passed on from parents to offspring), and competitive struggle (because typically life is hard, between the scarcity of food and an abundance of predators, diseases and so on).

SOCIAL DARWINISM: THE CONCEPT

The term *Darwinism* refers most centrally to the theory of natural selection, according to which only the fittest species in organic nature survive, whereas the unfit become extinct. The extension of these ideas to social thought is known as Social Darwinism. The application of models of evolution to human societies long preceded the publication of Charles Darwin's *Origin of Species* in 1859, however. Already in the eighteenth century, historians influenced by the Scottish Enlightenment—including William Robertson and Adam Smith—had constructed a universal vision of history in which all societies advanced through four stages (from hunter gathering to commercial society) as they progressed from "rudeness to refinement." This theory of development by stages influenced European notions of progress and of civilization among non Europeans: peoples engaged in trade were held to be superior to those who relied exclusively on agriculture while the latter, in turn, were considered more advanced than subsistence hunter gatherers.

In the early nineteenth century, the notion that world history and human society proceeded in evolutionary stages was purveyed in the works of Auguste Comte, G. W. F. Hegel, and Karl Marx, each of whom searched for general laws that underpinned social change. Unlike later theorists, these earlier political writers had a universal outlook that did not exclude non European peoples from following the road already taken by European nations. By the time of Herbert Spencer (1820–1903), however, this optimism had given way to a bleaker, Malthusian conception of competition between human beings for the scarce resources required for subsistence. In the late nineteenth century, this notion became linked directly to imperialism. It provided a framework for understanding the rise and decline of nations and enlivened competition among European nations.

Spencer—who coined the term *survival of the fittest* several years before Darwin set forth his theory— developed an all encompassing conception of human society and relations based on evolutionary principles. His conviction that a general law for all processes of the earth could be formulated led him to apply the biologic scheme of evolution to society. The principles of social change must be the same, he supposed, as those of the universe at large. Although Spencer clung to outdate scientific ideas, such as Jean Baptiste Lamarck's debunked thesis concerning the inheritance of acquired characteristics, it would be inaccurate to argue that he corrupted Darwin's pristine scientific ideas. Many of Darwin's ideas emerged from the social context in which he lived. As Marx noted, "it is remarkable how Darwin recognizes in beasts and plants his English society with its divisions of labor, competition, [and] opening up of new markets".

Spencer's ideas about selection also were born from his political beliefs: He repudiated government interference with the "natural," unimpeded growth of society. He maintained that society was evolving toward increasing freedom for individuals and so held that government intervention should be kept to a minimum. This belief led him to oppose all state aid to the poor, a group he maintained were unfit and should be eliminated. Spencer viewed state intervention to ameliorate their condition as the "artificial preservation of those least able to take care of themselves." As Spencer wrote, "the whole effort of Nature is to get rid of such, to clear the world of them, and make room for better".

Darwinism was put at the service of imperialism, as a new instrument in the hands of theorists of race and civilization struggle. Competition with other European states urged the securing of colonies to prevent raw material, land, and potential markets from being seized by rapacious rivals. In Theodore Roosevelt's "The Strenuous Life" (1899), the future American president warned against the possibility of elimination in an international struggle for existence. America, he said, could not shrink from "hard contests" for empire or else the "bolder and stronger peoples will pass us by, and will win for themselves the domination of the world". Successful imperial ventures thus were perceived to indicate the vitality, and hence "fitness," of a nation.

Social Darwinism also proved to be a justification for the subjugation of non-European peoples, who were deemed less "fit" than Europeans. Nature, theorists argued, intended the rule of superior European nations over inferior colonial races. Racial arguments permeated the language of adherents of Social Darwinism as well. The French political leader Jules Ferry (1832–1893) explicitly argued that "the superior races have rights over the inferior races". After World War I, the mandate and trusteeship system set up by the victors over much of the colonized world utilized arguments that derived from Social Darwinism. In 1922 Baron F. D. Lugard argued that the British Empire had a "dual mandate" in tropical dependencies "unsuited for white settlement," calling for the "advancement of the subject races" and "the development of [the territories'] material resources for the benefit of mankind." He insisted that indigenous populations were benefiting from "the influx of manufactured goods and the substitution of law and order for the methods of barbarism". Social Darwinism thus lent a pseudoscientific veneer to colonial subjugation and bolstered the alleged civilizing mission of Europeans to non-Europeans. The most extreme form of Social Darwinism was eugenics.

Proponents of eugenics claimed that particular racial or social groups were naturally superior, and sought the enactment of laws that would control human heredity by forbidding marriage between people of different races and restricting the reproductive activities of people they considered unworthy, such as criminals and the mentally ill. In the late 1920s and 1930s, Nazis drew on such extreme precepts of Social Darwinism in their attempt to create an idealized Aryan race, an effort that culminated in the Holocaust and the brutal deaths of millions of Jews, Roma (gypsies), and members of other groups considered inferior by the Nazis.

SOCIAL DARWINISM AND THE POOR

The impact of British biologist Charles Darwin's Origin of Species (1859), The Descent of Man (1871), and other writings went well beyond the audience of natural scientists to whom it was addressed. Throughout the western world, journalists, academics, and social reformers were quick to appropriate Darwin's theories about the evolution of life forms to explain trends in social and economic life. Under the circumstances, this is not surprising. The world was in the midst of vast and frightening changes — industrialization, urbanization, immigration, class war, and mass poverty — which no one understood and to which no one could offer solutions. Extrapolations from Darwinism, with its emphasis on evolutionary progress, offered reason for hope that a new and better social order could emerge from the turbulence. At the same time, by highlighting competition and the survival of the fittest as the drivers of evolution, it seemed to explain both the emergence of the fittest — fabulously wealthy elites and giant corporations, as well as the unfit — the masses of poor in the teeming city slums.

Social Darwinism, as it came to be known, served the purposes of both liberals and conservatives. Because conservatives believed that many of the traits associated with unfitness — propensities for idleness, criminality, sexual misbehavior, and alcoholism — were passed along from generation to generation by heredity, much like hair and eye color, they grimly predicted the growth of a permanent criminal underclass unless steps were taken to prevent it. They were particularly concerned with the impact of sentimental and impulsive charity on the poor. Spontaneous responses to suffering attracted impostors and vagrants from every direction to enjoy the public benefaction," drawing to the cities the floating vagrants, beggars, and paupers, who wander form village to village throughout the State. The streets of New York became thronged with this ragged, needy crowd; they filled all the station-houses and lodging places provided by private charity, and overflowed into the island almshouses. Street-begging, to the point of importunity, became a custom. Ladies were robbed, even on their own doorsteps, by these mendicants. Petty offenses, such as thieving and drunkenness, increased. One of the free lodgings in the upper part of the city, established by the Commissioners of Charities, became a public nuisance from its rowdyism and criminality.

Poor relief, conservatives believed, destroyed the work ethic that motivated the poor to work. "The public example of alms induces many to be paupers who were never so before, while they do not at all relieve the truly deserving, who hesitate to be exposed to such publicity. They are, in fact, an especial assistance to the idle, and a reward to the improvident. Preventing the growth of this criminal class called for strict measures, beginning with a thorough and discriminating supervision of all charities, public and private; the most careful attention to the education and employment of the poor and their children; the placing of pauper children in good families, at a distance, if possible, form degrading associations; a rigid and exact system of in-door relief, accompanied with labor; the reduction of out-door relief in cities, and the encouragement of emigration to rural districts from the crowded centres of poverty and crime, which most of our largest cities have now become. The position of New York in this respect is exceptional, because it yearly receives a quarter of a million immigrants from foreign countries, and this exposes it to peculiar evils and dangers. While this should be borne in mind, it should not be made an apology for neglect or an occasion for abuses, but should lead to increased vigilance and activity on the part of magistrates and citizens.

In a word, conservatives (then as now) not only blamed the poor for their poverty, but also the dispensers of "indiscriminate" and "sentimental" charity, whose well-intentioned, but ill-informed benevolence served both to perpetuate the sufferings that they sought to ameliorate and to compound them by encouraging the survival of the unfit. By the turn of the nineteenth century, the most extreme of the conservatives, combining ideas drawn from Darwin, with those of his contemporary Francis Galton, produced theories which urged actions to prevent the disabled and other "unfit" people from perpetuating their kind by segregating them from society in almshouses, asylums, and other congregate institutions and through sterilization. These practices were enacted into law by many states and were upheld by the U.S. Supreme Court, with Justice Holmes memorably defending government's right to incarcerate and sterilize by declaring "three generations of imbeciles is enough!"

Social Darwinism never constituted a formally articulated philosophy; it was used in a variety of often contradictory ways by writers and thinkers of the late nineteenth and early twentieth centuries. Regardless of the social and political agendas it gave rise to, the one thing all had in common was a scientific data-based approach to defining and offering solutions to social problems. Whether used to justify laissez-faire or activist public policies, social Darwinism provided a vocabulary and set of concepts that facilitated the emergence of the social sciences and their application to such pressing problems as poverty and social justice.

MISUSING DARWIN'S THEORY

Darwin's theory of evolution by natural selection is entirely focused on an explanation of life's biological diversity. It is a scientific theory meant to explain observations about species. Yet some have used the theory to justify a particular view of human social, political, or economic conditions. All such ideas have one fundamental flaw: They use a purely scientific theory for a completely unscientific purpose. In doing so they misrepresented and misappropriated Darwin's original ideas. One such distortion and misuse is the loose collection of ideologies grouped under the label of "Social Darwinism." Based largely on notions of competition and natural selection, Social Darwinist theories generally hold that the powerful in society are innately better than the weak and that success is proof of their superiority.

Darwin passionately opposed social injustice and oppression. He would have been dismayed to see the events of generations to come: his name attached to opposing ideologies from Marxism to unbridled capitalism, and to policies from ethnic cleansing to forced sterilization. Whether used to rationalize social inequality, racism, or eugenics, so-called Social Darwinist theories are a gross misreading of the ideas first described in the *Origin of Species* and applied in modern biology.
Social Darwinism declined during the 20th century as an expanded knowledge of biological, social, and cultural phenomena undermined, rather than supported, its basic tenets.

MORPHOLOGY OF CULTURAL LANDSCAPE (CARL O. SAUER)

CARL O. SAUER

Carl O. Sauer was one of the leading American geographers of the 20th century. For most of his academic life he was associated with the University of California, Berkeley. During his graduation course he attended lectures by Ellen Sample. He however, rejected the environmental determinism of Ratzel and Miss Semple. Most of his survey was done in Latin America and the less industrialized parts of U.S.A.

He contributed mainly in the field of origin of agriculture, the diffusion of plants and animals, and the impact of conquest upon indigenous American societies (Red Indians) The focus of his study was on the process leading to landscape change up to the present, beginning at the pre-human stage of occupance. Thus, the human geographers should make cultural processesthe base of their thinking and observation. His presentation in the symposium of Princeton (1955) on man's role in Changing the Face of the earth which was conceived by Thomas was highly appreciated. He "specifically insisted upon the use of the historical method in geography." In his opinion the present cultural landscape are to be studied in terms of their development from the original natural landscapes'. He also opined that "geography dissociates itself from geology at the point of the introduction of man into the aerial scene."

BACKGROUND: "THE MORPHOLOGY OF LANDSCAPE"

Landscape has been defined by different thinkers as a visible territory, as the visible part of the environment or as an individual's perception of the environment through his senses. It is the external environment, natural and/or anthropic, which may be directly perceived or experienced by an individual who observes or senses a portion of a larger physical environment. The landscape is an area or territorial unity, more or less well defined, which varies depending on the watcher and his position within the landscape itself, but, above all, on the representational context he shares with members of his own cultural milieu. The European Landscape Convention would define the concept merely as an area perceived by people. "Landscape means an area, as perceived by people, whose character is the result of the action and interaction of natural and/ or human factors".

This new interpretation of landscape is propitiating the creation of a new territorial culture. In fact, territory is starting to be considered and experienced in terms of landscape: as the scenic background within which people's lives evolve. This represents a new way of understanding the socio-physical environment in which people live. Territory and landscape have changed into correlative concepts while a population's territorial culture is measured by how it evaluates its own landscapes. This implies an interpretation of territory which incites conservation and sustainable management measures applied to valuable areas, indicating a much needed change of direction in our lifestyles, and encourages new more prudent and imaginative ways of dealing with our surroundings. It calls for an existential connection with the environment, a new outlook on Nature. This attitude depends on the territory's sustainability as a whole and as the consequence of its inhabitant's wellbeing and quality of life.

Throughout much of his professional life, Carl Sauer was at odds with mainstream American geography. While still in graduate school (1909-1915) he began to develop critiques of major currents within the mainstream. These departures or dissents were well informed and grounded in the philosophy and history of geography. Initially these came from close readings of the German and French geographical literature. Save for his advisor Rollin Salisbury, he was not impressed with the geography faculty at the University of Chicago, and even less taken by the cause-and-effect geographical determinism then enjoying paradigmatic dominance. He later remarked that by 1912 he had begun to distance himself from the Chicago program and spent his evenings reading the continental literature (Sauer, 1999b). It was not until after 1923, when he moved from the University of Michigan to Berkeley, that he began to publish critiques of mainstream currents and propose alternative approaches. His major statement during this period, "The Morphology of Landscape;' was published in 1925.

THE LANDSCAPE TURN

One of the main transformations in our contemporary world has been the emergence of ecological awareness; today's worldwide preoccupation with the environment. In the 70s a new idea emerges, claiming that a delicate balance of nature ─essential for our survival ─ is to be restored and sustained only by a global effort which involves each and every one of us. By reinforcing ecological awareness the learned concept of landscape becomes common among people who indistinctly discuss it in relation to location. People gradually become aware of how landscape impregnates their lives due to having been born and raised in its bosom, not knowing to what extent they are thus conditioned. They assume that throughout history the relationship between humans and their landscape is a constant flux of reciprocal energy, sometimes enriching and others degrading. Landscape thus becomes an essential part of their culture. They start considering it a determinant factor in the configuration of their own society, for they accept that it contains the deepest roots of the structure which defines society itself. "Landscape analysis, understood as the result of social practices, as a social construction, enables us to expose man's action throughout time and recognise aspects of our history within our current landscape. As a testimony of human action and the ways of life which have shaped it, landscape is intrinsic to cultural identity, and as it preserves our civilization's traces and remnants, it is a heritage of great value to be respected".

TERRITORY AND LANDSCAPE

Human beings constantly interact with their environment. This affects their senses, emotions and community relationships. Landscape has been defined as the invisible context of our lives, of which we are not wholly aware. Territory, as a physical entity, refers to the earth's surface which we consider as our own, and to which humans have assigned different functions throughout time. It alludes to the relationship between members of a given social group and their physical surroundings; the time-space location where members of the group live. Therefore, it is a space defined, produced and ordered, according to human actions which adhere to values projected upon the landscape by individuals themselves. Territory has been assessed in different periods and among different societies in multiple and creative ways: surface of the earth, resource, habitat, frontier, limit, foundation and background for vegetable and animal species, refuge, an area common to a human groups, socio-physical construction, playground, etc. The way in which the human species structures space and acts upon it is conditioned by the way in which he perceives and experiences it. The ability to adapt to a given environment depends on individuals' creative awareness, which serves as a foundation for their ability to decide and provide meaning for the elements to be found within a given territory.

Territory is conceived as a setting within which human life is framed and implies the existence of human subjects who project a given meaning upon it, which is fostered within a specific time and culture. The multiple meanings that men award to territory can be understood as one of the landscape's category. "The landscape is emerging as a very powerful concept to convey the relations between society and its territory, in two ways: on the one hand, complex, dynamic and changing interactions between a society and its territory, that is, the social and economic processes which shape the territory; on the other, the representations and images that the mentioned society attaches to its territory, that is, the social and cultural assessment of territory".

The landscape concept has gone through a period of swift and deep transformation until it has acquired high praise. Two factors have been essential for the emergence of this new identification of territory and landscape. On the one hand, radical changes that are taking place on a world scale and their consequences may endanger humanity's future. Environmental awareness is making us notice that the unprecedented worldwide changes that are taking place may jeopardize life on Earth. This threat is inducing reflection on the validity of our current territory models due to the environmental crisis our lifestyles have generated, founded on beliefs and values whose risks and undesirable consequences for humans are unforeseeable. It is true that within this controversial issue, myth and reality coexist, that we have scarce knowledge of the planet, that we do not possess a reliable and complete scientific model and that apocalyptic fear is well embraced in our consumer society.

CULTURAL LANDSCAPE

Environmental awareness has become one of the major motors of contemporary thought and social action. To its development has contributed the anthropological perspective which has always questioned the limits between humanity and nature, to expose the deep imbrications existing between culture and ecosystems. In order to explain human conduct, the Social Sciences have used three strictly intertwined categories or levels of analysis: person, society and culture. But we must acknowledge that the physical environment is also essential. Territory as cultural environment is an element on which human nature is based, as complexity science has shown. The latter has proven that a system isolated from its environmental conditions is unconceivable, and, consequently, human beings can not be understood as isolated from an ecosystem's interactions also shaping his nature. We can not understand people's individual and social life if we disregard the environment in which they are immersed, that is, socially and culturally constructed spaces inhabited by them. Thus, a genuinely anthropological approach must analyze human intention and action in the context of permanent and mutually conditioning interaction between people and their socio-physical environment.

Landscape's plural meanings, its different scales and the diversity of goals established by landscaping projects, explain the very open character of landscape methodology analysis and the variety of instruments, explicitly or implicitly related to landscape, destined to the defence of certain values and the order of their dynamics and transformations. The richness of possibilities any given landscape offers, either subjective or objective, are to be found in different treatments and methods identified with each professional collective, which has propitiated specified definitions and varied indicators to establish the essence of a given landscape. Based on these several viewpoints, landscape acquires particular values and self contained meanings. Thus, we may remember the meanings provided by art, philosophy, science, myth, the anthropologist's cultural references and interpretations, the identification and uses social actors make of it, etc.

The term "Cultural Landscape" is already officially recognized, being subject to growing scientific interest, while there is talk about the demand for landscape. But even today the Cultural Landscape category is an uncommon term and is even perceived as an opaque concept by most people. Many identify cultural landscape with a given historical setting, a geographical area, to sum up, an historical landscape whose fundamental components are aesthetic and cultural. They view cultural landscape as a trace of human activity within a territory, but reduced to a mere fossil of great value that should be preserved and protected. However, cultural landscape is a much richer concept than is commonly suggested. It contains and symbolizes a huge amount of meanings and values, and thus is used as the fertile basis on which to develop theoretical digression. This is due to the fact that it serves as a tool to look into ordinary landscapes, those within which people dwell.

The moment is ripe, for the way in which people perceive cultural landscapes is being shaped by the European Landscape Convention's decisive contributions. This is an innovative measure in contrast to other documents on the natural and cultural heritage because it refers both to landscapes that might be regarded relevant from an historical viewpoint ─singular elements of exceptional character─, and everyday cases, associated with quality or deterioration. "This new concept expresses, on the contrary, the wish to face, in a global and frontal way, the matter of quality in places where people live, recognized as an essential condition for individual and social wellbeing (in the physical, physiological, psychological and intellectual sense), for a sustainable development and as a resource which favours economic activity".

The cultural landscape category may be a solid foundation on which to build a model of progress that current society needs. It links us to the people's real economy and not to financial-speculative fictions generated by successive crises ─stimulated by the capitalist system itself. This key concept provides ideas and principles upon which to rationally and intelligently deal with local development, basis for economic and social progress. It is an essential concept to guide our environmental perception and land-use planning, as well as for the protection and management of cultural and natural heritage. In building land, planning based on cultural landscape criteria will help guide urban growth, for the formal and functional integration of new landscapes. In undeveloped land it serves to properly intervene, strengthening or creatively transforming existing landscape units.

Cultural landscapes are a testimony and legacy that our forefathers bequeathed us and that we ourselves are obliged to respect and recreate. Their conservation is a matter of self-respect as well as a form of embracing the culture they represent. Landscape is a society's cultural projection upon a given space. It is therefore one of the most exceptional elements of identity we possess and, in consequence, a form of cultural heritage. Hence, the convenience in strengthening and developing a territorial awareness, not just as a source of material gain, but as a way of awakening people's desire to know and enjoy their surroundings. "This raising awareness makes the individual go from being a passive object subdued by external processes and structures, to become an agent of change, an active agent of resistance, in the defense of the planet and of life, a possible change".

LANDSCAPES OF DESTRUCTION

If this refocusing of critical landscape studies on landscapes of destruction is to be taken seriously, then it must also take seriously its precursors. In nineteenth-century and early twentieth-century geography, the work of George Perkins Marsh (1864), Elisee Reclus (1905-1908), Jean Bruhnes (1910), and a number of other chroniclers of imperial and colonial "destructive exploitation" (to use the apt term of the times) provides the foundations for any contemporary study of landscapes of destruction. If the focal point is to be grounded in cultural landscape study, then Carl Sauer and his work indisputably need to be the starting point. Surveying his career, starting with his dissertation in 1915 and continuing until his death in 1975 (some sixty years of fieldwork, research, and publication), one will find that, for Sauer, landscape construction and destruction were central organizing concepts and were often conjoined to produce powerful and sometimes polemical critiques of European colonial expansion from the late middle ages onward.

CONCLUSION

An adequate account has not yet been written of Sauer's contributions to American geography as realized and expressed through the collective production and directions taken by his associates, students, and those inspired by the Berkeley school approach. If and when it is, it will encompass the work of several generations of scholars, whose numbers now total several hundred at a minimum. Sauer's direct progeny within the Latin American branch of his "academic genealogical" tree (his advisees and their advisers, et cetera, that wrote dissertations on Latin American topics) numbered over 150 by 2000.

Sauer's works and those of some of his students have been or are currently being translated into Spanish and Portugese and a whole new generation of Latin American geographers and students are being introduced to the Berkeley school, many for the first time.9 Nor has anyone yet attempted to tabulate the published contributions of Sauer, along with his associates and several generations of his legatees, to the multiple questions that have engaged this large group of like-minded scholars. It can be saying that, however, that this literature comprises some hundreds of books and monographs, and several thousand articles and lesser publications.

Viewed collectively, this corpus amounts to one of the larger bodies of published work in North American geography. It has not previously been implicitly recognized or acknowledged in these terms. It is also safe to say, despite retorts to the contrary, that Sauer's legacy is alive and well and is likely to persist as long as geographers and kindred scholars continue to take an interest in questions of culture and landscape and the history of humans' agency on earth.10 The quality and quantity of the Sauerian oeuvre taken as a whole-or even in parts-is large and complex enough to ensure both continuing criticism and enduring admiration. For example, Don Mitchell's 2003 call for landscape scholars to put questions of landscape destruction front and center provides not only a fulcrum to redirect cultural geography, but also an appropriate lens to reassess Sauer's and his adherents' contributions. Mitchell's call could be the grounds for a critical survey of past thought, in light of the future. And, this is not likely to be the last opportunity either.

UNIT-4: MAJOR PARADIGMS IN GEOGRAPHY AND THEIR SHIFT

Although 'the appearance of a "new geography" has been loudly and persistently proclaimed by almost every generation of geographers since ancient times', or so Preston James (1972,505) tells **us,** the last decade has witnessed the novelty of geographers seeking support for their claims from studies in the history and philosophy of science. Thomas Kuhn's *The Structure* of *Scientific Revolutions* has received special mention. Despite the frequent references to Kuhn's work, however, fundamental questions about the implications of his ideas for geography remain unanswered. One of these questions is prompted by the repeated appeals to Kuhn's scheme as support for programmes of reform in the discipline; a second arises from the lopsided treatment of his work in which his concept of 'normal science', the necessary corollary to scientific revolution, has been totally neglected; a third comes from the lack of any serious attempt to analyse the history of geography in terms of his scheme.

Has the dominantly prescriptive use of the paradigm concept in geographical writing been consistent with Kuhn's use of the term? What are the implications for geography of Kuhn's view of 'normal science'? Has the kind of scientific revolution envisaged by Kuhn really occurred in geography?

THE PARADIGM PANACEA

At one point in *The Structure* of *Scientific Revolutions* (p. 160) Kuhn mused that Probably questions like the following are really being asked: Why does my field fail to move ahead in the way that, say, physics does? What changes in technique or method or ideology would enable it to do so? Geographers certainly were and still are asking this question, and have seen in Kuhn's work a justification for their programmes for reform. Kuhn's ideas have apparently been welcomed *not* because of their value as analytical tools which might sharpen our knowledge of the development of geography, but rather because they have allowed recent trends in geography to be likened *to* phases of rapid advance in other, more scientifically respectable disciplines.

Rapid progress can be **(or** has been) made, we have been repeatedly told, by the adoption of some new paradigm or exemplar, By implication much traditional geographic work has seemed pre-scientific. Nonetheless, Haggett and Chorley (1967) have assured us, 'Geography, coming late to the paradigm race, has the compensating advantage that it can study at leisure the "take-off" paradigms of other sciences'. What geography has been transformed by (or still needs) is a new paradigm such as model building (Haggett and Chorley, 1967), or systems analysis (Chorley, 1971, 1973; Chorley and Kennedy, 1971), or a 'massive shift from descriptive geography towards more analytic work in which mathematical models of how regions grow and interact play a dominant role' (Haggett, 1970, 20), or even a 'process meta geography that leads to a paradigm of locational and environmental decision making in complex systems' (Berry, 1973). Nor have the exhortations been limited to academic circles: the secondary school teacher of geography has been told both that he 'must become aware of the paradigms he has internalised as a result of his educational experiences', and that his teaching will be greatly improved by 'selecting a (geographical) paradigm to achieve educational aims' (Biddle, 1976).

My concern here is not with an evaluation of these programmes for reform - they all may be laudable - but with the impression that they give of deriving support from Kuhn's study of the development of science. I will argue that the prescriptive use of his concepts in geographical writings has been unjustified (see also Guelke, 1971). To use a concept as an analytical tool is one thir.g; to proclaim it as a panacea is quite another. In Kuhn's own words: -

. . . the proto-sciences, like the arts and philosophy, lack some element which, in the mature sciences, permits the more obvious forms of progress. It is not, however, anything that a methodological prescription can provide . . . I claim no therapy to assist the transformation of a proto-science to a science, nor do I suppose that anything of the sort is to be had. If , . . some social scientists take from me the view that they can improve the status of their field by *first legislating on fundamentals and then turning to puzzle solving, they are badly misconstruing my point.* . . As in individual development, so in the scientific group, maturity comes most surely to those who know how to wait. (Kuhn, 1970b, 244-45, italics added; see also Kuhn, 1962a, especially 160-61; 1962b, 190).

Indeed the suggestion that a disciplinary matrix can be changed by such legislation is the very antithesis of Kuhn's argument. In any discipline, Kuhn argued, there are initially no clear cut rules of research, and fact gathering is a fairly haphazard activity usually restricted to data that lie readily to hand; though the practitioners are scientists, the result of their labour is 'something less than science' (Kuhn, 1962a, 13**).** Research, he continued, is transformed by an achievement sufficiently unprecedented to attract adherents from competing modes of work, yet sufficiently open-ended to leave a wide range of problems to be solved (see also Masterman, 1970). In his view of science shared examples serve the cognitive functions generally attributed to shared rules, and he asserts that such knowledge 'develops differently from the way it does when governed by rules' (Kuhn, 1974,482). The role of exemplars is not just a matter of practice in applying rules, it is rather a matter of acquiring the ability to see resemblances, of perceiving similarities that are 'both logically and psychologically prior' to any rules (Kuhn, 1974, 472).

According to Kuhn, scientific progress is not gradual, orderly, or necessarily cumulative, but rather **is** characterised by truly revolutionary changes as one research tradition or programme replaces another. Nonetheless, he specifically warns against the trap of seeing a paradigm as 'a quasi-mystical entity or property which, like charisma, transforms those infected by it' (Kuhn, 1974, 460-61). Yet this very trap seems to have ensnared some seeking new directions for geography.

NORMAL GEOGRAPHY

If indeed there have been revolutions in geography there must also, according to Kuhn, have been periods of 'normal' science; the dominant role of shared examples ensures this. In fact the novelty and significance of Kuhn's work lie far less in his treatment of revolutionary; phases than they do in his comments on the periods between major advances. There has, however, been an almost total neglect of the concept of 'normal' science in geographical writing. Apart from assurances that all will be well, we have been told virtually nothing of what will happen after a Kuhnian revolution in geography. Normal science, as Kuhn, sees it, has three main aspects: it includes the pursuit of facts that the discipline matrix shows to be worthwhile, the comparison of factual data with predictions, and the articulation of the disciplinary matrix by the resolving of ambiguities. Such activities entail the solving of puzzles, and not the testing of the essential assumptions of a disciplinary matrix. Kuhn is adamant here, insisting that while the striving to bring theory and fact into closer agreement may be seen as testing or as searching for confirmation, its object is the solution of a problem 'for whose very existence the validity of the paradigm (read disciplinary matrix) must be assumed' (Kuhn, 1962a, 80). And he is by no means simply repeating the old contention that selection, evaluation and criticism are guided by theory; he goes much further, claiming that **'**no process yet disclosed by the historical study of scientific development at all resembles the methodological stereotype of falsification by direct comparison with nature' (Kuhn, 1962a, 77).

Challenges to a disciplinary matrix, Kuhn (1962a, 78-9) argues, may go unnoticed, while the discovery of anomaly leads not to an immediate rejection of prevailing ideas but rather to their defence by the formulation of ad hoc hypotheses. In fact, he continues, a new paradigm or change in the matrix may initially offer no solution to questions successfully answered by the prevailing one. One must 'have faith that a new paradigm will succeed with many large problems that confront it, knowing only that the older paradigm has failed with a few. A decision of that kind can only be made on faith' (Kuhn, 1962a. 158). Faith, he emphasises, is necessary because of the 'incommensurability' of competing views. The proponents of each are always at least slightly at cross-purposes. Neither side will grant all the non empirical assumptions that the other needs in order to make its case . . . the competition between paradigms is not the sort of battle that can be resolved by proofs (Kuhn, 1962a, 148).

Briefly then this is the nature of Kuhnian normal science and it is a far cry from the delightfully rational rendition of paradigm-directed research and cost benefit weighting of competing paradigms (Haggett, 1977) presented in geographical writings. Even if we accept it as a true picture of our discipline, our sailing may be none too smooth because claims that the rejection / acceptance of paradigms is generally not subject to critical assessment and is always based on more than a comparison of theory to fact, or that research problems are made meaningful only by paradigms whose validity must be assumed, have prompted many rejoinders. Shapere (1964) comments that '. . . anything that allows science to accomplish anything can be part of (or somehow involved in) a paradigm', while Suppe (1974) suggests that the concept has become 'bloated to the point **of** being a philosophical analogue to phlogiston'.

Scheffler (1967) argues that in such a scheme 'independent and public controls are no more, communication has failed, the common universe of things is a delusion, reality itself is made by the scientist rather than discovered by him'. Berkson (1974) claims that belief in the necessity of suspension **of** criticism after the adoption of a paradigm can cripple fruitful critical activity, and also points out those significant contributions have been made by those who have no faith in the ruling paradigm. Popper (1968) maintains that the hallmark of scientific statements is that they are testable, not that they are grouped around a central paradigm. 'It is just a dogma (he insists) . . . a dangerous dogma . . . that the different frameworks are like mutually untranslatable languages' (Popper, 1970). Debate of these issues falls beyond the scope of my paper, but, and this is what the advocates of paradigm shift in geography have not made clear, it is precisely such criticism that geographers must face if they embrace Kuhn's theory (see Lakatos and Musgrave, 1970).

A KUHNIAN GEOGRAPHY

The need for an assessment of the development of geography in the light of Kuhn's arguments is clear enough. In fact Preston James has already attempted such an assessment, but his handling of Kuhn's terminology in a history of geography which emphasises continuity and gradual transformation (James, 1972, especially pp. 227, 395, 505) is strikingly incompatibIe with the stress placed on discontinuity and revolutionary change in The *Structure* of *Scientific Revolutions.* He tells us, for instance, that 'it was Alfred Hettner who elaborated Richthofen's concept of chorology into a paradigm for geographical study', then argues that 'when Hettner presented this view, it had apparently been accepted by so many generations ***of*** geographers that he felt no need to support it with references to earlier writings' (James, 1972, 226-27). Moreover the differences between the German school and *la tradition vidalienne* were nothing like the intellectual gulf implied by James's use of Kuhn's terms. If correct, and notwithstanding his use of that terminology, James's emphasis on continuity in the discipline demonstrates the irrelevance to geography of Kuhn's version of scientific revolution and paradigms. However James was content with listing notable achievements or changes of fashion in geography and made no real attempt to come to grips with the complexity in even the early version of Kuhn's scheme.

While James was apparently too hasty in cutting the geographical coat to the Kuhnian cloth, Harvey (1973) was no less so in using a Marxist cudgel to knock the stuffing out of a Kuhnian straw man. Harvey argues that the overriding weakness of this 'idealist's' interpretation is its neglect of the manipulation of nature in the interest of man (especially middle-class man), yet he ignores Kuhn's counter-argument that it is precisely the overwhelming concern with paradigm-directed research that may lead the scientist to neglect pressing social problems.

Clearly, caution will be needed in assessing the relevance of Kuhn's work to geography – just how much caution can be readily seen from one example which at first sight seems a likely contender for the title of paradigm (i.e. in the sense of a disciplinary matrix). W. M. Davis's work has had an undeniably profound influence, and has recently been portrayed as little less than the hub of the geomorphic universe (Chorley ***et al.,*** 1973). Nonetheless it seems on several counts to fall short of a revolution *a la* Kuhn. Firstly, it was essentially cumulative, involving no distinct break with previous work. The great watershed surely lies earlier, at the triumph of fluvialism over its rivals. Secondly, the Davisian approach was never universally accepted, nor were its opponents regarded as heretics beyond the pale of geomorphic decency. Thirdly, the recent waning of its importance was by no means the outcome of a fatal crisis; this much is admitted by one of its most trenchant critics (Hack, 1975).

Instead we have witnessed an increasingly precise recognition of the spatial and temporal Smits of cyclicaI interpretations (for example, Schurnm and Lichty, 1965; Jennings, 1973, 1978). In fact the growing awareness of those limitations is apparent in the later works of Davis himself. Though the rise and subsequent decline of the Davisian school thus seem to have little in common with Kuhn's revolutions, they do provide excellent examples of the major attributes of 'normal science'. Davis's work undoubtedly did open up previously unthought of problems, ranging from intercontinental comparisons of erosion surfaces down to the enumeration of 'cycles' of river terracing. Most of these problems did take the form of 'puules' in which the validity of the cyclical model was simply assumed; the taunt of 'how many peneplains can sit on the top of a mountain?' (Tuttle, 1975), was not far wide of the mark. There were few attempts to thoroughly test the cyclical model. while the challenges that did arise were generally ignored or met by recourse to *ad* hocdefence: consider the fate of Crickmay's challenge; or, for some Australian examples, see Young (1978).

If caution is needed in the interpretation of Davisian geomorphology, how much more is required when we turn to the so-called revolutions of the last two decades or so. Admittedly much of the work done in this period has the same balance of faith and critical temper as Kuhn says is typical of 'normal science', while an enormous effort has undoubtedly gone into 'puzzle-solving'. Yet this same period saw the publication of Glacken's Truces *on* rhe *Rhodian* Shore, a work untouched by the 'new geography' but with few if any peers. The changes of the 1960s may well have been revolutionary, but they seem to have been less profound and widespread than was thought a few years back (for example, see Jeans, 1978, for comments on historical geography). Moreover, I hazard the guess that the new movements of the 1970s have more in common with the anti-scientific groundswell expressed in the writing of Roszak (1973) than they do with the scheme that Kuhn derived from the physical sciences.

Acceptance of Kuhn's argument that a disciplinary matrix develops essentially from the study of exemplars rather than from some explicit catechism must also throw doubt on claims that geography as a whole has been or is being transformed by one revolution or another. Where are the new examples shared by human and physical geographers, alike? Even the textbooks which presented the novice with an image of a unified geography seem to be going out of style. Changes over the last two decades have in fact cut us off from the shared examples we once had. For, with apologies to both

Lewis Carrol and Dury (1970), is it merely from nervousness not from goodwill that we march along shoulder to shoulder?

CONCLUSION

Almost invariably the versions of Kuhn's work presented in geographical writings have been distorted. The prescriptive use of the 'paradigm' concept advanced in support of vaned programmes for reform in our discipline is at odds with his theory of the growth of science. This, of course, in no way detracts from the worth of these programmes, but it does require that they are stripped of any mystique that surrounded the term 'paradigm'. **A** thorough historical analysis must replace the mere listing of notable achievements before we will know whether geography has actually undergone revolutions in Kuhn's sense of the term. But I suspect that the periods of 'normal science', rather than the sequence of 'new' geographies, will yield the greatest returns **for** a Kuhnian analysis. For like Popper (1970, *52)*, I can see the dangers of paradigm-directed research (Young, 1978), but also express an 'indebtedness to Kuhn for pointing out the distinction (from revolutionary science), and thus for opening my eyes to a host of problems which previously I had not seen quite clearly'.

UNIT 5: DUALISM AND DICHOTOMIES IN GEOGRAPHY: PHYSICAL AND HUMAN GEOGRAPHY, REGIONAL AND SYSTEMATIC GEOGRAPHY, IDEOGRAPHIC AND NOMOTHETIC APPROACH

DUALISM AND DICHOTOMY

MEANING OF DUALISM AND DICHOTOMY IN GEOGRAPHY

The word *'dualism'* simply connotes the state of being divided. For any domain of knowledge therefore, it means two conceptually contrasted stances. Dualism finally leads to *'dichotomy'* which means the bifurcation of any subject into branches of knowledge.

Ever since its inception as a domain of knowledge, geography has been encountered with several methodological issues that eventually gave birth to several dualisms and dichotomies in the subject. Such sort of dualism was prevalent even in the classical or medieval periods of geographical history. Greek scholars like Aristotle, Herodotus or Hecataeus emphasized on physical geography; Roman scholars like Strabo insisted on regional geography while Ptolemy stressed on mathematical geography; and, the Arab scholars like Al-Masudi, Al-Biruni or Al- Idrisi highlighted on the importance of the physical environment. However, such dualisms were very equivocal and abstruse.

It was in the post-Renaissance period that geography witnessed the evident rise of dualism and since then, the subject has been branched off into several exclusive domains on methodological grounds. Over time the divisions have been further sub-divided into different sub-disciplines.

Some of the most conspicuous dualisms known to have existed in geography were:

- General (Systematic) Geography versus Regional Geography.
- Physical Geography versus Human Geography.
- Historical Geography versus Contemporary Geography.

PHYSICAL VERSUS HUMAN GEOGRAPHY

It was Varenius again as one of the first scholars to highlight on the differences in the nature and content of physical and human geography. He himself however was not much interested in the latter since human geography could not be subjected to mathematical laws to generate universal principles. He believed that the methods of the natural sciences could be successfully used to draw conclusions about natural phenomena with precision to a considerable extent. But they could not be applied to human groups because they were more subject to probability than certainty. Following him, was Immanuel Kant who offered a regular course of lectures on physical geography between 1756 to 1796 at the University of Konigsberg According to Kant, physical geography not only included the features visible on the earth's surface created by natural processes but also by human actions. Kant opined that physical geography was the first part of knowledge of the world and was essential to develop the basic understanding of the earth as the abode of humans and for furthering philosophical studies.

After Kant, it was Humboldt who stressed upon the study of physical geography. Since he believed in the *'unity of nature,'* in his opinion, physical geography was the study of phenomena arranged on the earth's surface and mutually related to each other that constituted the *'natural whole.'* Humboldt was of the view that differences in the economic, social and political conditions of different spatial units were largely a function of differences in natural conditions. Thus, according to him, human factors were subordinate to the natural factors. On the contrary, by upholding the teleological view that sought to provide a philosophical interpretation for geographical phenomena, Ritter's view of geography was anthropocentric in nature. Ritter conceived the earth as created by God with a *'purpose'* to educate humans and facilitate in their development. Ritter's anthropocentric geography stated that the way natural phenomena of any spatial unit affected its inhabitants, the inhabitants could also have an influence on the land.

Under the impact of Darwinian ideas, geographers focused more and more on physical geography. In 1848, Mary Somerville authored her *'Physical Geography'* and in 1877, Thomas Henry Huxley authored *'Physiography'* as the study of nature. In the second half of the 19th century, more and more geographers were inclined towards physical geography. It is believed that Hettner accorded greater importance to the physical environment in comparison to cultural environment. German geographer, Albrecht Penck coined the term *'geomorphology'* as the study of landforms and established it as a distinctive sub-field of physical geography. It was Penck who formulated the principles of *'landform evolution.'* He also highlighted on the importance of relief maps in the systematic study of geographical elements. Subsequently, American geographer William Morris Davis also put forward his *'normal cycle of erosion.'* There were other scholars like Koppen, Martonne, Mill, and Dokuchaiev who put greater emphasis on landforms or climate as the major focus of geography. Semple went forward to explain humans as *'product of the earth's surface.'* Mackinder, Chisholm, Herberton and Huntington----all of them recognized physical geography as the core of the discipline of geography. Over the years, several sub-fields of physical geography have evolved like *geomorphology, climatology, oceanography, pedology, biogeography and environmental geography.*

The human element in geography was formally introduced in the work of Ratzel which actually gave rise to the dichotomy between physical and human geography. Ratzel in his *'Anthropogeographie'* described geography as the study of humans in the context of different races. But, Ratzel too was influenced by the Darwinian views and incorporated two major Darwinian tenets in his works-----(i) *struggle and natural selection* and (ii) *association and organization.* He used these themes in drawing an analogy between the political units and living organisms and thus came to be known as the *father of political geography*. However, it was the French geographer, Vidal de la Blache who may be regarded as the founding father of modern human geography. However, Blache was of the opinion that human geography was a natural science. He adopted an inductive and historical approach in putting forth his propositions of human geography. His *'Principles de Geographie Humaine'*, had several parts each devoted to several aspects of human geography. The introductory part analysed the principle of terrestrial unity as well as the concept of cultural milieu. Part first focused on population clusters and density. The second part was a description of man-milieu relationship.

The third part dealt with transport and communication which was completed later by Martonne who also added the components of human races, diffusion of innovation and cultural regions in it. The Vidalienne tradition was carried forward by his ardent follower Jean Brunhes who propagated Blache's views of human geography not only within France but in other parts of the world as well. French historian Lucien Febvre was also inspired by the Vidalienne human geography. He put forward that humankind emerged as a powerful agent of modifying the earth's surface through centuries of their accumulated labour and decision-making. American geographer Isaiah Bowman also championed the cause of human approach in geography. In 1924, American geographer Carl O. Sauer propounded his *'landscape paradigm'* in which he highlighted on humans as agent of *'fashioning'* the natural landscape.

Over the years, studies in human geography has led to various sub-branches in this field as------population geography, settlement geography, economic geography, social geography, cultural geography, political geography, historical geography and so on.

The dichotomy that existed between two prominent philosophies of geography namely, *environmental determinism* and *possibilism* could be attributed to the dualism between physical and human geography. In fact, the two revolutions that geography underwent namely, *the positive-quantitative revolution* and the *critical revolution* were somehow related to this dichotomy. This may be justified by the fact that while the former attempted to introduce the methodologies of the systematic sciences in geography as had been mostly done in the field of physical geography, the latter developed as a critique of the former mainly emphasized on the *'humane'* essence of the subject.

REGIONAL VERSUS SYSTEMATIC GEOGRAPHY

The dichotomy between systematic and regional geography was essentially rooted in another dualism that existed in the *approaches to study geography.* This dualism was between the

Idiographic or Inductive Approach and the *Nomothetic or Deductive Approach.* The dichotomy between the two approaches may be explicated as—the idiographic or empirical approach did not seek to develop laws but mainly focused on the description of particular places in the context of their lands, seas or places and attempt to find its relation with other places. The nomothetic or deductive approach on the other hand, sought to establish laws and made general deductions based on those laws.

Dualism in geography was formally introduced in the 17th century which is often described as the *classical period of modern geography* by the German geographer, Bernhard Varenius. Using the terms of Bartholomew Keckermann a German philosopher, Varenius in his *'Geographia Generalis'* partitioned geography into-

Special geography essentially concerned with the description of particular places on the basis of direct observations. This branch of geography was assumed to have great practical importance for governance and commerce. *General geography* based on universally applicable mathematical or astronomical laws.

Gradually, general geography evolved into systematic geography by incorporating the methods of the systematics sciences, while special geography evolved into regional geography. In simple words, the two may be expounded as the study of the natural vegetation of the world is a systematic approach while the study of a continent with respect to its natural vegetation, landforms, climate etc. is a regional approach.

The prominent German geographer Alexander von Humboldt followed Varenius and laid the foundation of systematic geography. In his famous book *'Cosmos'* Humboldt asserted that geography was meant to understand the *'harmonious unity of the cosmos.'* He distinguished between *uranography* as descriptive astronomy dealing with the *celestial bodies* and, *geography* as dealing with the *terrestrial part* with the prime objective of deciphering the unity that exists in the vast diversity of phenomena. It was not only the natural phenomena that Humboldt spoke of but, he also asserted that there existed unity of the human races as well since all the races had a common origin and therefore, no race was superior to the other. The unity of the phenomena, a viewpoint that Humboldt obtained from the German philosopher Hegel was based on the conjecture that there existed coherence as well as some sort of causality among them. The understanding of that unity was supposed to be derived from an understanding of the unity that subsisted between humans and the physical landscape. In fact, Humboldt opined that like other phenomena, *humans were basically a part of the nature*. Knowledge of the natural or physical phenomena was categorized by Humboldt as:

Systematic Sciences: This included sciences like botany, zoology or geology that classified phenomena according to their form and grouped them on the basis of certain commonalities.

Historical Sciences: This dealt with the development of phenomena over time.

Geography or Earth Sciences: This concerned itself with the spatial distribution and spatial relationship and interdependence of phenomena. It included all earth phenomena whether organic or inorganic.

Humboldt in his Cosmos stressed on his views that, for a comprehensive knowledge of the cosmos it was necessary to pursue systematic studies of particular phenomena and their interrelationship with other phenomena rather than undertaking complete studies of specific areas.

According to Carl Ritter, a contemporary of Humboldt, geography was concerned with *'lokalverhaltnisse'* or local conditions which described a spatial unit on the basis of *three* characteristics---

- *topographical*, concerned with the delineation of natural divisions on the earth's surface;
- *formal*, which dealt with the distribution and movement of such phenomena as water, air etc. that constituted the bases of human life;
- *material*, which dealt with the distribution of biotic life, minerals etc.

Ritter provided the above purpose of geography in his famous *'Erdkunde.'* It was Ritter who introduced the inductive method in geography. He sought to develop a regional geography for which he used *'erdteile'* or continents as his units of study. He was of the idea that all continents had similar physical features and thus divided each continent into a highland core drained by major rivers of the land and, a low-lying coastland at the periphery.

Thereafter, in the late 19^{th} century, geographers were highly influenced by the Darwinian doctrine and made significant contributions in furthering systematic geography. The most prominent among them were Ferdinand von Richtofen and Friedrich Ratzel.

Richtofen perceived geography in the same line as Humboldt as, the science of the earth's surface as well as the phenomena on it that were causally interrelated with it. According to him, the purpose of systematic geography was to provide an understanding of the interrelationship and causality of phenomena on the earth's surface which could be used for deducing about individual regions as well. He provided a guideline for the systematic study of the earth's surface. Richtofen also differentiated between general or systematic geography as analytic and regressive that was based on general concepts and, special or regional geography as synthetic and descriptive dealing with the unique and peculiar.

Friedrich Ratzel in his *'Anthropogeographie'* set a framework for the systematic study of human geography and thus set a new trend in the subject. Prior to him, systematic geography only involved physical geography and, human geography was mainly confined within regional studies. His anthropogeographie was essentially a reflection of the Darwinian viewpoints and emphasized on the concept of natural selection that was used in the natural sciences. Ratzel was of the view that cultural differences of a land were much more prominent than the physical differences. Ratzel's concept of geography was based on *two* propositions---(i) the interrelation of environment and humans and (ii) the interrelations of humans.

Alfred Hettner distinguished between systematic geography as that which was interested in formulating general laws and theories and, regional geography as concerned with the study of peculiarities in which the generalisations were tested to improvise on the existing theories.

The regional tradition was again revived by the French geographer Vidal de la Blache. He introduced the concept of *'pays'* or small local units as ideal units of study for the geographers which could even be used to arrive at general conclusions. He was contested however, by Reclus with his *'Le Terra'* that was centered on systematic physical geography.

The dichotomy between systematic and regional geography subsequently led to the Hartshorne-Schaefer debate. While Hartshorne in his *'Nature of Geography'* advocated that geography was regional in its essence and put forward his concept of *areal differentiation'*, his views were refuted by Schaefer as *'Hartshorian Orthodoxy'* who called for a systematic scientific approach for geographical studies.

IDEOGRAPHIC AND NOMOTHETIC APPROACH

Nomothetic and idiographic are terms used by Neo-Kantian philosopher Wilhelm Windelband to describe two distinct approaches to knowledge, each one corresponding to a different intellectual tendency, and each one corresponding to a different branch of academe.

HISTORICAL BACKGROUND

Nineteenth century German philosopher Wilhelm Windelband, a neo-Kantian, introduced these terms and defined their distinctions. Windelband used nomothetic to describe an approach to producing knowledge that seeks to make large-scale generalizations. This approach is common in the natural sciences and is considered by many to be the true paradigm and goal of the scientific approach. With a nomothetic approach, one conducts careful and systemic observation and experimentation to derive results that can be applied more broadly outside the realm of study. We might think of them as scientific laws or general truths that have come from social science research. In fact, we can see this approach present in the work of early German sociologist Max Weber, who wrote about the processes of creating ideal types and concepts meant to serve as general rules. On the other hand, an idiographic approach is one that is specifically focused on a particular case, place, or phenomenon. This approach is designed to derive meanings particular to the research target, and it is not necessarily designed for extrapolating generalizations.

IDEOGRAPHIC CONCEPT

Idiographic is based on what Kant described as a tendency to specify, and is typical for the humanities. It describes the effort to understand the meaning of contingent, unique, and often cultural or subjective phenomena. The idiographic approach is focused on an individual case and aims at finding out details about the individual case, this may be a person, as in the case of psychology where it is used to understand the behaviour and personality attributes of the individual, or a society, as in sociology. Ethnographers usually carry out such research; they use the methods such as participant observation, interviews. The ethnographers typically go to the field/society or the group of people that they wish to study and spend years at a stretch trying to observe the most intricate details of their everyday life- such as their patterns of interaction, their rituals, and their kinship ties, etc. thus one can say that the ideological approach studies the society at a micro-level. The study of Trobriand Islanders by Malinowski thus uses the idiographic approach as it is the detailed study of the life of the people there, the results derived from this study are specific to the field and generalizations cannot be applied to other societies.

NOMOTHETIC CONCEPT

Nomothetic is based on what Kant described as a tendency to generalize, and is typical for the natural sciences. It describes the effort to derive laws that explain types or categories of objective phenomena, in general. The nomothetic approach, on the other hand, is an approach typical to the natural sciences, where broad generalizations are to be made. Sociologists also have come to use this approach; in this, the research is carefully conducted to bring out results that one can apply to the broader society, e.g. the concept of ideal types as given by Max Weber, which he said to be common to all societies and thus a broad generalization. This approach aims to study society at a more macro level, using quantitative methods, such as objective questionnaires that are easier to analyse and help us make generalizations. The aim is to be able to apply such generalizations even outside the field of study, thus generalizations which are universal, such as those of the scientific laws.

NOMOTHETIC VS. IDIOGRAPHIC CONCEPT

Idiographic and nomothetic methods represent two different approaches to understanding social life.

- An idiographic method focuses on individual cases or events. Ethnographers, for example, observe the minute details of everyday life to construct an overall portrait of a specific group of people or community.
- A nomothetic method, on the other hand, seeks to produce general statements that account for larger social patterns, which form the context of single events, individual behaviours, and experience. A key debate is the one between the two research methods: nomothetic and idiographic. The debate concerns which method of enquiry is more important and which would allow greater and more valid investigation into the field of psychology.
- Nomothetic research is about attempting to establish general laws and generalizations. The focus of the nomothetic approach is to obtain objective knowledge through scientific methods. Hence, quantitative methods of investigation are used, to try and produce statistically significant results. The subsequent laws that are created can be categorized into three kinds: classifying people into groups, establishing principles and establishing dimensions. An example of this from the world of psychology is the Diagnostic and statistical Manuals of Mental Disorders (DSM), which provides the classifications for mental disorders, hence classifying people into groups.
- The method of investigation used by the nomothetic approach collects scientific and quantitative data. To do this, experiments and observations are used, and group averages are statistically analysed to create predictions about people in general. An example of this is Milgrim‘s experiments on obedience. From his scientific experiments he found that 65% of his participants would harm another person, (via a 450v electric shock) potentially killing them, within the presence of an authority figure. Although there were many ethical issues with his experiment including the deception involved and potential harm to the participant, this is an example of nomothetic research. Milgrim repeated his experiments many times and as a result created his laws of obedience.

CONCLUSION

Idiographic and nomothetic approaches should not be seen as conflicting. It is more helpful to see them as complementary. The insights from an idiographic approach can shed more light on the general principles developed using the nomothetic approach. For example, Bowlby’s theory of maternal deprivation suggests that extreme maternal deprivation is irreversible. This theory was developed using a nomothetic approach.

However, the case of Koluchova's twins demonstrates that in this single, idiographic, case maternal deprivation was overcome. Bowlby's theory can therefore be challenged. As always, it is best to take a combined approach. Millon & Davis (1996) suggest research should start with a nomothetic approach and once general 'laws' have been established, research can then move to a more idiographic approach. Thus, getting the best of both worlds!

DUALISMS IN GEOGRAPHY---A MYTH OR A REALITY

It is a fact that methodological differences had given rise to several dualisms in geography but the question that arises is that, whether the dichotomies that resulted from such dualisms had produced exclusive fields of knowledge or whether they are mutually related and transcended into one another. To be the precise, the question is whether dualisms in geography were a myth or a reality.

Varenius who actually introduced the tradition of dualism in geography, while categorizing between general and special geography asserted that they were mutually interdependent branches of geography. He stressed on the fact that special geography provided the database based on which general geography could infer the general hypotheses and laws. Humboldt recognized the interdependence of areal phenomena and opined that to understand a whole comprised of multiple phenomena it was essential to have knowledge about the constituents of that whole. Similarly, though Ritter adopted an inductive approach he acknowledged the contributions of Humboldt's systematic studies that enabled him to undertake special studies of regions. Richtofen on one hand while attempted to follow the precedence of Humboldt in establishing the affinity between geography and the natural sciences, on the other hand he also tried to restore the Ritterian tradition. It was Hettner who actually removed the dualism of the idiographic and nomothetic approach in geography. He stated that geography could involve both idiographic and nomothetic methods.

Both systematic and regional geographies could be regarded as the two extreme points of a continuum that gradually merged into the other. In regional studies, the concept of *'compage'* was introduced by Derwent Whittlesey to explain that regional geography was not a mere description of phenomena characteristic of any spatial unit but studied the functional association that existed between human beings and their physical, biotic and social environment. Therefore, through these arguments the dualism that is known to have existed between systematic and regional geography seemed to have been blurred. The general could be deciphered only through the particular which in turn, was not independent of the general.

The dualism to follow was between physical and human geography. In this too, the basic question posed was whether humans could be studied in exclusion from nature. At the same time, natural landscapes were occupied by humans. It was not possible to study human phenomena independent of the natural landscape and natural phenomena without their relation with humans. So the major thrust was on the relationship between humans and environment that constituted the central thesis of geography. This relationship however underwent several modifications sometimes according greater importance to nature and sometimes to humans or placed humans in harmony with nature. In physical geography, while explaining the normal cycle of erosion or landform evolution analogy was drawn with the lifecycle of humans, the concept of *'pays'* in human geography involved small *'natural'* regions. Thus, physical and human geography instead of being in contrast rather complemented each other which in turn faded away the dualism between physical and human geography.

Finally regarding the dualism between historical and contemporary geography, it may be stated that contemporary geography would become a part of historical geography over time. To comprehend the present it was highly essential to know the past. Therefore, historical geography provided a base for studying contemporary phenomena and how they have evolved over time. Hence, **Mackinder** asserted that historical geography was basically the study of the historical present. So, even in case of the dualism between historical and contemporary geography it was evident that one eventually led to the other and hence, their dualism also stood as illogical.

UNIT 6: POSITIVISM AND QUANTITATIVE REVOLUTION IN GEOGRAPHY

THE CONCEPT OF POSITIVISM

The origin of positivism as a well-established philosophy can be accredited to French philosopher August Comte in the 1830s. Positivism as a philosophy was mainly initiated as a polemical instrument against the romantic and speculative tradition that prevailed prior to the French Revolution. Its main purpose was to distinguish science from metaphysics and religion. Thus, positivism may be precisely described as a philosophical movement that emphasized on science and scientific method as the only source of knowledge and, which stood in sharp contrast to religion and metaphysics.

Comte rejected metaphysics for *two* reasons (i) its abstract nature with no grounding in reality; and, (ii) for being more concerned with emotional than with practical questions. He sought for *'sociocracy'* dominated by scientists for the unity and progress of the entire humanity. Since a lot of social disorder was created following the French Revolution, Comte attempted to establish positivist philosophy as an organizational tool that would lead the society through an organized development. This was much against the metaphysical principle that sought to change society through utopian solutions to the existing situations. Therefore, Comte argued that philosophy was an *'immature science'* and metaphysics should hence be replaced by a scientifically dominated *'positive'* outlook.

August Comte delivered lectures on positivist philosophy which was published as a book with the title 'The Course of Positive Philosophy.' His positivist philosophy included the following *five* basic guidelines:

- All scientific knowledge was to be based on direct and empirically verifiable experience of reality (*phenomenology)*. This was supposed to provide an edge over theoretical conjectures. Scientific methods, he asserted, was to combine both reason and experience—reason to formulate the hypothesis; and, experience to do away with falsifications.

- There was to be a unified scientific method or *'le certitude'* acceptable to all the sciences. This implied that the different branches of knowledge were to be distinguished by their subject matter or the object of study rather than their method of study. In other words, branches of knowledge differed from each other not on the basis of how they studied but on the basis of what they studied.

- This was possible only when there was *'le precis'* or a common objective of formulating scientific theories that could be subjected to empirical testing and utilized for proposing universal laws. This meant that ethical-based value judgements (beliefs, customs, norms etc.) were not to be considered as part of scientific knowledge since they were not based on direct observations and thus, were not capable of empirical verifications.

- The empirically verifiable theories so developed were supposed to be based on the tenet of *'le utile'* meaning that it should have had some utility to serve as an instrument of social engineering.

- Finally, the positivist philosophy was supposed to follow the doctrine of *'le relative'* which implied that scientific knowledge was never complete but rather relative. It kept on progressing with time through the unification of scientific theories which increased human awareness about the social arrangements that in turn, required more inclusive theories.

The philosophy of positivism challenged several taboos and religious beliefs that existed against empirical investigations. The above five postulates provided some sort of transition from the immediate through the unitary to the universal.

Comte opined that development of the society took place in *three* stages:

- *theological* when everything was described as God's will;

- *metaphysical*; and,

- *positive* when attempts were made to find out some sort of causal relations between the observed phenomena.

Comte advocated that it was true that the social phenomena were more complex than the natural phenomena, yet, there should have been a science of social relationships to be developed as parallel to and in the same principles as the natural sciences. The purpose of such social sciences would be to explore the laws governing human society through the scientific investigation of social communities. These ideas of Comte were much in tune with the proposition of **John Locke** that knowledge could only be derived through direct observations of actual situations and whatever were not supported by empirical facts could not considered as knowledge.

In a nutshell the basic tenets of positivism were:

- Positivism was also described as *empiricism* (derived from the Greek word *'empeire'* meaning experience) since it promoted science and scientific methods as a source of knowledge. It averred that science only dealt with *'empirical questions'* that were based on experiences of real conditions as they existed and that which could be tested through experiments or some other measures. It enabled to discover the causal connections between the facts to arrive into some conclusions that were supposed to be value-free, unbiased and unprejudiced.

- The positivist philosophy proved to be *anti-idealistic* that is, it stood in contrast to anything that was abstract and essentially a mental construct. Therefore, positivist philosophy did not deal with the *'normative questions'* since they could not be tested empirically and could not be established with scientific evidence. Positivism thus rejected metaphysics for being unscientific.

- Since positivism declared anything as unscientific and abstract until it could be verified with empirical evidence and tested through experiments, it did not accept authority just because it was declared as authority. This brought them in conflict with the Nazi Movement in Germany and positivist philosophy was branded as *anti-authoritarian* and the term *'positivism'* was used as an abusive term.

In the 1920s, positivism witnessed some sort of deviation from the classical Comtean ideas when, a group of scientists created the *'Vienna Circle'* and identified themselves as the *'logical positivists.'* German philosopher and physicist Moritz Schlick was the founder of this group which also had another German philosopher Rudolf Carnap as a prominent member. They upheld the viewpoint that some knowledge could also be gained without relying on experience, through formal logic and pure mathematics. Hence they distinguished between:

- *analytical statements* which were, a priori propositions whose truth could be verified through tautologies and, which were essentially the domain of the formal sciences like logic and mathematics; and,

- *synthetic statements* that were supposed to be established empirically through hypothesis testing and these in turn, were supported by the analytical statements.

Hence, logical positivism offered a much more authentic basis for scientific investigation. The essence of logical positivism was acquiring knowledge through a combination of both experience and analysis and using such knowledge to alter phenomena so as to yield a desired outcome. This philosophy included *three* interrelated precepts:

- *Scientism*: This meant that the positive methods alone were the methods of acquiring knowledge.

- *Scientific Policies*: This implied that only positivism was the key to social engineering or modification as it provided rational solutions to all social problems.

- *Value-Freedom*: This implied that scientific judgements derived through positive methods were neutral, unbiased and objective and hence were free from any moral or political binders.

POSITIVISM IN GEOGRAPHY

There was a great deal of efforts in the latter half of the 19th century to develop the discipline of geography as a *nomothetic science*. This was largely the impact of the Darwinian tradition that invigorated the scientists to search for the governing laws of nature and in the same tune, the social scientists to explore the laws determining social arrangements. The *hypothetic-deductive approach* of study that was especially characteristic of the natural sciences, replaced the inductive methods in the social sciences. Thus there was an effort to accommodate social sciences within the framework of positivism. It must be pointed out here that the geographical developments that took place in the 1950s and 1960s were mainly committed to logical positivism. The researchers sought to develop a priori models about reality for which they devised a set of hypotheses that were to be authenticated, validated or discarded through testing of empirical data. Once verified, they were validated as laws until their eventual refutation through further research. The logical positivists conceived that some order persisted in the objective world that needed to be explored and discovered through scientific investigation- spatial patterns of variation in geography----that could not be manipulated by the observer. Geography soon became *'positivist-led.'* The hypothetic-deductive approach led the discipline particularly human geography to develop as a model building and theoretical science since it dealt with phenomena that were familiar with reality both spatially and temporally.

During the 1950s and 1960s, the essence and purpose of Anglo-American geography witnessed a drastic transformation with the replacement of the *idiographic approach that focused on areal differentiation* with the adoption of the *nomothetic methods that sought to explore models of spatial structure.* This change was initiated by Schaefer with his critique of Kant's exceptionalist views that placed history and geography as exceptional and different from the other systematic sciences. Schaefer put forward his *'spatial organization paradigm'* and conceived geography as a spatial and social science primarily concerned with the formulation of laws that governed the spatial distribution of any phenomenon as they were found on the earth's surface. Hence, Schaefer set off a sort of *'revolution'* in geography that was basically *'theoretical and quantitative'* in nature. This revolution in geography sought to provide the discipline a scientific approach with the application of mathematical and statistical methodologies. It largely accepted the tenets of positivism of unified scientific methods acceptable to all the sciences----natural and human. The quantitative schools undertook to construct models and theoretical structures within which geographical realities were supposed to be incorporated.

The quantitative revolution that geography underwent by adopting the viewpoints of positivism was set off mainly by the mathematicians. It was mainly the outcome of the impact of the non-geographers on geography. As was evident in many other disciplines, it altered an already existing knowledge base with a mathematical approach. William Bunge in his *'Theoretical Geography'* (1962) went as far as to describe geography as a science of spatial relations and geometry as the mathematics of space. So logically, geometry was supposed to be the language of geography.

What followed was the development of the concept of *'space'* as the basic concept for organizing the subject matter of geography. *Two* major approaches were identified in the study of geography, namely----

- ***Spatial Analysis***: This referred to the application of quantitative or more specifically statistical methods and techniques in locational analysis.

- ***Spatial Science***: This concept was largely akin to the positivist philosophy and presented human geography as a social science with its prime focus on space as the guiding principle behind the organization and operation of the society and the behaviour of individual human being.

These two approaches related to space led to the development of the following *two* aspects of space that became the central theses of geography:

- *Spatial Interaction*: This referred to the interdependence between spatial units and was gratuitous to the interaction between humans and environment within a particular area.

- *Spatial Structure*: This referred to the spatial arrangements or more precisely the geometric pattern of any phenomena on the earth's surface since geometry was regarded as an important tool in geographical studies.

Inspired by positivist thinking, major advances towards a unified methodological and philosophical basis of the quantitative schools were rendered by Peter Haggett, Richard Chorley and David Harvey in the 1960s. The discipline of geography witnessed major theoretical and methodological developments. A new domain of knowledge emerged that came to be known as *regional science*. It was basically an assemblage of geography, economics and planning with its main concern for regional problems. The pioneer for this new discipline was Walter Isard (1956).

The most important theoretical development that fundamentally incorporated the philosophy of positivism was the *locational analysis* of Peter Haggett. The concept was put forward by him in the book *'Locational Analysis in Human Geography'* (1965). Following the geometric tradition this approach in human geography, more popularly termed as *spatial science* concerned itself with, the spatial arrangements of phenomena on the earth's surface. In addition to this, it also dealt with the interaction between places within a spatial pattern, the dynamics of such patterns as well as the creation of alternative patterns through model building to provide for a better possibility. The Central Place Theory postulated by August Losch (1954), the Gravity Model by Stewart and Warntz (1959) or the Diffusion Theory of Hagerstrand (1953) were all formulated using locational analysis.

Another concept that was intrinsically associated with positivism was the concept of systems. A system was defined as an array of entities that had specific relationship among them as well as with their environment. Richard J. Chorley was the first geographer to introduce general systems theory in geography. His paper *'Geomorphology and General Systems Theory'* (1962) was developed within the framework of the systems approach in which he tried to apply the concept of open and closed systems to geomorphology.

A major contribution to the positivist theory was made by David Harvey in his *'Explanation in Geography'* (1969). He opined that reality was a set of complex phenomena particularly so far as the relationship between the phenomena were concerned. However, it was possible to decipher such complexities with the aid of system analysis which explored the structure and function of a system. Every system was supposed to have three fundamental aspects structure,function and development. The structure of the system was the set of elements it was comprised of and the relationship between them; function was the exchange or the flow between the elements while development meant the changes in the structure and function of the system over time.

Since geography studied the relationship between humans and the environment, systems analysis was supposed to have a wide range of applications especially in human geography. This was because the systems analysis was based on an implicit assumption of positivist philosophy and drew analogies between human societies on one hand and natural phenomena on the other.

This drawing of analogies led to the model and analogue theory that had close connections with positivism. A model was basically a structured conceptualization of reality that represented particular attributes of reality and, analogue theory was the formal theory related to building of models. A model or an analogue ranged from a structured idea to a hypothesis to a law to a theory. Following the positivist outlook, a model could be used as a guide to validate a set of hypothesis through empirical testing and to establish a theory as it contained some resemblances with the reality. Though model building had been used in many sciences since long back but, its use in geography was of comparatively recent origin and could be attributed to positivism.

CRITIQUE OF POSITIVISM

The positivist philosophy by rejecting metaphysics provided a sound philosophical, methodological and scientific base to the discipline of geography. Knowledge based on the observations of real situations that could be easily verified empirically was highly objective, unbiased and unprejudiced and could be readily utilized for the formulation of universal laws and theories. Positivism encouraged the use of statistical and mathematical techniques that provided precision to research and enabled to analyse a geographical system in a much more simplified form. It provided a kind of framework within which theoretical statements could be formally presented. However, the critique of positivism was highly intense and convincing. Positivism was criticized mainly on *three* fronts:

Empiricism**:** Positivism recognized the fact that theory building was essentially based on the direct observation of reality which could be subjected to statistical procedures for empirical verification. But this approach proved to be very superficial against which new philosophical and theoretical frameworks were designed and for which alternate methods than statistical techniques were required. These, like realism and structuralism offered a much more insightful exposure of human society. In contrast to positivism that concerned itself with *'how'*, they were concerned with *'why'* and went beyondthepositivists.***Exclusivity*****:** The positivists' claim that the methods of natural sciences could be extended into the domain of the social sciences including humanities to establish a unified scientific method was also criticized. Positivism excluded the normative questions like beliefs, values, emotions, attitudes and so on. But in reality, much of human behaviour and social arrangements was to be guided by such questions. Hence, it provided a very parochial approach to the study of any domain of social sciences.

Autonomy: The assertion of positivism that knowledge based on direct observation and verified empirically would yield a scientific discipline that would be objective, neutral and unprejudiced was widely challenged.

Among the social sciences, human geography was the one to adopt the positivist doctrine in a great way as it provided a systematic and scientific approach to the discipline and where reality could be verified. This new paradigm was widely accepted particularly in the fields of urban and economic geography. Schafer's paper on exceptionalism opened the door to the domain of (logical) positivism. The geography that developed by adopting the positivist doctrine emphasized on analyzing spatial data and developing spatial theory based on empirically tested mathematical models. However, in the 1970s and 1980s, there was increasing dissatisfaction among the geographers with over emphasis on spatial views and they sought to explore alternative approaches to geographical problems. They argued that human geography employing the spatial view was actually a sort of *'fetishism'* that alienated, diverted and obscured the fundamental social questions. For this reason, even David Harvey later deviated from the positivist stand to focus on the question of social distribution.

The critique of positivism in geography mainly emanated from *two* sources:

- Its acceptance of statistical techniques for making inferences about reality; and,
- its acceptance of the assumption of the methodological unification of the sciences.

Regarding the first criticism, Bennett in 1985, highlighted the following points:

- Positivism created a false sense of objectivity. The models constructed using statistical techniques that were considered as an effective tool of theory building actually deviated the observer from the observed by giving more prominence to some elements and undermining others. This paved the way for controlling and manipulating society. They were regarded as grossly inadequate for geographical enquiry as well due to non-repeatability of experiments and data.
- By employing quantitative techniques positivism largely eliminated the social and humanistic concerns and reduced humankind as decision-makers or workers to mere passive agents. Such models mostly turned out to be the result of economic determinism.
- Positivism described the existing real situation and thus encouraged status quo in society especially with respect to the distribution of social well-being.
- Since it excluded the normative questions thus it deprived human society of the norms and values based on which it should have been organized. With its overenthusiasm with empirical questions, sometimes it overlooked many good qualitative statements that would have otherwise proved to be effective in describing regional entities.
- Positivism attempted to construct theories with universal acceptability by moving from particular to the general which had reduced validity in real world owing to the spatial character of geographical data. Hence, this resulted into overgeneralization.
- Availability of an extensive and reliable database was an important pre-requisite for the application of statistical techniques in human geography in the absence of which the models or theories developed were supposed to portray an erroneous and distorted picture of reality.

So far as the second criticism was concerned, some geographers like **Peet** or **Slater** derided the positivist-spatial science tradition in geography and even challenged its methodological base as a whole. Their rejection of positivism implied a rejection of the concept of space and hence the rejection of the entire subject matter of geography. They argued that spatial science with geometry as its language was not adequate for addressing geographical problems. Though space was regarded as the central theme of the subject, yet the discipline could not be distinguished from others solely on that ground. The concept of space merely reduced geography to a heterogeneous amalgam of spatial models that yielded no process of understanding or specific theoretical object to geography.

The separate development of physical and human geography also eroded the possibility regarding the unification of the subject. In most situations, human geographers minimized the role of the environment while physical geographers sought to develop theories detached from human and social needs.

The critique of positivism revived the social foundations and responsibilities of social sciences due to which geography went through an anti-positivism and critical revolution in the 1970s. A lot of humanistic approach was proposed to counter and replace the concept of an objective world highlighted upon by the positivist paradigm.

QUANTITATIVE REVOLUTION IN GEOGRAPHY

INTRODUCTION

In the 1950s and 1960s, a revolutionary change described as "*quantitative revolution*" occurred in the discipline of geography. It replaced the ‘idiographic’ approach based on areal differentiation by ‘nomothetic’ one, which had its roots in the search for models of spatial structure and phenomenon. The *quantitative revolution* led the basis of geography as a spatial science that dealt with the spatial analysis of phenomena that existed on the earth surface. In simpler words, it gave geography a scientific vision through the application of methodology rooted in statistical methods. Some of the elements of positivism, which had previously been not accepted at some point in time, were now accepted open-handedly. In the words of Burton (1963) this school had set out to discover universals, to build models and to establish methods and theoretical bases on which geographical realities could be erected. Traditionally, geography was a discipline that studied and described the surface of the Earth, but in due course of time, its definition and nature have changed. It was now related to providing accurate, systematic, rational descriptions and explanations of the variations in the geographical phenomenon that occurred over the Earth's surface. The most obvious change has occurred due to the *quantitative revolution* that brought changes in the methods and techniques used to explain the geographical phenomenon in a spatial framework.

The movement that led to the occurrence of quantitative revolution in geography was initiated by natural scientists specifically physicists and mathematicians. It expanded and led to change physical sciences followed by biological sciences. By the late 1960s, it became a feature of most of the social sciences. These include economics, psychology, and sociology; though had faint impressions in the disciplines of anthropology or political science, has not occurred in history.

The main objectives of this paradigm in geography were first, to change the narrative character of the subject (geo + graphics) and make it a scientific discipline. The second objective was to explain and interpret the spatial patterns of geographic phenomena in a logical and objective pattern way. The third objective deals with the use of mathematical and statistical techniques; fourthly, to make accurate statements (generalization) about location order; fifthly, to prepare estimates, principles and laws for testing estimates and estimates and forecasts and lastly to provide a sound philosophical and theoretical base to geography, and to make it a scientific discipline.

These objectives lead a number of dichotomies within the discipline apart from the quality dichotomy. Now, these included measurement by instruments versus direct sense-data; rational analysis versus intuitive perception; cold scientific constructs developed in the laboratories versus rich daily sensed – experience from the real world itself; constantly changing phenomena versus discrete cases; nomothetic versus idiographic, to mention a few. If one tries to seek answers to these dichotomies he gets trapped within them and is unable to understand the movement towards quantification in geography. Thus, to avoid this we shall concentrate on how this movement became part of the discipline and slowly engulfed it in such a manner that it led to the spread and growth of scientific method in geography.

QUANTITATIVE REVOLUTION IN GEOGRAPHY

Traditionally, Geography has been a "following" discipline; the main streams of ideas had their roots in other disciplines. The doctrine of environmental determinism was represented in the writings of Semple, Huntington, G. Taylor, and Ratzel (if he can be considered a determinist). They were busy with the idea of a causal relationship and were regularly demanding and looking for "laws". A similar mechanical flavor existed in the works by "Quantifiers". It seems as if geography is re-emerging after it got soaked in ideographic approach, which created a distance between geography and environmental determinism. It seems in some way or other; the quantitative revolution took geography closer to environmental determinism especially as this revolution occurred simultaneously with the upsurge of neo – determinism. The Quantitative Revolution, but natural, was strongly opposed and the dominance of environmental determinism delayed the process of establishment of the scientific basis that the quantifiers wanted to provide.

It was vehemently opposed in the United States as determinism had its strongest base there. Still, new techniques were been used and others were being developed as part of the prevailing probabilistic trend in contemporary science. In the words of Bronowski (1959) in simpler terms, statistics replace the notion of inevitable effect by probable trends. As the revolution progressed the use and purpose of use of statistical techniques that are quantification became more and more indeterministic.

In geography, the revolution began in the late 1940's and culminated in the period from 1957 to 1960; finally, over in 1963, the year Burton wrote his paper. In between these years, it did gain momentum especially after Ackerman and Schaefer favoured in making geography more theoretical and systematic in nature. Ackerman commented, "although the simplified forms of statistical assistance have been part of geographic distribution analysis in the past; discipline is beginning to move towards more complex statistical methods-a completely logical development'. Burton further commented that both Hartshorne and Spate also agreed on the usage of these techniques in geographical thinking.

The reference of Hartshorne (1959) is being made to his statement, which says, "To raise ... thinking above the scientific knowledge level, it is important to establish generic concepts, which can be implemented with maximum objectivity and accuracy through quantitative measurements which can be subjected to comparisons through the mathematical logic".

Spate (1960) in his paper on "Quantity and Quality in Geography''; published in the Annals of the American Geographers seems somewhat skeptical about quantification in geography. The report of a National Academy of Sciences – National Research Council Committee on 'The Science of Geography' (1965) also discussed the influence of quantitative revolution in geography. They stated that geographers believe that correlation of spatial distributions, considered both statistically and dynamically, maybe the keys to understanding the development of living systems, social structures and environmental changes that occur over the earth surface. In the past progress was slow and gradual as the number of geographers was less while the problems were numerous. Moreover, the methods of analyzing these multi-variate problems were rigorous. It was only recently that systematic concepts and approaches have been adopted to analyze these multifaceted problems.

THE PATH OF THE QUANTITATIVE REVOLUTION IN THE DISCIPLINE OF GEOGRAPHY

The roots of the revolution were in the following publications, which had their significant influence on the incidence and growth of quantification in geography. These are – Neuman and Morgenstern's *Theory of Games and Economic Behavior* (1944); Weiner's volume on *Cybernetics* (1948); *Human Behaviour and the Principle of Least Effort* by Zipf (1949) and Stewart's paper entitled *Empirical Mathematical rules Concerning Distribution and Equilibrium of population'* (1947). Stewart's paper needs special mention as he put forward a new way to raise the old geographic questions.

The effect of quantification began to be felt immediately in geography. Rather its rise has been startling in its suddenness. Quantification did increase in geography and one should accept it as it had a valuable role to play. For example, in 1936, John Ker Rose argued in his paper on corn cultivation and climatic conditions that "the methods of relational analysis would be particularly promising tools for geographical investigation." This call was largely ignored. Strahler initiated an excellent petition when he attacked Davis's explanatory and descriptive explanation of geomorphology and supported G. K. Gilbert's dynamic-quantitative systems.

A) QUANTITATIVE REVOLUTION IN THE BRANCHES OF GEOMORPHOLOGY AND CLIMATOLOGY

Strahler claimed that Gilbert's paper was more apter than Davis's work; then what was the reason that it was not accepted as a sign post in geomorphology for future work; rather it has been forgotten and neglected for nearly thirty years. The answer is with Strahler himself who opines that thinks that geomorphology was a part of geography. The physical geographers did not adopt these ideas rather they followed Davis. Some of the prominent followers of Davis include *Douglas Johnson, C A. Cotton, N. M Fenneman,* and *A.K. Lobeck*. Strahler finally states these geographers made "excellent contribution to descriptive and regional geomorphology" and has provided a solid foundation for study in "human geography", but did not lay the basis for scientific study within the geographical thinking. This does not mean that prior to Strahler; geographers were not using quantitative techniques in geomorphology. Quam and Woolridge vehemently criticized his views. Quam (1950) states that mathematical formulae and statistical analysis in geomorphology may result in showing an unrealistic picture of reality that might not be accurate and objective.

Similarly, Woolridge (1959) critics Strahler's views and states that although there is the prevalence of a 'new' quasi-mathematical geomorphology; it is inadvisable to use mathematics at a higher level as these are not apt in explaining the geomorphologic phenomenon. He further states that whatever the case may be they will continue to regard W. M. Davis as their founder and would criticize all those who do not agree with the methodology of Davis's interpretations of a different phenomenon occurring over the earth surface. It is not that geomorphologists did not adopt quantification; Strahler did find his support in L. King (1962) who writes that statistical methods are useful for bulk studies and can be well appreciated if used to study complex phenomenon and processes that constitute a large number of variables or indicators. Although few studies in the branch of geomorphology can apply them, they should be used with great precision so that results are not superficial in nature.

Many geomorphologists in addition to Strahler like Chorley, Dury, Mackay and Wolfman, used quantitative methods and it seemed that the practice would spread. In the case of climatology, there is little dispute about the use of quantification. This branch of geography whole-heartedly embraced these new statistical techniques to interpret various climatic phenomena. Examples can be cited from the works of Thornthwaite, Mather and Green, Bryson who have successfully implemented quantitative techniques to seek answers for climate problems; thus silencing their critics.

B) QUANTITATIVE REVOLUTION IN THE BRANCHES OF HUMAN AND ECONOMIC GEOGRAPHY

So far, the biggest struggle for approval of quantification has been in human and economic geography. It is not surprising that in view of the possibilist tradition; it is here that the revolution runs against the ideas of independence and the uncertainty of human behavior. Here comparisons with physical sciences are useful. Physicists who work at a microscopic level, with quanta and energy, face similar problems that social scientists face with people. Such parallels when recognized are a reason for happiness and not for disappointment. In order to make a reputable place in human society, social science must get direct results in the form of a prediction science that does not need any kind of control, restriction or regiment the person. A social science that distinguishes random behaviour at the micro-level and is even able to foresee results at this level is nothing but the consequence of quantitative revolution.

Several works can be cited which used statistical techniques in a positive manner. Most interestingly large number of debates took place between scholars that appeared in the literature (Burton, 1963). Some of these are worth mentioning – Garrison's and Nelson debate on Service classification of cities; Reynolds – Garrison's deliberation on the modest use of quantification in geography. The Spate – Berry argument in *Economic Geography* that ends on the agreement that statistics are half of a filled glass, the other half is understanding and interpretations. The list is endless but some of the other debates that need to be mentioned include the contest between Zobler and Mackay on the use of chi-square in regional geography and the dispute of Lukermann and Berry on 'geographic' economic geography.

The deliberations were done through professional magazines, which got them the much-needed attention. The result was the establishment of the Regional Science Association in 1956 that promoted quantification in geography. Moreover, it made quantifiers an essential part of the geographical thinking and giving them appreciation and approving their work part of the geographical academia.

Although most of the literature cites that, the revolution is over, it has remained active in several sub-branches of geography particularly transport, economic, and urban geography. This is evident from the fact that writings with quantitative methods have been regularly published in acclaimed journals in geography, including *Annals of the Association of American Geographers, Geographical Analysis, Environment, and Planning A, The Professional Geographer, Journal of Geographical Systems, Urban Geography,* and many others (Kwan and Schwanen, 2009).

Although quantitative geography is generally "perceived as a relatively static research area," it is actually "a vibrant, intellectually exciting, area in which many new developments are taking place" {Fotheringham, Brunsdon, and Charlton (2000); Clark (2008); Golledge (2008)}. Interestingly, quantification in geography has changed its course in due course of time. It now an ally of critical geographies - for example, the emphasis has shifted from global generalizations to local levels dealing with local relationships in a spatial framework. It has also become sensitive to variables like gender, race, ethnicity, sexuality, and age; and even pays attention to processes which shape individual's spatial behaviour (Kwan and Weber 2003; Poon 2003; Fotheringham 2006).

Quantitative research is still dominant in the fields of transport, economic, and urban geography in the writings of McLafferty and Preston (1997), Rigby and Essletzbichler (1997), Plummer and Taylor (2001), Schwanen, Kwan, and Ren (2008) and Bergmann, Sheppard, and Plummer (2009). In this regard, Kwan and Schwanen (2009) are of opinion that knowledge in statistical methods is essential for decoding and challenging regressive political agendas; often supported by numbers and quantitative analysis. Quantitative geography, when incorporated with a critical sensibility and used suitably, can be a powerful device for encouraging progressive social and political change.

THE CRITICISM OF QUANTIFICATION IN GEOGRAPHY

The quantitative revolution was initially propounded to make the discipline of geography a scientific discipline where the validity of the knowledge that was generated was justified according to the principles of positivism. Although many geographers like Plummer and Sheppard (2001); Kwan (2004); Fotheringham (2006) have argued that quantitative geography does not necessarily have to be based on the epistemological premises of positivism. Whatever the case may be it is to be understood that when positivist epistemology was adopted, the purpose of the geographic research was to seek universally applicable generalizations. The criticisms became more prominent as critical geographers started questioning the relevance and value of spatial science in the early 1970s. Now quantitative geography was labeled as positivist and empiricist because it was based on the principles of scientific objectivity, value neutrality, and the search for universally applicable generalizations. One of the groups that criticized quantification was the group of feminist geographers that was critical of the tendency to draw conclusions based on the principle of universal causality from inferential statistics (Kwan and Schwanen, 2009).

Quantification was also criticized for other reasons. For instance, there were those who thought that this method would mislead geography towards a futile course. Some like Stamp argued that quantifiers were too busy in sharpening their instruments that they forgot the real purpose of the revolution. Few opponents also commented on the suitability of statistical techniques for all kinds of geography. They opine that these techniques were appropriate for some branches and not the entire geographic paraphernalia. Another group condemned this revolution on a note that there was a confusion of ends and means. In the words of Spate (1960), 'it is important to count what can be counted'. Another dichotomy lies in classifying and understanding; classification should never be mistaken for comprehension. Goodall's (1952) point is worth pondering where he states that quantitative methods or statistical techniques are only adjunct to elucidations or descriptions; they can neither provide explanations nor replace them. Therefore, these methods should be observed only as useful tools and not keys to universal knowledge (Spate, 1960).

These criticisms clearly point out that the quantitative methods have some severe limitations, especially when applied to the study of certain kinds of phenomena—for example when the purpose is to uncover the complex gendered, racialized or sexualized experiences of individuals or the socio-spatial construction of identities. However, this does not imply that quantification is not in a position to make valuable contributions in the field of geography. The difference lies in the time period if we talk of the 1950s or 1960s maybe this was not possible but in contemporary geographic research, it is possible to reconnect the critical geographies with quantification. Another point of deliberation is that within the discipline of geography several subfields, like transport geography, are historically more quantitative in nature than others; this happened because of the influence of allied fields such as civil engineering and neoclassical economics (Kwan and Schwanen, 2009).

It can be said that the revolution had an early demise; it means that the purpose of the revolution was achieved or not. If seen from the point of view of Burton (1960) its basic purpose was to make geography more scientific and to develop a body of theory. Discontent with the idiographic approach in geography is the root of quantitative revolution; the development of theoretical and model-building geography with a nomothetic approach was the expected result. The basic rationale was to develop scientific method; to develop the theory and to test the theory with the prediction for which the logic of mathematics is the best tool available.

CONCLUSIONS

The use of statistical or quantitative techniques is one of the most suitable methods for the development of theory in geography. The revolution can never be over until it is able to seek answers and aid the theoretical development of the discipline. Moreover, theory development and its testing are the only ways of creating new knowledge and subsequently verifying it. Models have just formalized ways of descriptions that an author has visualized and represented through his arguments and justifications. In geography, quantification brought this revolution where the ideographic base was replaced by theory building in a nomothetic approach.

Geographers started developing theories and created '*new*' geography that focused on the philosophy as well as methods. These scholars were of the view that mere description, mere quantification, and mere abstraction were valid to a certain extent; but repeated use of these methods makes one an obscurantist. *Theoretical geography* got its philosophical base in Bunge's monograph published in 1962, which identified geometry as the mathematics of space and hence made spatial science the language of new geography. Harvey's *Explanation in Geography* (1969) provided a more inclusive channel for the methods and philosophy of new geography. Apart from these scholars, the Department of Geography at Lund University, Sweden became a centre for quantitative and theoretical geography under the leadership of Hagerstrand and Morill. Hagerstrand although based in Seattle provided an academic support to the geographers working in this field at the Lund University. To conclude, whichever method one, the purpose of geography is to seek answers to questions pertaining to problems of quantity and value. Most of our experiences are qualitative in nature and when everything is, reduce to numbers; some essential attributes are lost (Huxley, 1951). Thus one needs to maintain balance as still new worlds are to be conquered and new contributions to be made (Bansal n.d.).

UNIT – 7: SYSTEMS APPROACH IN GEOGRAPHY

INTRODUCTION

The word system has been derived from the Greek word "system" which means a set of rules that govern structure and behaviours. In other words, the system is termed as a unified whole (working body) which consists of interdependently functioning elements. The element is very basic part of a unified whole. For example, the human body is a biological system involving various elements (parts) like cells, tissues, blood, bones, and muscles. These elements (parts) are functioning interdependently. Likewise, the Earth itself is the largest system which is made of lithosphere, hydrosphere, atmosphere, and biosphere. The biosphere is the largest ecosystem made of interconnected sub-systems (both terrestrial and aquatic ecosystems) viz., forest, grassland, desert, ocean, lake, pond etc. These systems vary greatly in size and scale ranging from microscopic to micro, meso and macro. For instance, biosphere forms an ecosystem of macro size and rivulet can form a micro one.

SYSTEM APPROACH TO GEOGRAPHY

The concept of the system approach has been fundamentally derived from the general systems approach or theory. A biologist named Ludwig von Bertalanffy's seminal paper on open systems is attributed as a seedling for the rise of the system movement. He has published various papers on a system approach to biology between the 1920s and 1950s. His papers aimed at giving account for the key distinction between the organismic systems of biology and the closed systems of conventional physics and understanding common laws that govern the life of organisms. Through his general system approach, he comprehended intrinsic unification of different streams of sciences and fusion between science and environment. For von Bertalanffy, the main propositions of general systems approach or theory were:

1. Isomorphisms between the mathematical structures in different scientific disciplines could integrate and unify the sciences;

2. Open systems require consideration of the flow of energy, matter, and information between the system and its environment;

3. Within open systems, the same final system state may be reached from different initial conditions and by different paths – open systems exhibit equifinality;

4. Teleological behaviour directed towards a final state or goal is a legitimate phenomenon for systemic scientific inquiry;

5. A scientific theory of organization is required, to account for wholeness, growth,

disorientation, hierarchical order, dominance, control and competition; and

6. GST could provide a basis for the education of scientific generalists

As per the above-mentioned propositions, a general system is a unified whole of elements bound together by specific linkages. It is higher order generalization of a multiplicity of systems, their complex structures, and functions. This is an analytical framework to unify various sciences. It has a self-sustaining mechanism. Such systems may be open or closed and may change over the period of time. Most general systems, however, are open. As discussed earlier the earth is an open system in which there are inputs, outputs, and energy flow through a variety of mechanisms. The linkages, or connections, that bind entities together into a system, are paths through which matter, energy, ideas, and people pass from one element to another.

As very early, Ludwig von Bertalanff y had realized the importance of system approach to non-biological science. Over the period of time, geographers introduced system approach to various geographical studies. R. J. Chorley, Leopold and Landbein, Wolderberg and Berry were some prominent geographers who made its application possible in geography. In fact, it was R. J. Chorley, who become the first geographer brought about this approach or theory to geography. His paper "Geomorphology and General system theory" (1962) focused on the mainly application of the concept of open and closed system in Geomorphology. In addition, Leopold and Landbein applied entropy and steady state in the study of the fluvial system. Berry developed a basis for the study of the city as a system within the system of cities in spatial form by using the two concepts viz., organization and information.

Wolderberg and Berry have applied system concept to analyze central place and river system. Curry has tried to use this concept to the spatial location of settlement. Even new age geographers solve numerous spatial problems through a general systems approach in both human and physical geography. They use the system as analytical models to understand and explain spatial patterns and interactions. Human geographers, for example, use the system model to examine human migration patterns, the diffusion of ideas, and the spread of information. Moreover, researchers in physical geography trust this approach in understanding natural set up in which physical system operates. In short, researchers in both physical and human geography are interested in identifying, explaining, and predicting flows in human and physical systems. They also search, identify, describe, and explain cycles and patterns of geographical phenomena in different branches of geography.

In other words, the systems approach is used in a variety of applied branches of geography viz., land-use planning, natural resource management, watershed management, regional planning, management of pollution (water, air, soil, sound), environmental management, climate change etc.

These areas involve the study of elements and sub-system of general environment like quantitative, qualitative, behavioural, socio-economic, and political subsystems. The qualitative subsystem encompasses finite space including urban, rural, empty or filled places, psychological spaces, and their various-use. It also includes a variety of concepts like capability, carrying capacity and stability.

The economic subsystem comprises decision-making processes based on well-tested economic theories whereas citizens, governments, civil administration, and civil societies form the political subsystem, which plays a vital role in geopolitics of a state (nation) being studied in political geography. The behavioural subsystem consists of attitudes, values, beliefs, customs, and traditions, which are integral parts of behavioural geography. A general systems model is a composite one in which physical and socio-economic variables are intricately linked. Some of the variables may be measured quantitatively and some may not. The systems approach involves a number of relationships (links) between variables (elements). Now geographers use statistical tools like multi-variable analysis, principal components analysis (PCA), probability theory, Chi-Square and Gini-coefficient to analyze data on geographical variables explaining interlinks between them in a system.

I. Systems Analysis

A. Basic Elements

System, as discussed earlier, is a unified whole or working body, which consists of interdependently functioning elements. There are multiple variables or elements that form a system. A system forming basic elements is as follows:

1. **Inputs and Outputs:** Every system requires a regular flow of inputs for producing an amount of output. Without inputs, no outputs can be produced in a system. For example, a fixed agricultural output needs a certain amount of inputs of seeds, water, fertilizers, labor etc. in the assumed constant external environment.

2. **Processors:** It involves the actual transformation of input into output. It is the operational component of a system. Processors may process the input totally or partially, depending upon nature, amount, and requirement of the output.

3. **Control:** This is an important element, which guides and controls the system. It is basically the decision-making subsystem that governs the pattern of activities like processing input and producing output.

4. **Feedback:** Feedback is an indication for characteristics, amount of produced output

against the set parameters and standard. Feedback is conducted on the principle of cybernetics which comprises communication and control. Feedback is of two types., viz., positive or negative. It is good to have positive feedback, which strengthens system's performance. Negative feedback gives the wrong signal if the system is not functioning well and it also makes available the controller with information on actions required for its correction.

5. **Environment:** Every system is operating in a unique environment. It is a broader framework often called "supra-system", Which affect operating system and determining routes and rules of its functioning.

6. **Boundaries and Interfaces:** A system has delineated boundaries or limits through which it identifies its components, processes, and interrelationships when it interfaces with another system or with its environment.

B. Components of a System

A set of elements form component and a set of components form a system. All systems of varying scales are having three basic components as follows:

1. A set of elements

2. A set of functioning links

3. A set of links between system and external environment

All systems have internal and external environments. The external environment influences the internal environment of the system. Some systems are close and some are open. A closed system can easily be created in science laboratories. In other words, the closed system exists mainly in a controlled environment like chemical labs. For example, chemists conduct chemical tests in their labs but such kind of total control is not possible in an open environment like agricultural or forest land, but in open environments like agricultural land, some elements can be partially controlled. Such partially controlled environments are of great importance for semi-scientific experiments aiming at socio-economic development. This is to be understood with this example. If farmers want to know impacts of certain inputs on a crop production. Impacts of inputs like fertilizers, pesticides, high-yield seeds, labour on crop production can be known by controlled and regulating inputs in a farm. With above discussion, the basic characteristics of a system can be inferred as follows:

1. System is a well organized and an integrated ideal body;
2. Systems have multiple elements and components;

3. The components in a system are interdependently functioning;
4. Systems have a structure and pattern of behaviour;
5. Systems have a boundary and interfaces;
6. Systems are at balancing and enduring state;
7. Systems affect and are affected by their internal and external environment;
8. Systems exhibit feedback;
9. Systems are either closed and open and
10. Predominant systems in the environment are open-ended.

C. Essential Features

Above mentioned basic characteristics exhibit some interlinked essential features of a system as follows:

1. Environment of a System

2. The Behaviour of a System

3. The State of a System

4. Organization and Information in System

5. Structure of a System

1. Environment of a System

The environment of a system is supra-structure in which system operates. There is countless system and their sub-systems are working in our environment. The environment is classified as physical environment, socio-cultural environment, political environment. Socio-cultural environment and political environments operate in their physical environment. In fact, each system has its own physical, socio-economic environment that affects the performance of that system positively or negatively. For example, farmland is a system, its agricultural productivity depends upon its physical, socio-economic environment. A system may have the internal and external environment. For example, the productivity of an industry as a system is affected by both internal and external environment.

2. The Behaviour of a System

It refers to all dynamic activities of a system like the introduction of new stimuli, flows, and responses, inputs, and outputs etc. It studies the flow of energy between the elements of a system and between a system and another system. Functions within a system are called internal behaviour and outside a system is external behaviour. Internal and external behaviours are interlinked. For example, an element of a system is the part of the external environment, change in external environment will bring some behavioural change in one element and change in one element will affect all inter-connected elements of that system. Such behavioural change can be explained by its flow from input (simple stimulus) to the output (response).

The input-output analysis in economics is a popular example for explaining the flow of behavioural change. Increase in final demands (e.g. derived from exports, home consumption, or another way around) is working as stimuli to rise in final outputs in various sectors in an Indian economy (a system).

3. The State of a System

Each system wants to achieve its state of equilibrium. A slight change in one element of system disturbs its equilibrium. A disturbed system may experience morphogenesis to gain a complex level of equilibrium. 'Equilibrium' denotes maintaining a kind of balance in a system. Equilibrium is two categories, viz. stable and dynamic. Further, the stable equilibrium includes both homeostasis and steady states. A homeostatic system shows always activity, but it does not alter the balance between the system's components. A system in a steady state is also stable, but it may change in an orderly way. Dynamic equilibrium represents a process by which a slight disturbance causes a constant change in throughout the system.

4. Organization and Information in System

Normally system is well organized, that's why one can predict the possible amount of change in all set of elements of a system if a change occurred to an element is known but such prediction is not possible if the system is not organized. 'Information' is considered as 'the measure of the amount of organization' (as opposed to randomness) in the system. Good information means well-organized system. In addition, the word 'entropy' and 'negentropy' are associated with organization and information. Changing level of energy in a system creates a disorder, hence disorganizing the system. In other words, entropy (a measure of unavailable energy) is regarded to be a measure of disorder or disorganization of a system. In the contrary, the negative entropy or negentropy, on the other hand, is a measure of order in a system. The close system may have highest entropy hence making a system inactive. While interacting with the environment, it is good for the open system to have the optimum level of entropy creating a more complex system. As entropy brings disorder and negentropy brings order, these concepts can be used in different branches of geography like an ecosystem, river system, and socio-economic systems.

5. Structure of a System

Structure of a system depends on how element and components of a system are arranged and interlinked. Therefore, the structure may vary in terms of its shape and size. The structure could be hierarchal or parallel. For example, in Walter Christaller's Central Place Theory, settlements are hierarchal arranged in an urban system. Large urban cities in few number lie at the top of the order and small cities in large number at the bottom.

D. Common Relationships

The links of elements shape components and structure of a system. A different pattern of links between elements forms a variety of relationships, some of the most common relationships are being illustrated through Figure no. 3, 4, 5, 6 to 7.

1. **Cause and Effect Relationship:** This is the simplest relationship which is also called 'Series' relationship in which elements are connected by an irreversible link.

For instance, rainfall affects the rate of soil erosion but soil erosion directly does not have an effect on rainfall.

2. **Parallel Relationship:** When two elements affect third element making relationship called parallel one. For Example, rainfall and temperature affect vegetation and vegetation, in turn, directly or indirectly affects the amount of rainfall and local temperature.

3. **Feedback Relationship:** This is newly introduced relationship into analytical structures. Two elements get mutually affected. For example, farmers grow pulses (leguminous plants) which enriches nitrogenous fertilizer in the soil and in turn, production of pulses increases because of enriched soil.

4. **Simple Compound Relationship**: In a simple compound relationship, components are modified by itself and influenced by a set of other external components. Both processes operate simultaneously. For example, industries in India are removing their old technologies and adapting new foreign technologies to increase low-cost production in the competitive global market.

5. Complex Compound Relationship

This is the most complex relationship of elements in which elements of the internal and external environment is mutually affected and influences each other. In our environment, all real systems have complex compound relationships amongst their element and complement. Our ecosystem is the best example of a complex compound relationship, which is very difficult to interpret correctly.

E. Classification of Systems

On the basis of above-mentioned relationships, salient features, and characteristics, Systems can be classified as homeostatic, dynamic, self-regulatory, adaptive, controlled systems etc.

1. **Homeostatic System**: A constant balance maintained in a system is named as a homeostatic system. Such system by its constant operating environment restores its equilibrium or steady-state behaviour if it faces some external interventions. As per its nature, it resists an alternation caused by internal environment but if it faces new change, its processor restores previous equilibrium or steady-state. There are innumerable homeostatic systems in our environment. For example, the human body is a homeostatic system that maintains its equilibrium in its temperature at about 98.2 degrees Celsius. Temporarily, it might change but body againrestores equilibrium in its temperature.

Likewise, innumerable geographical systems operating in our environment are known as a homeostatic system which maintains equilibrium or steady-state. The geomorphic cycle of erosion is the homeostatic system. In the cycle of erosion, if any element like the amount of water, slope, suspended particles etc. changes, the entire system gets affected but with certain changes, cycle maintains steady-state.

2. **Adaptive System**: It is a system which has adaptive capacity to changing external environment. It's some characteristics are similar to the homeostatic system. This system sustains a constant operating environment to achieve the desired state which has been emerged because of certain change in external environment. For example, our socio- economic systems are becoming adaptive to climate change. New technologies are introduced in agriculture and energy sector in wake of climate change. The direction of the adaptive system depends upon the feedback, it is getting in form of increased or decreased productivity.

3. **Dynamic System**: It is different from both homeostatic and adaptive systems which experience some change over the period of time in achieving steady or desired states. The dynamic System shows a chain of continuous changes along with a line behaviour in the entire system over the period of time. For example, the vicious cycle of poverty and cumulative causation as Economic growth models.

4. **Controlled System**: It is a system in which elements or inputs can be regulated to achieve goals (results) for socio-economic development. Normally such kind of system lies in the close environment like laboratories. For example, scientists, doctors, and chemists conduct experiments to assess the impacts of certain chemicals as medicines on animals or human bodies in a controlled environment. In a study of system engineering and cybernetics (the study of communication and control mechanisms in machines and living beings), partially controlled systems are of great importance. Even in geographical subject matters like resource management, regional and economic planning, partially controlled systems can be created and applied. Example, economically backward

region can be developed by pushing huge investment in infrastructures hence creating employment opportunities for local people but such environments cannot be completely controlled. Therefore, partially controlled environments are of great importance for human well-being.

In conclusion system approach and its analysis offers a simplified theoretical and conceptual framework to study the subject matters of geography like study of landforms, river system, ecosystem, regional and economic planning and social and economic development etc. Geographers can apply this approach in all three stages of research viz., descriptive, analytical (explanation and seeking governing natural laws and undegrading orders in the real world) and predictive (how existing orders are likely to change in future?).

ADVANTAGES AND DISADVANTAGES

Application of system approach and analysis in geographical studied has various pros and cons. It discloses inherent information on current state, structures and dynamic behaviours of various geographical systems. Our open existing environment is so complex that it goes beyond our understanding. Therefore, system approach simplifies existing environment in order to make it easy for students to understand. In words, it is a technical tool to comprehend interaction between elements of any complex geographical structure in simplified ways. It also helps us to develop a variety of abstract geographical theories. More importantly, its mathematical languages like geometry and probability theory are widely used in solving numerous geographical problems like rising pollution and prediction of climate change and understanding affecting factors. Despite these advantages, this approach is criticised because its overemphasis on positivism and quantification social science (quantitative revolutions) and avoiding normative values (beliefs, attitudes, desires, hopes, fears). Nevertheless, system approach is still relevant in geography (Jiwan n.d.).

UNIT - 8: CRITICAL REVOLUTION IN GEOGRAPHY; HUMANISTIC GEOGRAPHY; RADICAL GEOGRAPHY; BEHAVIOURAL GEOGRAPHY

CRITICAL REVOLUTION IN GEOGRAPHY

Critical geography is based upon the notion that humanity has the potential to transform the environment. It challenges the dominant ideologies that characterise international political structures, hence contesting traditional categories and units of analysis in IR such as anarchy, security and the concept of the state. Critical geography is based upon the principle that questions about spatial relations, which refer to how an object located within a particular space relates to another object, are important because political behaviour is embedded in socio-political structures based on ideas about space. Following from this, if scholarship and political behaviour are ingrained in socio-political structures, an objective analysis of international politics becomes impossible. IR theory cannot reflect the global situation from a neutral standpoint. Critical geographers suggest that alternative ways of thinking about space have the potential to change fundamental ideas, theories and approaches that dominate the study of international politics. In turn, they hope that this alternative scholarship will help to transform international politics and reduce human inequality.

THE BASICS OF CRITICAL GEOGRAPHY

Critical geography emerged in the 1970s as a critique of positivism, which is a form of scholarship based upon the idea that the world exists independently of observers. Critical geography is rooted in neo-Marxism and draws upon the ideas of Jürgen Habermas and the Frankfurt school, who expanded upon ideas within classical Marxism by exploring how freedom from inequality could result from peaceful processes rather than revolutionary action. At this time, scholars began examining how dominant political structures and scholarship perpetuated existing political inequalities.

The end of the Cold War in 1991 saw new global economic developments, accompanied by changes in global demographics. In the early 1990s, the increasing importance of non-state actors such as non-governmental organisations and multinational corporations accompanied by increasing ethno-nationalism – whereby nations are defined on the basis of ethnicity rather than civil state membership – fostered new ideas about security and the role of the state.

Critical IR scholarship began focusing on how dominant theories like realism reinforced unequal power relations by favouring the states that dominated international politics. Drawing on the ideas of Ken Booth and Richard Wyn Jones from the Welsh school, they argued that human insecurity was perpetuated by existing political structures (Booth

1991 and 1997). From this, scholars began looking towards critical geography and Lefebvre's (1991) critical theory of space to examine how assumptions about space perpetuated these existing insecurities and inequalities. Two important scholars associated with this are David Harvey and John Agnew, who highlight how traditional conceptions of space decontextualise processes of state formation and cement traditional polarised conceptions of space between East and West, North and South, developing and developed countries in International Relations thinking (Agnew 1994; Harvey 2001 and 2006).

For example, think about how the world is represented on a standard political map. A political map represents the world in terms of individual states separated from each other by territorial borders. An absolute view of global space takes this mode of representation as fixed, meaning it would not consider the possibility of alternative ways of mapping the world. This fixed view also ignores how international politics changed throughout history, altering the shape of the global space as new states and international institutions emerged.

CRITICAL GEOGRAPHY AND INUIT VIEWS OF SPACE

The Inuit are a group of culturally similar indigenous people living in the Arctic regions of Alaska, Canada, Greenland, Denmark and Russia. Their view of territorial space is based upon cultural similarity and use of land for traditional hunting practices rather than nation-state boundaries. The spatial extent of Inuit occupancy of Arctic territory reaches across five states, illustrating their historic sovereignty over a large area. Yet political maps of the world do not represent this area as Inuit territory. Rather, the area that Inuit territory covers is broken down and subsumed within individual state boundaries. When Inuit territories were colonised by European, American and Russian powers, their territories became part of colonial nation-state territories and the Inuit became subject to the colonial state governments. Today, the legacy of colonialism can still be seen in representations of the international political space, as the majority of membership within international political institutions continues to be designated on the basis of sovereign states, resulting in the ongoing political marginalisation of the Inuit.

Without adequate representation at the international political level, Inuit concerns about security and environmental sustainability cannot influence international policy to the same extent that state governments can. Furthermore, the extent to which Inuit interests are represented in the decisions made at the international level is poor. This is especially so when

Inuit interests conflict with the interests of governments, such as over pipeline constructions through Inuit territory to transport oil between states. However, by adopting an Inuit perspective of territory that rejects nation-state delineation of the global space, critical geographers can offer alternative definitions of territory and provide more accurate representations.

When viewed from this perspective, the security risk to the wellbeing of people right across such a large area of the globe appears much more prominent than that afforded by most other IR theories. When scholars adopt traditional spatial definitions, they over-simplify the global space and, as we can see in this example, oversimplify the geographic extent of threats to human security. Furthermore, when scholars define space as existing solely of independent states, it limits the examination of the impacts of environmental disaster to simple comparisons between two or more nations, such as between Canada and United States. This undermines differences in the severity of impacts of natural disasters within particular regions of the world. Moreover, this traditional method of analysis also overlooks how the human security threats posed by environmental disasters are not evenly spread within individual state territories. For example, it downplays the fact that the Inuit living in Alaska are at risk of far greater disruption from the effects of melting sea ice than people living in other areas of the United States. It also downplays how coastal communities within Alaska are at a greater risk from the devastation caused by flooding and erosion than communities located within the interior of the state.

However, despite the success of Inuit representation in the Arctic Council, the vast majority of indigenous governmental bodies continue to fall outside the formal political representational structure in larger international climate change negotiations. The Inuit Circumpolar Council (ICC) is a United Nations- recognised non-governmental organisation that defines its constituency as Inuit populations in Greenland, Alaska, Canada and Russia. However, their participation is restricted at UN summits on climate change to that of 'observer' status as it is not a sovereign state – thereby constraining its voice. It is on this basis that the state system of representation within the United Nations climate summits can be argued to further marginalise indigenous groups like the Inuit. As representation is afforded on the basis of state territory rather than Inuit conceptions of territory it reinforces the decision- making power of the former colonial governments, enabling them to exercise greater control over international affairs, which hinders Inuit self-determination efforts.

The power of the Inuit to shape international political decision-making risks becoming further marginalised if IR scholarship does not critically question nation-state ideas about territory and representation. By bringing alternative conceptualisations of territory to the foreground, critical geography opens up a space for recognising and exploring alternative modes of representation that reduce inequality between indigenous people and state governments. If the Inuit are at greater direct risk from the impacts of global climate change, representational reform would enable them to have a greater voice in managing these risks.

CONCLUSION

By drawing attention to alternative ways that space can be imagined, critical geographers have sought to transform international politics and the global space. Critical geography highlights how issues of economics and climate change impact upon people and shows that the spatial effects of these processes differ to how they are dealt with by states, international organisations and within academia. The unique vantage point of critical geography provides useful ways to rethink what we know about International Relations in both theoretical and empirical terms. It challenges assumptions about space and territory, offers new conceptual and analytical tools and encourages students to question mainstream thinking.

HUMANISTIC GEOGRAPHY

Humankind as an agent of change on the earth's surface was first identified by Comte de Buffon as early as in the 18^{th} century. Inspired by his ideas, Immanuel Kant developed his physical geography that was essentially *'anthropocentric'* in nature and content. According to Kant, physical geography not only included the features visible on the earth's surface created by natural processes but also by human actions. Kant was also of the opinion that empirical knowledge could be obtained in two ways—either (i) through pure reason, or (ii) through the senses. Senses again could be divided into—(i) inner senses and, the (ii) outer senses. The world as perceived by the inner senses was the *seele* (soul) or *mensch* (man) while as perceived by the outer senses, was the Nature. The concept of Kant's anthropocentric geography was subsequently adopted by Carl Ritter. In his famous *'Erdkunde'*, he asserted that the central theme of geography was the element of reciprocity that is believed to have existed between the natural phenomena and humanity. Subsequently, Friedrich Ratzel in his *'Anthropogeographie'* set a framework for the systematic study of human geography and thus set a new trend in the subject. Prior to him, systematic geography only involved physical geography and, human geography was mainly confined within regional studies. His anthropogeographie was essentially a reflection of the Darwinian viewpoints and emphasized on the concept of natural selection that was used in the natural sciences.

The human approach in geography was greatly popularized by the French geographer Paul Vidal de la Blache in 1899 with his introduction of a new dimension to the possibilistic philosophy. Blache may rightly be called the *father of modern human geography*. He advocated *'genre de vie'*, a concept akin to human culture, inherited and developed over time to convert natural *'possibilities'* into elements of fulfilment. Nature was conceived as a mere adviser and humanity, an active force of change. Blache's possibilist philosophy was carried forward by Jean Brunhes throughout France and other parts of the globe. His main emphasis was on the exploitation of the earth by humankind for satiating human needs and desires.

In fact, it was the French historian, Lucien Febvre, who is actually credited for coining the term *'possibilism.'* In his *'Geographical Introduction to History'* (1922), Febvre accorded to the Vidalienne tradition of possibilism. He put forward that humankind emerged as a powerful agent of modifying the earth's surface through centuries of his accumulated labour and decision- making. In 1924, American geographer Carl O. Sauer propounded his *'landscape paradigm'* in which he highlighted on humans as agent of *'fashioning'* the natural landscape.

The discipline of geography underwent several paradigm shifts and revolutions over time. The 1920s witnessed the revival of the positivist philosophy after. The concept was however, introduced in the 1830s by August Comte. The aftermath was a theoretical and quantitative revolution in geography in the 1950s. Schafer's critique of Kant's exceptionalism and the introduction of his *'spatial organization paradigm'* opened the door for such revolution. However, in the 1970s there was yet another revolution in geography which was essentially anti- positivist in nature. It came to be known as the *'critical revolution'* as its origin was rooted in the criticism against the positivist-quantitative-spatial tradition of geography. The effort was on replacing the quantitative methods with a variety of humanistic approaches. This was supposed to ascribe a pivotal role to humankind in the subject particularly to *'human awareness, human consciousness and human creativity'* and freed human beings from geometric determinism. Thus, the modern humanistic geography was mainly an outcome of the growing dissatisfaction against the quantitative revolution.

Effort was made to revive the *'normative statements'* of values, attitudes, beliefs and so on. It aimed at *'verstehn'*, that is, understanding humankind within the surrounding environment in which humankind by using his rationality could improvise on the conditions of their lives.

The proponents of humanistic geography asserted that, humanistic geography should not be considered as an earth science in its scope and content. Instead of viewing geography as the study of the earth, it treated geography as the study of the earth as the home of humankind.

Hence, the main focus was on how humans perceived the place they inhabited through their thought processes, consciousness and experiences.

Humanistic geography sought to be more than a mere critical philosophy. Therefore, Anne Buttimer in 1978 attempted to resuscitate the Vidalienne tradition and asserted that any spatial unit should be studied from a *local perspective* (similar to Blache's concept of *'pays'*) and with a *historical approach*. This was possible because some affinity was discovered between Vidal's *le geographie humaine* and humanistic geography. But, there were grounds of departure between the two. Firstly, Blache considered human geography as a natural science and, many of Blache's work contained the elements of functionalism which the humanistic geographers renounced.

Humanistic geography also contained elements of *neo-Kantianism* and *pragmatism* in it owing to its emphasis on human consciousness and experience which were reflected in human actions and, which in turn, were directed towards alleviating human problems.

Though humanistic geography started on the same platform as of behaviourial geography, the two soon parted ways as, humanistic geography according to Entrikin concerned itself with the *'subjectivity'* of both the researcher and the reconnoitered. It digressed from the formal structures of behaviourialism which otherwise was supposed to have a strong connection with the positivist/spatial tradition and was rather considered an outgrowth of that tradition.

One of the first geographers to advocate humanistic geography was the Irish geographer, William Kirk as early as in 1951. He published his ideas in his essay, *'Historical Geography and the Concept of the Behaviourial Environment.'* But perhaps the time was not appropriate since by then, geography was greatly inspired by the positivist tradition to initiate the quantitative revolution. Later in 1976, it was Yi-Fu Tuan who argued for humanistic geography as concerned with people and their conditions. He opined that humanistic geography sought to achieve an understanding of the world through an insight into the human-nature relationship and the geographical behaviour of humankind as well as their perception about space and place. Geographical activities and phenomena were treated as the manifestation of human awareness and knowledge.

After the 1980s, humanistic geography advanced further from its early position of a critique of positivist philosophy to attack on structuralism. At the same time it developed an insightful methodology for empirical research. Two prominent streams of work were identified in humanistic geography. One stream tried to connect with the *humanities* by investigating knowledge that emanated from human feelings and experiences regarding being a human being on this planet. The other stream tried to connect with various philosophies of *human and social sciences.*

APPROACHES TO HUMANISTIC GEOGRAPHY

Humanistic geography was developed as a conceptual perspective that highlighted on the thorough understanding of human-environment relationship particularly on the basis of individual or group awareness and experiences regarding different spatial units and related geographical phenomena. The main emphasis was on humans as rational being with the power to think and perceive rather than as mere responders to stimuli as was presented within the positivist and behaviourial framework. According to Ley and Samuels, humanistic geography incorporated a wide range of philosophical approaches within it ranging from idealism, existentialism and hermeneutics to phenomenology; the connection with pragmatism has already been mentioned before. At the same it ascribed a central role to human beings and was a people's geography with human development as its principal objective.

Humanistic geography imbibed in it the philosophy of *existentialism* that urged on human quality and subjectivity. It was based on the doctrine of *'existence before essence'* which implied that humans existed first and, thereafter were responsible for their every action. It stressed upon personal freedom, personal decision-making and personal commitment. In other words, the purpose of humanistic geography in its affinity with existentialism, was to analyse the existential space as occupied by humans and the ways they defined their relationship with their space. This approach was essentially historical in that, it attempted to reconstruct space through the experiences of its denizens.

As a counter to the postulates of positivism, Leonard Guelke propounded the philosophy of *idealism* and urged the human geographers especially the historical geographers to probe into *what* humans, as decision-makers believed in and not *why* they believed. Thus, human geographers were not supposed to engage themselves in developing theories as, the pertinent theories that resulted in the geographical activities under study were already extant in human minds. Humanistic geography inspired by the idealist philosophy upheld that reality was basically a mental construct and a pattern of human behaviour actually reflected the underlying ideas. Idealism according to Guelke was based on *two* propositions---(i) a *metaphysical proposition* which asserted that an idea or a mental construct had a particular duration which was however, independent of material things and processes; and, (ii) an *epistemological proposition* which believed that knowledge was derived indirectly from the subjective human experience of the world and was an outcome of human thoughts and ideas. It upheld that the existence of a *'real'* world was actually mind-dependent.

Idealism was basically a sort of *hermeneutics* that dealt with the theory of interpretation and clarification of meaning. It developed in the German tradition of *'geisteswissenchaften' or* *human science*. The contention between the objectivity and subjectivity of human discourses led to *'double hermeneutics.'* Hermeneutics was applied in contrast to the positivist-spatial science methods as advocated by humanistic geography through, a presuppositional approach directed by social conscience. It provided an epistemology that aided in restructuring regional geography by speaking of the spatio-temporal aspect of a region. At the same time, it expressed its concern regarding any spatial unit with respect to its culture as developed by humans occupying it over time particularly language.

In the 1970s, another philosophy that was more popular among the human geographers than idealism was *phenomenology*. Though the term was first used by Sauer in the 1920s, it became widespread in the 1970s when Relph tried to introduce the approach within geography. The objective was similar to the above approaches---to present a critique of the positivist tradition. It presented an alternative to positivist philosophy that was based on the premise that there can be no objective world without human existence. Kirk in 1963 identified *two* different yet mutually dependent environments (i) a *phenomenal environment* that included everything on this planet; and, (ii) the *behaviourial environment* that was the perceived and experienced part of the former. Phenomenology in geography was concerned with the phenomenal environment the elements of which were considered distinctive for every human since, it was the outcome of individual perception and action. Therefore, the phenomenological approach in geography sought to explore how individual human being structured the environment in a subjective way.

THEMES AND METHODS OF HUMANISTIC GEOGRAPHY

Humanistic geography originated as a perspective against that form of human geography that was reduced to an abstract study of space and structures. Sometimes, humanistic geography could be used interchangeably with humanism because it accorded central role to humans. But precisely, humanistic geography was mainly concerned with the outcomes of human activities. According to Ley and Samuels, humanistic geography was based on three basic precepts:-

(i)anthropocentrism; (ii) subjectivity; and (iii) the concept of place.

Humanistic geography did not consider humans as mere *'economic man'* but attempted to investigate as to how geographical activities and phenomena were a manifestation of human awareness and creativity. As a propounder of humanistic geography, Tuan identified the following *five* major themes of humanistic geography:

- *Geographical knowledge or personal geographies*: Humans were to be treated as rational beings with the ability to think and perceive. The main task of the humanistic geographers therefore, was to study the ideas and thoughts that emanated from human minds since these ideas constituted geographical knowledge. Each and every human being possessed such knowledge though their perception varied. They utilized their geographical knowledge for their biological survival. Hence, geographical knowledge was conceived as personal.

- *Role of territory and creation of place identities*: As mentioned earlier, sense of place was an intrinsic aspect of humanistic geography. Every human being occupied and utilized some space with which they developed a strong sense of emotional bonding. Much of his biological needs were satiated in that space. Hence, a particular space constituted the territory of humans which was not only a confined area in its literal sense but a place with which human beings identified themselves. It was here where humanistic geographers stepped in to analyse how a mere spatial unit turned into a place identity for individual human being.

- *Crowding and privacy*: Crowding of a place resulted in physical as well as psychological tensions which were eased out by cultural, social institutions and infrastructures. In a similar way, privacy and seclusion also influenced the thought processes and actions of humans. Privacy was thought to be required by every individual. Within the private space individuals developed their own personal world.

- *Role of geographical knowledge in determining livelihood*: For sustenance humans engaged themselves in economic activities. They utilized their geographical knowledge to decide their economic activities. Thus, accordingly they planned their action for sustenance which was the essence of pragmatism. In doing so, they were in an position to distinguish between life-sustaining and life-destroying activities.

- *The impact of religion*: Religion was supposed to be subjective and associated with the normative elements of values, beliefs or ethics. Religion was conceived as the desire for coherence. The variation in this desire, which differed with individual persons and culture, provided a field of investigation for the humanistic geographers.

Four conceptual and methodological themes were identified as inherent of humanistic geography.

- According to humanistic geographers, human life and experiences were regarded as dynamic and multivalent that had cognitive, attitudinal and emotional elements attached to them. Humanistic geographers asserted that the task of a comprehensive human geography was to identify these elements and understand how they contributed to human experiences and actions, as well as, how each of these elements were connected to each other in a supportive or contradictory manner. This may be made clear in the words of Tuan that every individual human was at the same time a biological being, a social being and an inimitable personality and all these three aspects were believed to be a function of environmental perceptions, attitudes and values.

- Since human experiences were indefinable, humanistic geographers departed from the scientific methods employed under the positivist regime in which everything was to be explicated and verified empirically using statistical techniques. On the contrary, humanistic geographers adopted the ontological-epistemological perspective to encompass a much wider range of experiences; which would have created a framework within which the investigators would be able to study the experiences of their subjects with greater precision.

- The humanistic geographers advocated that humanistic geography should have originated from the self-knowledge and first-hand experience of the investigator. At the same time, it should have also incorporated the experiences of the *'others.'* The others could range from people, places, any natural phenomena or any aspect of human-environment relationship. This approach of humanistic geography brought them in sharp contrast with the objective approach of the quantitative paradigm in which the experiences of the researchers were greatly minimized. With regard to this, Tuan asserted that through an understanding of geographical experiences individuals developed a sense of environmental humility and acted more compassionately towards other humans and the place and the environment they occupied.

- Humanistic geography employed the usage of *two complimentary research methods*--- one that involved the *explanations of experiences*. It was based on a multitude of descriptive sources like first-hand experiences of individuals, archived reports and literature, evidence gathered through photography, films or any other forms of media. Its emphasis was to highlight the commonalities that existed in experiences related to a place or an environment. The other method that involved the *interpretation of the social world* was based on philosophical arguments rather than experiential evidences. It involved a wide range of philosophical traditions ranging from existentialism, pragmatism, idealism to post-structural Marxist approach.

CRITICAL APPRAISAL OF HUMANISTIC GEOGRAPHY

Human experience and human actions have always been the focus of humanistic geography. The central thesis of humanistic geography was provided by the criticisms rendered against positivism. It ardently highlighted upon human as *'living, thinking and acting being'* and insisted that human conditions could only be suggested through humanistic endeavours expressed in human attitudes, impressions and sense of place which otherwise could not be articulated through positivist methods.

However, humanistic geography was not free from criticisms.

- The first and the basic criticism rendered against the humanistic methods is that the researcher was not in a position to ascertain whether the real and the true explanations had been provided or not. It is true that humanistic explanations could not be established with certitude but this again provided a field of criticism by the positivist-quantitative approaches where everything could be verified empirically and thus had a greater certainty. In fact, the natural sciences whose methods were adopted by the positivist regime were mainly comprised of theories that were abandoned through further research which in turn enhanced the scope of study and resulted in more authentic and powerful theories. But with humanistic methods this was not possible.

- Secondly, on methodological grounds humanistic geography differentiated and distinguished between physical and human geography which diluted the core of the subject and gave rise to some sort of dualism in the discipline of geography. This dualism sometimes proved to be detrimental in the development of geography. Physical geography mainly dealt with inanimate objects and so its methods were mainly scientific and mathematically verifiable. On the contrary, since human behaviour was difficult to predict and varied over space and time, such quantitative techniques were not always applicable in human geography. However, humans as the prime focus of humanistic geography and physical environment of physical geography were not mutually exclusive but rather related and, could not be studied independent of the other.
- Humanistic geography was criticized as *'methodologically obscure'* since its main focus was on subjective rather than objective research. Humanistic geography was largely based on the experiences and perceptions of the humans which were mental constructs and were essentially abstract in nature with no practicality as such. Any method was acceptable to interpret the meanings of human experiences. Thus, humanistic geography had no sound or valid methodological base on which the theories developed by it could be successfully and authentically grounded.
- This gave rise to another criticism against humanistic geography that it had limited applied aspect. The investigator could have numerous interpretations of reality and in that situation it was really difficult to ascertain reality. Under such circumstances, it was rather challenging to identify the geographical problems and frame alleviating policies accordingly.

Though humanistic geography attempted to combine several philosophical traditions along with an incisive methodology, yet as pointed out by Entrikin, it failed to provide a suitable and viable alternative to the scientific methods. It was better described as some kind of critical philosophy that originated against the positivist tradition with a purpose to revive the *'humane'* element in geographical research.

The concept of place as enshrined in humanistic geography was static and exclusive. This was criticized by several post-structuralist geographers who presented a progressive and dynamic concept of space that was responsive to wider social and environmental contexts. The sense of place of humanistic geography was also questioned by the post- modernists on the ground that the distinction between perceived and real space was no longer valid in the world of booming hyperspace comprised of digital environments and virtual realities.

Humanistic geography has been subjected to criticisms and rejection by modern day geographical research due to its unscientific character and its associated gross inability to provide generalisations, laws and theories. However, since any philosophy is largely an outcome of thoughts originating in human minds, the importance of human ideas can in no way be undermined. It is true that post 1990s humanistic geography disappeared as a distinctive sub- branch of geography, but interests in humanistic themes still persists particularly among the phenomenological philosophers regarding the phenomena of space. Interestingly with time humanistic geography with its continued focus on human action, human beliefs and awareness; human interaction with their place in space and, the interpretation of that place within space, have adopted psychoanalytic theories. The objective behind this has been to do away with the criticisms regarding their obscured methodological and poor theoretical base. It had also started focusing on the increased interaction between human and physical geography particularly in determining the role of individuals' perception in creating the physical landscape.

RADICAL GEOGRAPHY

INTRODUCTION

In the Leftist ideological group, there were two sections – Left-liberal and Left-radical. Both the sections were concerned with inequalities, deprivation etc, i.e. problems pertaining to rich-poor divide. They were against capitalism. They both criticized Positivists because Positivism could not answer the questions of deprivation, dislocation, crime, problems of female issues, class differences etc.

Left-liberals were those people who want minor adjustment in society for the benefits of have- nots. But Left-radicals wanted to change the entire social order.

CONTEXT

Amid mass demonstrations against government's social policies, for which people came out on the streets of American cities, political radicalism through the revival of socialist parties happened due to certain reasons, as follows:

1. After the World War-II, there was a steady economic growth for two decades. Then an

economic slowdown or slump started to happen. In such a situation, people became conscious of the role of the government – its successful schemes and failed projects. It was widely felt that the fruits of economic growth were not shared equally, and a substantive chunk of society was facing economic hardship. This fueled grievance against the government, and that's how Civil rights Movements took place in almost all American cities during the late sixties.

2. Another point of discontent was Vietnam War where USA's aggression was viewed as an imperialist hegemonic pursuit.By and large, it was against the essence of democracy which the USA preached and practiced. It not only led to the destruction of lives and properties of Vietnamese but also led to the death of US soldiers fighting in Vietnam. American people, especially the youth, revolted against the government for Vietnam War. Student protests were not only limited to its place of origin i.e. USA, but it expanded to several European countries also.

3. Problems of Black population, who lived in the shabby physical environment, started to emerge. Such problems pinpointed the failure of economic growth-centric government policies, which was running under the profit maximization policies.

SOCIAL RELEVANCE REVOLUTION

Given this context, a reassessment of purpose and methodologies of natural as well as social sciences began. It was felt that human being and their environment as a part of the earth is the most important subject that natural and social sciences should enquire in details. Geographers, who were working on the themes of "optimum location" of infrastructural facilities, now started to focus on the physical and social environment that surround people. This phase of revolution in geography, after the much-acclaimed *quantitative revolution,* is known as "radical revolution" or "social relevance revolution".

A-THE RADICAL STREAM OF THE RELEVANCE MOVEMENT:

In fact, "radical revolution" emerged as a critique of *quantitative revolution*. During the 1950s, the philosophy of positivism and empiricism became very influential. Geographers, while interacting with other disciplines, also started applying various tools and techniques to analyze and explain the spatial variation of man-nature interaction. They got so engrossed in model building, that the theoretical approach towards looking at socio-economic problems was sidelined, and availability of data and application of techniques started to guide research procedures. In such a context, through "radical revolution", a new discourse started that reminded geographers it is theoretical understanding that shows the path of research through an exploration of suitable dataset and methods, and not the other way round.

Radical viewpoint started through William Bunge's work who wrote about Radicalist ideas in his book *Theoretical Geography* in 1962 and who founded Society for Human Exploration at Detroit in 1968. This Society urged geographers to undertake fieldwork in areas where poorest people live or the areas which are most backward and depressed. Such expeditions targeted to acquire firsthand and unbiased information of these areas so that a collective engagement with local people can bring meaningful inputs and bring about sound policy and planning framework. Few expeditions were carried out in Detroit. For providing training to aspirants who shown interest to participate in such expeditions, a course was opened at the University of Michigan. As university officials did not cooperate at the later stages, such expeditions were stopped in the USA. However, the expeditions continued in Toronto (Canada), Sydney (Australia) and London (England). Moreover, the Union of Socialist Geographers (USG) was established in 1974. Members of USG also participated in special sessions of AAG conventions and IBG annual meetings.

Radical ideas flourished in the hands of David Harvey and Richard Peet. Harvey wrote *Social Justice and the City* where he talked about Black people living in Ghettos. Richard Peet started to publish articles in a famous journal known as *Antipode* in Clarke University in Massachusetts in 1969. The issues in *Antipode* were quite revolutionary. They talked about urban poverty, discrimination against Blacks, feminism and cruelty against women, crime, deprivation, problems pertaining to minorities etc. Therefore, geography again got a breakthrough from its original systematic or regional approach when it started incorporating new social issues.

Due to increased poverty and inequality, especially poverty among the people of Ghetto and rural areas, Radicalists tried to perceive planning from a new viewpoint i.e. planning *with* the people rather than planning *for* the people. According to Harvey, geographers should consider the question as to who is going to control whom, in whose interest the controlling is going to be exercised and if it is exercised in the interest of people, who is going to take it upon himself to define that public interest.

Important features and objectives of the radical stream of relevance movement are following:

1. To expose the issues of discrimination, deprivation, inequalities, crimes, issues pertaining to health and mental degradation in the capitalistic society.
2. To pinpoint the weakness of Positivism and Quantitative Revolution in geography which emphasized geography as a spatial science and did not deal with the human issues?
3. To remove regional inequality
4. Radicalists opposed economic and political concentration, imperialism and nationalism.
5. They opposed superiority of a particular race.

6. They also prescribed revolutionary changes in the work order to develop a tension-free peaceful environment for all.

Radicalism was developed as a critique of existing models, because such models especially those adopting a positivist methodology which was supposed to be value-neutral, was helping the imperialistic forces to maintain the status quo. Geography was a tool for imperialists. Radicalism was critical to this system.

Radicalists always talked from the standpoint of those people who were not in control of means of production (land, labour, capital, organization), and they always supported the downtrodden group of society. Radicalist thinking always went against nationalism. Before Radicalist thinking, geography was a science which protected the ideology of majority who owns the means of production. Radicalists criticized this scenario. It also developed as a protest against data. Radicalists thought of a society which is controlled by all.

Radicalists like James Blaut (1970) attempted to link the issue of *imperialism* with capitalism. Imperialism denotes domination and subordination of one country to another – be it in economic or political terms. More developed countries had a tendency to control less developed ones, by exploiting natural resources and setting terms of trades often biased against less developed countries. Capitalistic countries, through this control, created a monopoly situation. Another issue was *ethnocentrism*, where an ethnic group was considered superior to another group(s). European ethnocentrism pointed towards the superiority of Europe over Asia and Africa, the superiority of Whites over non-Whites. It also showed the world how and why development persists in Europe. Blaut was very critical to this unicentric model and explained how Europe progressed at the cost of disrupting African and Asian countries. Imperialistic hegemony, through colonization of many African and Asian countries, paved the way for European countries to access billions of wealth. This led to the progress of Europe, in terms of expansion of industries, commercial activities, education, and technology. As the entire idea of racial superiority and ethnocentrism was based on certain prejudices, Radicalists opposed it.

Radicalists also opposed the way females were oppressed in developing and developed countries. Females were found to have an unequal role in terms of decision-making in households. They were systematically exploited, as their role was defined from a male perspective. They had relatively less mobility, and their role was defined to be restricted within household – cooking and taking care of children and so on.

To a certain extent, Radicalism was linked with *anarchism*. Anarchism called for the removal of state, and its replacement by voluntary groups of individuals. These individuals could work without external pressure and maintain social order. In a way, anarchism promoted individual liberalism and socialism. Peter Kroptokin and Elisee Reclus elaborated on the way by which such social orders can be maintained. Kroptokin attacked capitalism on the ground that it increased competition and inequalities. He commented that mutual cooperation and support help a community or a society to live peacefully. Cooperation based production, decision-making at grassroots level, the spread of democracy, greater integration of short-distanced workspace and living space were some of the ideals many Radical geographers followed.

B-THE LIBERAL STREAM OF THE RELEVANCE MOVEMENT:

Liberalism, although beliefs in democratic capitalism, advocates executive actions for minimizing social and spatial inequalities in the levels of human well-being. It shows a commitment towards ensuring a basic minimum level of standard of living for all. In this context, it prescribes state action in helping less privileged section of the human society. Statistical techniques, involving multiple variables, were applied to map levels of human being (Thompson and associates, 1962). The work done by Smith (1973) and Knox (1975) are often referred in geographical literature. Measurement and mapping of variables related to human well-being became important, and such variables were categorized into three sub-sets- "physical needs" (nutrition, shelter, and health), "cultural needs" (education, leisure, recreation and security), and "higher needs" (through surplus income). Their works show that geographers can play a significant role in informing policy-makers about the spatial implications of inequalities so that better decisions can be taken for improving policies and schemes further. Another part of these works is raising awareness among citizens so that they become better informed on welfare issues. Cox (1973) and Massam (1976) looked at how efficiently public services can be provided, by redrawing of administrative boundaries or changing the location of public facilities. In his seminal book, *Human Geography: A Welfare Approach*, David Smith (1977) focused on "who gets what, where and how", and this reoriented the goal of human geography towards making a society where spatial malpractices and injustices are done away with. Therefore the "distribution" of fruits of economic growth emerged as an important issue.

Moreover, this stream of social relevance movement advocates that our surrounding environment should be looked after well. Therefore, issues like environmental degradation-conservation- management are often discussed.

TOWARDS MARXIST GEOGRAPHY:

A great contribution of guiding geography towards Marxism happened through the works of David Harvey. In his book on Ghettos located in American cities, he pointed towards roots of problems that lie in capitalism. According to him, the capitalistic system created such a market-based mechanism, that regulates land use, and this is biased against the poor Black population. He argued that once a geographer adopts Marxist approach towards looking social problems, he or she cannot detach himself or herself. That's why a political awareness is generated within them, and they get actively involved in making a society with more justice. Harvey's influence was so strong that some practitioners of social relevance research started adopting a Marxist approach. Nowadays, radical geography is more aptly known as Marxist geography.

CRITIQUE:

Social relevance movement, especially radicalism was able to usher in some fruitful changes in the methodological discourse of geographical studies. These are:

- From the rhetoric of quantitative technique based analysis of geographical attributes, it reoriented human geography towards prominent social and environmental issues, thereby broadening the scope of geography to interact with other disciplines of social sciences.

- The classical tradition of fieldwork in a small region was altered in the sense that more in-depth and participatory planning oriented studies were encouraged. This fieldwork entailed new pattern where respondents were involved in the process of surveying. One needs to understand that this new pattern was quite challenging. The expeditions, promoted by the Society for Human Exploration, could not go on extensively due to multiple reasons (including existing power structure), even though it received a certain chunk of academic interest.

Some of the limitations or weaknesses of radicalism are:

- First criticism came from Russians who claimed themselves as true Communists and Marxists. Radicalism was entirely an American enterprise. Though Radicalists in America talked about the social change they never talked about an armed revolution which is a basic component of Marxist ideology.

- The theoretical base of radicalism was very weak. Basically, they were dependent on other social sciences. Whatever Harvey discussed in *Social Justice and the City* were basically sociological, political or economic analysis.

- Though the topic of radicalism was varied, the techniques and methodologies were not very path-breaking.

- Radicalism gave over-weight to Marxism. Geography, by virtue of its subject matter, is a spatial science. It cannot be explained totally with the help of Marxist thinking.

- The ultimate question was 'Who will guard the guardians'? Even socialist governments following models of Marx could not solve problems of the oppressed class.

- Humanistic geographers criticized radicalists because the former gave more emphasis on people, not as an ideology like the Radicalists. Humanistic geography says that geography cannot be explained through any generalized theory. It is human-specific.

- Positivists criticized radicalism because they don't take help from any empirical science.

- After the fall of USSR and East European nations in the end of 1980s, the worldwide impression was that socialism has no value and capitalism has won its final victory.

Therefore, geography is essentially a locational science which is based on empirical positivist values, which is the tool of capitalists.

BEHAVIOURAL GEOGRAPHY

By the mid-1960s use of statistical techniques in research for precision has been largely accepted by geographers. The duality of systematic versus regional geography was resolved as both were now accepted as important components of the discipline though interdependent and equally useful. It was increasingly realized by the geographers that the models propounded and tested with the help of quantitative techniques, provided poor descriptions of geographic reality as well as the man-environment relationship. Consequently, progress towards the development of geographical theory was glaringly slow and its predictive powers were weak. Theories such as Central Place Theory, based on statistical and mathematical techniques, were found inadequate to explain the spatial organization of society. The economic rationality of decision-making was also criticized as it does not explain the behaviour of man. It was a psychological twist in human geography which emphasized the role of subjective and decision-making processes that mediate the association between environment and spatial behaviour of man. It can be said that the dissatisfaction with the models and theories developed by the positivists, using the statistical techniques which were based on the 'economic rationality' of man led to the development of behavioural approach in geography.

The axiom of 'economic person' who always tries to maximize his profit was challenged by Wolpert. In his paper entitled '*The Decision Process in Spatial Context'*, Wolpert (1964) compared the actual and potential labour productivity of Swedish farmers and came to a conclusion that optimal farming practices were not attainable. He concluded that the farmers were not optimizers but, satisfies. Thus human behaviour was seen to be a product of decision-making and it was a human tendency to have incomplete information, to make imperfect choices and even then be satisfied with sub-optimal options.

BEHAVIOURIALISM IN GEOGRAPHY

Behavioural geography banks heavily on 'behaviouralism'. Behaviouralism is an important approach which is largely inductive, aiming to build general statements out of observations of ongoing processes. The essence of behavioural approach in geography lies in the fact that the way in which people behave is mediated by their understanding of the environment in which they live or by the environment itself with which they are confronted.

In behavioural geography, an explanation for the man-environment problem is founded upon the premise that environmental cognition and behaviour are intimately related. In other words, the behavioural approach has taken the view that a deeper understanding of man-environment interaction can be achieved by looking at the various psychological processes through which man comes to know the environment in which he lives, and by examining the way in which these processes influence the nature of the resultant behaviour.

One of the most interesting and applied aspects of behavioural geography was work examining the human perception of environmental hazards. The pioneering work by Robert Kates (1962) on floodplain management is one of the bases of this approach. He states the manner in which human beings perceive the uncertainty and unpredictability of their environment play a significant role in the process of decision making. He developed a scheme that had relevance to a wide range of human behaviour. This scheme of Kates was based on four assumptions –

1. Men are rational while taking decisions.
2. Men make choices.
3. Choices are made on the basis of knowledge.
4. Information is evaluated to pre-determined criteria.

Subsequently, Kirk (1952-1963) supplied one of the first behavioural models. In his model, he asserted that in space and time the same information would have different meanings for people of different socio-economic, cultural and ethnic backgrounds living in a similar geographical environment. Each individual of a society reacts differently to a piece of information about the resource, space, and environment. This point may be explained by citing the following example.

The highly productive Indo-Gangetic plains have different meanings for different individuals belonging to a various caste, creed and religion. Jats, Gujjars, Ahirs, Sainis, Jhojas and Gadas living in the same village perceive their environment differently. A Jat farmer may like to sow sugarcane in his field, a Gada and a Jhoja may devote his land to sugarcane, wheat and rice, an Ahir ma Saini is invariably interested in intensive cultivation, especially that of vegetables. For a Saini (vegetable grower), even five acres of arable land may be a large holding, while a Jat who uses a tractor considers even 25 acres a small holding. The perceived environment of each of these farmers living in the same environment thus differs from each other both in space and time.

The aspect which was most enthusiastically adopted by geographers from behavioural analysis was the concept of *mental maps*. The paper of Peter Gould (1966) was the seminal contribution in this regard. He points out that since decisions on location were guided by the manner in which a human being perceives the environment, it becomes essential for a geographer to have a mental image of how one perceives his environment while making decisions. Therefore, *mental maps* are not just images or maps but an amalgamation of information and interpretation that a person has on a particular thing as well as how he or she perceives that place (Johnston, 1986). This was further developed by Gould (1966), Downs (1970), Downs and Stea (1973), Gould and White (1974) and Saarinen (1979) through their writings.

Gould opines that *mental maps* are not only means of examining a person's area of a spatial preference but also provides insight into the processes which led to that particular decision. He states that *mental maps* may provide a key to some of the structures, patterns and processes of man's work on the earth surface. The conceptual framework provided by Downs (1970) has been illustrated in Figure 1. This framework proposes that information from the environment (real world) is filtered as a result of personality, culture, beliefs, and cognitive variables to form an image in the mind of a man who utilizes the environment. On the basis of the image formed in the mind of the utilizer about the environment, he takes a decision and uses the resources to fulfil his basic and higher needs. Downs' framework also suggests that there exist an 'objective' and a 'behavioural' environment.

Pred (1969) presented an alternative to this inductive approach of behaviouralists on theory building on 'economic man'. In his work *Behaviour and Location,* he proposes a *behavioural matrix* (Figure 2) to give a structure in which decisions of locations can be analyzed. The axes of the matrix are quality and quantity of information available and the ability to use that information; man as an economic being is at the right-hand corner. As there is variation in the quantity and quality of the information, the position of man on the axis would also change. His position would reflect his aspiration levels, experience and even norms of the group to which he may belong. His further states, that even same individual would not be in the same position as his decisions may vary over time as spatial patterns are never static in nature.

During the 1970s, a range of related personality assessments, such as personal construct theory and the semantic differential were employed, and in this work geography and psychology became close neighbours (Aitken, 1991; Kitchin, Blades and Golledge, 1997). In particular, this productive interdisciplinary relationship was developed through the annual meetings of the Environmental Design Research Association and in the pages of the new journal, *Environment, and Behavior*. Since that period, behavioural geography has continued to diversify, even if its position has been less elevated than in the 1960s and 1970s when many disciplinary leaders worked in this sub-discipline. More recent research has included analysis of environmental learning, spatial search, developmental issues in spatial cognition and cartography and Golledge's (1993) important work with the disabled and sight-impaired. But some of the lustre has left the field. In part, this may be related to the methodological sensibilities of post-positivist human geography. In part, it is due to the growing conviction of the inherently socialized nature of geographical knowledge, which challenges the individualism of psychological models. In part, it emanates from a suspicion of the adequacy of an epistemology of observation and measurement that may leave unexamined non-observable and non-measurable contexts and ideological formations. Nonetheless, behavioural geography has a continuing legacy, comprehensively itemized and integrated into the massive compilation of Golledge and Stimson (1997).

THE OBJECTIVES OF BEHAVIOURAL APPROACH WERE

1. To develop models for a human phenomenon which would provide an alternative to the spatial location theories developed under the influence of positivism.

2. To define the cognitive (subjective) environment that determines the decision-making process of humans;

3. To come up with psychological and social theories of human decision-making and behaviour in a spatial framework;

4. To change the emphasis from aggregate populations to the disaggregate scale of individuals and small group

5. To search for methods other than those popular during the quantitative revolution that could uncover the latent structure in data and decision-making;

6. To emphasize on procession rather than structural explanations of human activity and physical environment;

7. To generate primary data about human behaviour and not to rely heavily on the published data;

8. To adopt an interdisciplinary approach for theory-building and problem-solving.

The fundamental arguments of the behavioural geography to achieve these objectives are that:

(i) People have environmental images;

(ii) Those images can be identified accurately by researchers; and

(iii) There is a strong relationship between environmental image and actual behaviour or the decision-making process of man.

The behavioural paradigm has been shown in Figure 3. In this paradigm, man has been depicted as a thinking individual whose transactions with the environment are mediated by mental processes and cognitive representation of the external environment. In geographical circles, this concept is derived primarily from the work of Boulding (1956) who suggested that over time individuals' developmental impressions of the world (images) are formed through their everyday contacts with the environment and that these images act as the basis of their behaviour.

SALIENT FEATURES OF BEHAVIOURAL GEOGRAPHY

The salient features of behavioural geography are discussed in the following section:

1. The behavioural geographers argued that environmental cognition (perception) upon which people act May well differ markedly from the true nature of the real environment of the real world. Space (environment) thus can be said to have a dual character:

(i) As an objective environment—the world of actuality—which may be gauged by some direct means (senses); and

(ii) As a behavioural environment—the world of the mind— which can be studied only by indirect means.No matter how partial or selective the behavioural environment may be, it is thismilieu which is the basis of decision-making and action of man. By behavioural environment, it is meant: reality as is perceived by individuals. In other words, people make choices and the choices are made on the basis of knowledge. Thus, the view of behaviour was rooted in the world as perceived rather than in the world of actuality. The nature of the difference between these two environments and their implications for behaviour was neatly made by Koffka (1935-36) in an allusion to a medieval Swiss tale about a winter travel.

2. Secondly, behavioural geographers give more weight to an individual rather than to groups, or organizations or society. In other words, the focus of the study is the individual, not the group community. They assert that research must recognize the fact that the individual shapes and response to his physical and social environment. In fact, it is necessary to recognize that the actions of each and every person have an impact on the environment, however, slight or inadvertent that impact may be. Man is a goal-directed animal who influences the environment and in turn, is influenced by it. In brief, an individual rather than a group of people or social group is more important in a man-nature relationship.

3. The behavioural approach in geography postulated a mutually interacting relationship between man and his environment, whereby man shaped the environment and was subsequently shaped by it (Gould, 1980).

4. The fourth important feature of behavioural geography is its multidisciplinary outlook. A behavioural geographer takes the help of ideas, paradigms, and theories produced by psychologists, philosophers, historians, sociologists, anthropologists, ethnologists, and

planners. However, the lack of theories of its own is coming in the way of rapid development of behavioural geography. Therefore, one can say that the behavioural approach in geography is a fruitful one and it helps in establishing a scientific relationship between man and his environment. The broad scope of behavioural geography is remarkable even by the standards of human geography.

CRITICISMS

There are, however, overall, biases in content towards urban topics and towards developed countries. One of the main weaknesses of behavioural geography is that it lacks in the synthesis of empirical findings, poor communication, inadvertent duplication, and conflicting terminology. In behavioural geography, the terminology and concepts remain loosely defined and poorly integrated, primarily owing to the lack of systematically-organized theoretical basis.

Another shortcoming of behavioural geography lies in the fact that most of its data are generated in laboratory experiments on animals and the findings are applied directly to human behaviour. Koestler (1975) pointed to the danger of this strategy, in that behaviouralism "has replaced the anthropomorphic fallacy—ascribing to animals human faculties and sentiments—with the opposite fallacy; denying man faculties not found in lower animals; it has substituted for the erstwhile anthropomorphic view of rat, a rato-morphic view of man". In short, behaviouralist theories are elegant but unhelpful when it comes to understanding the real world man-environment interaction.

Behavioural geography has too often put too much emphasis on ego-centred interpretations of the environment. Specifically, scholars are critical of two assumptions on which a great deal of behavioural research in geography is based. The first assumption is that there exist identifiable environmental images that can be accurately measured. It is not clear whether an environmental image can be extracted without distortion from the totality of mental imagery. Moreover, not enough effort has gone into checking and validating the methods by which images are elicited.

The second critical assumption is that there exists a strong relationship between revealed images or references and actual or real-world behaviour. The main objection to this assumption is that it is an unfounded assumption because extremely little research has been undertaken to examine the congruence between image and behaviour.

Another significant deficiency in behavioural geography has been the gap between theory and practice. This has been most noticeable over the question of public policy. In fact, behavioural geographers remain observers rather than participants. There is a serious lack of knowledge of planning theories and methods amongst behavioural geographers, which is an impediment to more active involvement.

It is a barrier that can be removed only by developing the requisite understanding of the planning processes; it cannot be camouflaged by noble sentiments and moral tone. For instance, it will be only rarely that a small survey carried out upon a sample of students will supply the basis for far-reaching policy recommendations, yet the final paragraphs of many such works contain this seemingly obligatory element. Despite several constraints and methodological limitations, behavioural geography is now widely accepted within the positivist orientation. It seeks to account for spatial patterns by establishing generalizations about people-environment interrelationship, which may then be used to stimulate change through environmental planning activities that modify the stimuli which affect the spatial behaviour of us and others.

The research methods of behavioural geography vary substantially but the general orientation—inductive generalization leading to planning for environmental change— remains. Eventually, it is hoped, a 'powerful new theory' will emerge. Golledge argued that substantial advances in understanding spatial behaviour have already been made by studying 'individual preferences, opinions, attitudes, cognitions, cognitive maps, perception, and so on—what he terms processes variables.

UNIT -9: WELFARE GEOGRAPHY

INTRODUCTION

Welfare geography is an approach to geography where the emphasis is on spatial inequality and territorial justice. Destined up with the rise of radical geography in the early 1970s, welfare geography stresses the need to identify and explain the existence of crime, hunger, poverty and other forms of discrimination and disadvantage. Welfare geography sought to reveal who gets what, where and how. This early work was largely descriptive and developed the abstract formulation used in welfare economics, grounding it empirically but maintaining the use of algebraic representations. It provided a basis for evaluation. Current welfare configurations, regarding who gets what, where and how, could be judged against alternatives. This preoccupation with description eventually aimed to match, and then superseded, by work on the processes through which inequality is shaped.

Marxist economics replaced neo-classical economics as the basis for illustrative analysis, which takes place at two different levels. The first involves understanding how the whole social, economic and political system functions, and teasing out universal tendencies. In the case of capitalism, this level of analysis reveals that inequality is endemic. Uneven development is the spatial imprint, the geographical result of the restlessness of capitalism as a system. The second level of explanation attends to the details of particular social, economic and political systems; for example, how housing policy under capitalism advantages some people in some places and disadvantages other people in other places. The analysis of the politics behind these policies has recently been strengthened as part of renewed interest in the relationship between social justice and the state. Accompanying an attention to the restructuring of the welfare state, which characterizes much of this recent work (Peck, 2001), have been endeavored to theorizing a relational ethics of care. Illustration of feminist theory, this work seeks to uncover the social relations behind construction of care and justice. Understanding politics as an integral part in the daily deed, the emphasis is on the connections and relations rather than the difference between categories, such as private and public, state and market (Smith and Lee, 2004).

Welfare geography focuses on the connection between the spatial variation of need and structures of a provision in the creation of geographies of welfare (Smith 1973). A rather late arrival of welfare approach in humanities and social sciences and particularly in geography has several political, historical, and psychological reasons, e.g., the Vietnam War, crime explosion, environmental degradation. The manifestation of social, political and economic injustice through these crises in cities and towns led a group of social scientists to come up with a new idea and promote the radical approach. Especially, with geography, the issue of distribution was taking new urgency (Smith, 1977). Before the dawn of the Quantitative revolution, geography, like all other main sister disciplines from

Humanities and arts, faced many philosophical and methodological problems. Geography did not progress as a well-regulated discipline.

In the recent years, geographers have, however, adopted new strategies by restructuring their courses of the study and designed the themes around contemporary issues like socio-economic development, rural-urban studies, making the subject a primary source of awareness of local surroundings and regional milieu. During the last five decades, the subject matter of geography has experienced immense changes in the subject-matter, philosophy, and methodology. The issues of primary concern on which the geographers are concentrating nowadays include hunger, poverty, racial discrimination, pollution, environmental pollution, social inequality or injustice and use and the overuse of depleting resources, etc.

Some of the leading works and issues which have been useful in the public policy making are Black-Ghetto, Geography of Crimes and Geography of Social Well-being. The quantitative revolution of the 1960's infused a vigor into geography, which was vastly essential for the indepth and comprehensive analysis required in any public context and the formulation of proposals for public policy.

Scientific revolution paved its way in geography in the early 1960s. The pragmatists introduced the use of scientific methods (positivism) for finding solutions to the problems faced by human beings. It is with this intent that scholars like David M. Smith have embraced the welfare approach while debating the problems, prospects and the future scenario of human geography.

The welfare approach in geography has been defined differently by some eminent scholars of geography. Mishan was of the view that, "theoretical welfare geography is that branch of study which endeavors to formulate positions by which we may be able to rank, on the scale of better or worse, alternatives in the geographical situation open to society." In the spatial context, Smith defined welfare geography as the study of "who gets what, where and how."

The geographers whose prime concern are the problems of society and are trying to formulate more realistic plans for public policy by giving the description and explanation of the phenomena. Through such analysis, they evaluate their plans and suggest suitable strategies for the balanced development.

The explanation involves the empirical identification of territorial levels of human development and the human condition. This is a major and instantaneous research area in which astonishingly little work has been done in India and other developing countries as well as developed Explanation covers the how? It involves in identifying the cause and effect relationship links among the different activities undertaken in society, as they contribute in determining who gets what and where. This is where the analysis of the economic, demographic and social patterns mentioned above logically fits into the welfare structure.

Geographical distance and ease of understanding mean that some people will be enjoying the better place to for advantages or disadvantages, whether the structure is a road, railway, hospital, school, theatre, community hall, cinema, park or a recreational place. Therefore, locational decisions and comprehensive plans for spatial allocation of resources must be made with utmost care and dedication, if the benefits and penalties are found to be proportional among the population in a more or less predictable and reasonable manner. In such public policy decision making, geographers' role becomes authoritative as they have the necessary expertise in the Spatio-temporal analysis of any such phenomena.

Spatial allocation problems are related to the identification of priority areas, planning routes, the location of factories or other sources of employment, the spatial arrangement of facilities providing medical care, housing complexes, shopping centers and allocation of land for different urban and recreational uses. Each of these decisions could be made in some ways, and every decision can have a different influence. Geographers by their expertise can build up more sophisticated knowledge and models of the process of development. This involves unscrambling and complex networks of economic, social and cultural relationships and also the ecological relationships in equilibrium, so easily disturbed by ill-conceived 'developmental' projects.

Geographers by sharing out, analysis, and synthesis of space can contribute, successfully, meaningfully and more efficiently to the formation of the policies for the public, property, etc. In developing countries like China, India, and Brazil there is relatively a high degree of internal inequality. On the other hand in the Third World nations, wealth and power and other facilities of public interest are still largely in the hands of urban elites or big landlords. In India, more than 50 percent of the population is still below the poverty line and on the contrary over 50 percent of the total national assets are in the hands of only a dozen families. Moreover, in India, most of the economic activities are concentrated in metropolitan cores, although still, more than 70 percent of the total population is residing in the rural areas. The urban-based industrial and social infrastructural policy adopted by planners is widening the already wide gap on the one hand, between the rich and the poor and on the contrary between rural and urban population. The highly advanced and developed countries like U.S.A., Russia, Australia, and Japan also have spatial disparities in levels of human development. In the United States, the overall material standard of living is higher than anywhere else in the world. Millions of Americans, especially Negroes (black people), live in poverty and social denial in ghettos (city slums). In many parts of the rural south of U.S.A. (Texas, Georgia, etc.) the living conditions of some people are as bad as anywhere in the African continent. In these ghettos, the rate of crimes like drug addiction is very high.

The perseverance of widespread poverty in American slums, the most affluent society in the world is a paradox which underlines the failure of economic growth under a capitalist system to uplift the lives of all people to a current standard of decency. In 1976, according to the U.S. Census Bureau, about 12 percent (26 million) Americans were below poverty line.

One of the opinions put forward by the capitalist for the existing regional inter-regional and intra-regional disparities is that human beings are not born equal, and hence they cannot be equal in their societies. This situation gets further serious if the social, political and economic organization is intended or formed with an urban-biased or rich people-centered policy. The planners with the help of geographers can construct general social amenities which can benefit all sections of the society.

Geographers, however, cannot be the cure for all the ills, inequalities and socio-economic imbalances that are persistent. Geographers can analyze the spatial dimension of environmental problems, natural hazards and more particularly they know how to handle, analyze and interpret spatially distributed data. This consciousness of and facility of tackling the spatial dimension, which is a major component of all problems of resource and environmental management, is something not provided by those in other disciplines and have a tendency to be overlooked if a geographer does not arrange it. A welfare society needs better sharing of commodities, better distribution of commodities and better of means of manufacture among individuals (groups or classes) and places. All these things are more easily attainable if geographers who are dealing with the man-environment interface and elaborate the spatial distribution of phenomena are actively involved in the procedure of planning and formulation of public policies at different levels of development, i.e., the local, regional, national and international levels.

In countries like Sweden, the Netherlands, France, Norway, Israel, Denmark, U.S.S.R., Australia and New Zealand where geographers in collaboration with the scholars and scientists of other fields to design public policies. Which is effective and beneficial and reaching all sections of the societies. Similarly, Geographers in India can also provide practical proposals for solving the various social, economic and infrastructural problems that are caused by rapidly increasing population.

DIFFERENCE BETWEEN WELFARE HUMAN GEOGRAPHY AND HUMANISTIC GEOGRAPHY

By their efforts, geographers can bring the causal relationships between inequality arising between the spatial organization of society and social structure. Public policies about restructuring and rearrangement can be designed properly by the experts through planning.

Although, human geography has appeared from earth sciences and has persistent links with physical geography. Hitherto the core aim of this particular branch of knowledge is to examine the various problems of different social groups about their environment. At present, especially after the

1960s, the geographers have adopted welfare approach as a go-to approach for the study of the human behaviour. The welfare approach, in fact, emerged as the response to quantitative revolution, spatial science, positivism, and model-building which was thought to be unsatisfactorily concerned with contemporary glitches of human societies.

The 1970s saw a chief redirection of human geography in the direction of 'welfare' issues such as hunger, deprivation, malnutrition, poverty, crime, distribution of income, assets, and access to social services (e.g., education and healthcare). This corresponded to a major change in social concern, from constrained economic conditions of development or progress to wider aspects of the quality of life. Spatial distribution of phenomena and distributional issues have presumed new importance in the present era of slow economic growth, for in these conditions policies of distribution in favor of the poor or socially deprived can be instigated only at the expense of the rich or better-off members of society. This is also known as Pareto optimality—a condition in which it is not possible to make some people rich without making others poor. The Pareto model assumes that one society has touched the edge of production possibilities, i.e., if there is no more growth; the poor cannot be made rich unless at the expense of the rich.

The basic emphasis of the welfare approach is who gets what, where and how. The 'who' refers to the population of the area under review (a city, region or country, or the entire world), subdivided into groups by caste, class, race, or other relevant characteristics like religion. The 'what' refers to the different goods and bads enjoyed or tolerated by the population, in the form of services, commodities, social relationships, environmental quality, and so on?

The 'where' reflects the fact that living standards differ according to the place of residence. The 'how' refers to the procedure whereby the observed differences arise. The initial task posed by the welfare approach is descriptive. The present state of society concerning the fact that who gets what? Where? Maybe signified by the extension of the abstract interpretations of welfare economics, and hence, the practical objective is to give these empirical substances to the people. In a spatially disaggregated society, the general level of welfare may be written as:

According to the Dictionary of Human Geography edited by R.J. Johnston, D. Gregory and David M. Smith (1994), "in a spatially disaggregated society, the general level of welfare may be

written as:

$W = f(S1 \ldots\ldots Sn)$,

Where S is the level of living or social well-being in a set of n regional subdivisions. In other words, it can be said that welfare is the function for the distribution of goods and bads among different groups of the people as defined by the area of residence. Social well-being may be defined regarding what people get, as follows:

$$S=f\,(X1\ldots Xm),$$

Where 'X' represents the quantity of them goods and bads consumed or experienced. Social wellbeing may also be expressed regarding the distribution within the area in question:

$$S = f\,(U1\ldots Uk),$$

Where 'U' is the level of well-being, 'utility' or satisfaction of each of the k population subgroups. In all the above expressions, the terms may be weighted differentially and joined according to any function, to denote the combination of territorial or regional levels of wellbeing, goods, and bads that maximize the objective function (W or S)."

For identifying a discrepancy in territorial distribution, development of social indicators is of great importance. Such indicators may be as follows: housing, income, education, employment, social orders or social participation, etc.

The welfare approach found Neo-classical economics as the least suitable one to explain social inequality. The Marxian economics provides a useful tool for analyzing social problems such as housing, income, education, employment, etc., because capitalism has an inherent tendency of to create disparity. The second level of explanation deals with the process of how specific elements of a socio-political and economic system operates. D.M. Smith (1977), in his Human Geography: A Welfare Approach, first suggested the approach which later amalgamated with other approaches of geography dealing with the issues of inequality. The issues dealt by welfare geography demand an interdisciplinary approach of the highest order. Moreover, in a rapidly changing era of globalization where the developing South stands deprived vis-a-vis the advanced North, there has been a transformed interest in welfare geography.

WELFARE AND SOCIAL WELLBEING

The welfare geography approach deals with the issues related to inequality and injustice. The approach grew up as a reaction to the quantitative and model-building traditions of the 1960s. In the 1970s there was a major redirection of human geography towards social problems, viz., poverty, hunger, crime, racial discrimination, access to health, education, etc. The issues such as the distribution of the fruits of economic development received attention mainly as a result of dramatic socio-political changes in Eastern Europe and South Africa. Therefore, the basic emphasis of welfare geography is on who gets what, where and how. The 'who' suggests a population of an area under review (a city, region or nation)? The 'what' refers to various facilities and handicaps enjoyed and endured by the population in the form of services, commodities, social relationships, etc. The 'where' refers to the differing living standards in different areas? Moreover, 'how' reflects the process by which the observed differences arise. The empirical identification of inequality in territorial distribution involves developing social indicators. These may combine particular elements of social well-being in a composite manner.

Conditions that may be included are income, wealth, employment, housing, environmental quality, health, education, social order (i.e., the absence of crime, deviance and other threats to social stability and security), social participation, recreation, and leisure. Alternatively, the focus may be on individual aspects of social well-being, such as inequalities in access to health care or the differential experience of a nuisance such as noise, air pollution and so on. Descriptive research of this kind is justified because it provides information on aspects of life hitherto neglected in geography. It also provides a basis for evaluation, whereby the existing state is judged against an alternative (the past, predicted or planned) according to some criterion of welfare improvement. Thus, the impact of alternative plans for facility location or closure (e.g., hospitals, schools) could be judged by the test of which would most equally distribute the benefits (such as access to health care) among the population of various sub-divisions of the area under review. This raises the question of rules of distributive justice and the manner in which they are applied (explicitly or otherwise) in the political process. Although originally proposed as an alternative framework for human geography, the welfare approach has now been merged with other lines of inquiry within geography directed towards the fundamental problem of inequality.

Implicit in 'welfare geography' is recognition that the issues in question extend beyond the limits of a single discipline, and in fact, render disciplinary boundaries increasingly irrelevant. The welfare approach logically requires a holistic social science perspective. In order to achieve the welfare target, geographers are attacking social problems and exploring the causes of socio-economic backwardness, environmental pollution, and uneven levels of development in a given physical setting. Now, the main objective of geographical teaching and research is to train students in the analysis of phenomena, so that they can take up subsequently the problems of society as the fields of their research and investigation, thereby helping the local, state and national administration. Problems are being tackled with approaches ranging from positive to normative, from radicalism to humanism, and from idealism to realism.

In brief, geographers are increasingly concerning themselves with the problems of society, conditions of mankind, economic inequalities, social justice, and environmental pollution. For reduction of regional inequalities and for the improvement of the quality of life, the main concern of geographers is with what should be the spatial distribution of phenomena instead of with what it is. It is in this context that the spatial inequality in social amenities and living standards is investigated by geographers to trace the origin of disparity rather than to condemn injustice.

The geographers who are mainly concerned with the problems of society and trying to formulate pragmatic proposals for public policy clarify the description and explanation of the phenomena. On the basis of such analysis, they evaluate their plans and prescribe suitable strategies for balanced development. Description involves the empirical identification of territorial levels of human well-being i.e. the human condition. This is a major and immediate research area in which surprisingly little works has been done in India and in other developing countries.

UNIT - 10:FEMINISM AND FEMINIST GEOGRAPHY

INTRODUCTION

It is very important to find answers to certain queries before going into a detailed discussion about feminist geography as, the key concept of the discipline may be rooted in it. Several statistics across the globe pose certain questions before us as to why there are lesser number of females in certain parts of the globe as compared to males; why the prevalence of illiteracy is more among females than males; why females in younger age groups tend to be more unemployed than their male counterparts; or why females are most often under-represented in governments and politics. In short, whether in terms of birth, education, economy or politics, opportunities and power are unequal between the sexes. It is this *'inequality'* that forms the subject matter of what is known as *'feminism.'* The most important feature of feminism is that it challenges the traditional thinking by connecting issues of production with the issues of reproduction; and the personal with the political.

The feminist theory is essentially based on *three* assumptions:

- ✓ Gender is a social construct that oppresses women more than men.
- ✓ These constructs are shaped by patriarchy.
- ✓ Women's knowledge about these constructs helps in envisioning a future non-sexist egalitarian society.

Thus, two relevant concepts that need to be understood here are that of *'gender'* and,*'patriarchy.'*

The word gender is often used interchangeably with sex, though the two have different connotations altogether. While sex is biological, natural and remains constant over space and time; gender is a social construct that may vary with time, space and culture. Gender is a social classification of the sexes into masculine and feminine. Different masculine and feminine qualities may have their impact on the social and spatial relations between and among the sexes. When such relations are approached by geographers from within the realm of the principles and concepts of feminism, what arises may be termed as *feminist geography*. Since feminism always deals with women's position vis-à-vis men; there may be another simultaneous field of study within geography, that is, the *geography of masculinities*. Together they constitute what can be precisely called *'gender geography.'*

The term *patriarch* originally derived from the Old Testament means the rule of the father (*pater* in Latin meaning father). However, the feminist use of the term was introduced by Kate Miller in her groundbreaking book, *'Sexual Politics'* in 1970. The term may be well understood in the words of Marilyn French as the manifestation and institutionalization of male dominance over women and children in the family and the extension of this dominance in the society as a whole.

The following aspects of women's lives may be under patriarchal subjugation:

- ✓ Women's productivity and labour power.
- ✓ Women's reproductive capacity and sexuality.
- ✓ Women's mobility.
- ✓ Women's access to economic resources.
- ✓ The social, cultural and political institutions.

To develop a proper understanding of the subject matter of feminist geography, it is necessary first to understand the true meaning of the feminist theory, its development through time, the different schools of thought that emanated within it and how its methods can be used in geography.

THE CONCEPT OF FEMINISM

Feminism as a concept is often misunderstood as an approach with extreme hatred for men and that a feminist is essentially a female. But in reality, there is no biological pre-requisite to be a feminist---males can also be feminists and in fact some are, just the way some women are not. The feminist theory upholds that inequality exists between the sexes. It has *four* notable features:

- ✓ It is intensely interdisciplinary in nature ranging across various disciplines.
- ✓ Certain themes are recurrent in it----reproduction, representation, sexual division of labour.
- ✓ It imbibes in it new concepts like sexism which are not only created to address the gaps in existing knowledge but also to describe forms of social discrimination.
- ✓ It draws upon women's subjective experience to enrich knowledge.

The idea of *'women'* as a distinct social group dates back to the 18th century. The first full political argument for women's rights and individual development was inspired by the French Revolution. At that time, Mary Wollstonecraft described in her *'A Vindication of the Rights of Women'* (1792), the psychological and economic damage experienced by women owing to their forced dependence on men and exclusion from the public sphere. Over time, the ideology of feminism has passed through several *waves or phases* that resulted in the development of its different variants.

The *first wave* of feminism started with the liberal principles of individual rights and freedom for women. The *liberal feminists* contrasted the concept of servitude of women that was considered as *'natural'* and protested against all forms of subordination that reduced women to adjuncts of their husbands or fathers. The roots of this stream of feminism can be traced in 17th century British liberalism and the French Revolution. Wollstonecraft, a liberal feminist advocated for the protection of women under civil laws, their right to be politically represented and to be engaged in well-paid work and respected professions so as to reduce their dependence on the institution of marriage. Harriet Taylor argued that women should be allowed to work even after their marriage because, not only will her economic contribution to the family promote her status within it, but it would also enhance her freedom of choice. Domestic violence and the tyrannical behaviour by the husbands was a central theme of focus for John Stuart Mill.

By the 1960s, though the first wave of liberal feminism had achieved its basic goals in Europe, women still suffered from various forms of legal discrimination and were grossly unequal in both economic and political terms. The *second wave* of feminism thus, that started in Europe towards the end of the 1960s, sought to adopt a socialist and radical standpoint. Since 1970s, many feminists had started questioning the relevance of liberalism as a possible remedy to women's subjugation. Hence, *Marxist feminism* emerged as a dominant strand of feminist ideology in the 1970s and 1980s. This variant of feminism, as the name suggests, drew its ideas from the theories of Karl Marx. It attempted to link the situation of women's oppression to class struggle and economic development. Though Marx himself did not have much to say regarding the situation of women, his methods and concepts were universally accepted and applied. This method argued that the key to comprehend the women's question is laid in the development of production, that is, economy and technology.

Therefore, like any other social organization, the relationship between the sexes is a function of a particular stage of economic development and cannot be altered on its own but only through socio-economic changes resulting from class conflict and revolution. Engels believed that women's oppression did not exist through time but only started with the creation of private property and a class-based society. Hence, only with the overthrow of capitalism, such oppression would disappear as, women would be no longer economically dependent on men and socialization of housework and childcare would free them from domestic chores. Therefore, women instead of fighting for their own causes should stand by working men for a revolutionary transformation of the society. This strand of feminism ruled out the idea that the interests of working men and women might conflict and that, women can have group interests beyond class lines or gender relations.

By this time, another group of feminists were developing their theories asserting that patriarchy, and not class was the oldest form of oppression. They constituted the *radical feminists* who originally worked within the Marxist set up in which they found that women's issues were treated as trivial. They were of the view that Marxism and feminism were not compatible with each other. However, in response to this there were some Marxist feminists who rejected the concept of patriarchy as historical and opined that women's issues could not be isolated from a wider socialist movement. They tried to analyse women's work both in home and in paid employment which eventually gave rise to the domestic labour debate and there was a demand of *'wages for* *housework.'* Some of the key ideas associated with radical feminism may be listed as (i) unity of theory and practice; (ii) linking the personal with the political; and (iii) the fundamental nature of women's oppression and subordination.

By the 1990s, there was a deep distrust for any metanarratives or any universal philosophy as Marx's. This was the beginning of the post-modern era. Jean Francois Lyotard's *The Post Modern Condition* (1984), laid the foundation for post-modern feminism which believed that, women like race, class or ethnicity could not be used cross-culturally to describe the practices of human societies and that it was not a universal category. Lyotard criticized the Marxist philosophy for propounding a homogenous society which was believed to be created only through coercion. Post-modern feminism upheld that social identities were heterogeneous and complex, and it was thus impossible to create a totalizing social theory.

EVOLUTION OF FEMINIST GEOGRAPHY

By the 1970s it was increasingly felt that very little attention was being paid to the matter that whether the methods of mainstream research and theoretical approaches could be applied in feminist studies. Prior to this, it was a widely held notion that women were not capable of political thinking or economic decision-making and, even in academia the discipline of geography was no exception to this. It was realized that since there were very less women academicians in geography, women's issues were not sufficiently studied in it. The preliminary objective was therefore, to make women visible in the field of geographical studies. What followed was a series of articles that attempted to probe the position of and acknowledge the presence of women within geography. One of the pioneering works was *'The Strange Case of the Missing Female Geographer'* (1973) by Wilbur Zelinsky.

Drawing inspiration from the development of feminist theory in the social sciences and the welfare, radical and Marxist streams of geography, soon there were works produced by members of several women study groups and professional geographical associations in United States, Canada and Britain. Mention in this regard, may be made of *The Women and the Geography Study Group* of the Institute of British Geographers (IBG) who presented a series of researches on feminism and geography at the annual meeting of the IBG in 1981. In 1983, they also organized a series of sessions on feminism as a mode of geographical thought and thereafter in 1984, published their landmark work, *'Geography and Gender: An Introduction to Feminist Geography.'* In 1982, Janice Monk and Susan Hanson collaborated to produce an outstanding article *'On Not Excluding Half of the Human in Human Geography.'* Mazey and Lee's *'Her Space, Her Place* (1983) provided one of the best introduction to this emerging branch of geography.

Taking recourse to conventional geographical methods, they tried to map the geography of women's rights; status of abortion laws; economic and political participation of women; their differential access to education, income and health services; their daily travel patters as well as long-term migration patterns. In 1984, two important works of feminist geography came forth in United States---a Ph.D. thesis was written in the department of geography at the University of California in Berkeley that was devoted entirely on feminist geography; and, a special edition of the geography journal *Antipode* (mouthpiece of the radical geographers) was published dealing exclusively with feminist geography. Following their British and American counterparts, a new specialized study group named the *Canadian Women and Geography (CWAG)* was created within the Canadian Association of Geographers in 1985. All these, greatly inspired the initiation of a multitude of research on women's topics by feminist academics in geography like—urban environment, housing, transportation, women in labour force, access to social services, violence, family structure etc.

By the 1980s, more advanced and theoretically sophisticated works began to be produced in this field. The celebrated article---*'A Woman's Place?'*—by Doreen Massey and Linda McDowell may be cited as an example. McDowell also published another work titled *'Coming in From the Dark: Feminist Research in Geography'* which itself is explanatory about the position of feminist studies in the recent past. Gradually geographical studies were being discussed more and more in feminist contexts. By this time, feminist geography was quite well- established and some feminist geographers wanted to extend the arena of this discipline beyond the Anglo-American circuit to the developing world as, in Africa, Asia and Latin America.

As the 1990s approached, feminism in geography was strongly grounded. This fact can be substantiated by the launching of a new and exclusive journal on feminism *Gender, Place and Culture: A Journal of Feminist Geography* in 1994. This journal was totally devoted to issues of feminism, gender, sexuality and so on within geography.

Three interrelated observations stimulated the growth of feminist geography. Firstly was the presumption that the spatial layout is essentially gendered. To elaborate, *'private', 'home', 'suburbs'* are always associated with women in the public-private, work-home or city-suburbs relations. Secondly, it was observed that culturally specific notions about gender behaviour are greatly shaped by spatial relations. Women's access to social services is largely determined by her location and associated gender roles. Thirdly, it was found that a person's relationship to the environment is largely a function of gender. For example, the idea of safe and unsafe environment may be different for women and men.

ECOFEMINISM—A MANIFESTATION OF FEMINIST THEORY IN GEOGRAPHY

Man-environment relationship has always been one of the prime themes of geography. Ecofeminism may be viewed as a feminist perspective to the relationship between nature and humans or in short, environmentalism. During the time when feminist issues began to appear in the discipline of geography, '*green politics'* in the West also assumed the character of mainstream politics with heightened concern for an ecologically balanced earth. Both the movements---environmentalism and feminism found a common ground of subordination by *'man'* (humans in case of environmentalism and man in case of feminism)---and joined hands to give birth to a new socio-political philosophy called *'ecofeminism.'* The original expression of the term was '*ecological feminiane.'* It was coined by the French feminist Francois d' Eaubonne to express a strong parallel between the subjugation of women in family and the society as a whole and, the degradation of nature. The term ecofeminism appeared for the first time in her book titled *'Feminism or Death'* in 1974. However, the term was popularized following the first ecofeminist conference that was held at Amherst in 1980 when large number of women across USA came together to launch their protests against environmental destruction. The basic essence of this concept is that, the devalued status of women in society and the degradation of nature are the two sides of the same coin. The nature was epitomized as feminine and male ownership of land and other natural resources were considered to give rise to a dominator culture. Hence, they used such terms as *'rape the land', 'tame nature'* and like. As described by Warren (1987), the *basic tenets* of ecofeminism may be elucidated as:

- ✓ Women are akin to nature whereas men are closer to culture.
- ✓ Both women and nature are conceived as *'producers of life'* that is ideologically rooted in their reproductive powers.
- ✓ A strong parallel exists between the oppression and domination of women and the degradation and exploitation of nature.
- ✓ Understanding this connection between women and nature is the basic requirement to understand the oppression of women and exploitation of nature.
- ✓ Hence, feminist theory and practice should have an ecological association.
- ✓ Likewise, environmental issues in turn, should be approached with a feminist perspective.
- ✓ Because, of the close link between women and nature, women can be perceived as important stakeholders in environmental protection and conservation.
- ✓ Finally, there should be the establishment of an egalitarian society in which there is no dominance on women or nature by *'man.'*

In the 1990s, by the time feminist geography was well-established, two prominent works on ecofeminism was produced. The one by Irene Diamond and Gloria Orenstein in 1990 named *'Reweaving the World: The Emergence of Ecofeminism'* laid out *three* strands of this concept:

- ✓ Social justice has to be achieved in collaboration with the well-being of the Earth as human life is dependent on the planet.
- ✓ The spiritual aspect emphasized on the sacrosanct earth.
- ✓ The third strand highlighted on the necessity of sustainability.

In the co-authored book *'Ecofeminism'* by Vandana Shiva and Maria Mies in 1993, they spoke of *three* kinds of domination prevalent in the world:

- ✓ Nature by humans.
- ✓ Women by men.
- ✓ Global South (the developing nations) by the Global North (the developed block)especially in terms of access to natural resources and controlling the world economy. Shiva asserted that one of main motto of ecofeminism was to modify the outlook of the society regarding the activities and productivities of women and nature, both of whom are misconceived as '*passive'* resulting in their exploitation. Mies described women's work as producing sustenance and advocated that women and nature worked as partners to give rise to a new kind of relationship in which there is an essence of reciprocity. Although women usurp nature, but there is no sense of domination. Rather, there is a sense of *'to let grow and to make grow.'*

Sometimes ecofeminism is also linked with *deep ecology* in their contrast to male chauvinism and both being forms of radical environmentalism. The term *'deep ecology'* was introduced around the same time as ecofeminism by the Norwegian philosopher Arne Naess in an article titled *'The Shallow and the Deep, Long-Range Ecology Movement.'* What was common between the two were---both were critical of any kind of hierarchy, be it nature- humans, man-women, nature-culture and so forth; and, both sought to establish an egalitarian system free from any form of domination with equal right to live and flourish. But there was a basic difference between the two. While deep ecology was against *anthropocentrism,* ecofeminism was against *androcentrism.* Deep ecologists considered human population as the root cause of biospherical destruction with humans selfishly multiplying at the cost of other forms of life causing their numbers to cross the carrying capacity of the Earth. However, they failed to provide an answer to the question as to why humans reproduce even in areas with huge population size that may often lead to food shortages, overcrowding, poor health and hygiene, degradation of land, destruction of species etc. For this, the ecofeminists found their answers in a multitude of human social factors many of which were akin to issues of gender and oppression. They may be highlighted as:

- ✓ ***Sexism:*** This means the *sexual disempowerment* of women. Glorification of male virility which is an expression of sexism associates reproductive capacities and abundance of offsprings with male prestige.

- ✓ ***Motherhood:*** Social stigma is attached to women not bearing children. Motherhood is an integral part of female identity which is often considered as the most meaningful purpose of their lives. ***Cultural Factors:*** Certain forms of birth control may be forbidden and may be treated as signs of collusion by some culture. Human reproductive capacities and having children are highly valued as per some cultural norms and beliefs which again are, formulated by men.

- ✓ ***Racism and Class Oppression:*** The reproductive issues become even more complex for women of colour, poor women or those from the developing block. Imperialism has left behind a disastrous level of poverty in which the need for children is intensified by economic stress and more children are regarded as constituting a large familial workforce.

Like feminism in which lies its root, ecofeminism too is not a single line of thought. Carolyn Merchant described about the variants of this ecosophy (ecological philosophy) in her book, *'Radical Ecology: The Search for a Liveable World'* (2005).

- ***Liberal Ecofeminism:*** This school of thought was in belief that if women were given equal educational opportunities as men, they could prove themselves as important stakeholders in natural conservation and improvement of the environment to ensure a higher quality of human life. They have to transcend the social stigma and biological constraints to join hands of men as scientists, lawyers, regulators and legislators for environmental conservation. This school of thought was thus in tune with reform environmentalism that sought to alter human-nature relationship within the existing framework of governance through the passage of new laws and regulations.

- ***Cultural Ecofeminism:*** Cultural ecofeminism spoke of an era prior to the emergence of the patriarchal system when femininity was held in high esteem and nature was portrayed through female deities. But, with the advent of the Industrial Revolution, the concept of the nurturing Earth was replaced with the metaphor of a machine to be controlled from outside by humans.

- ***Social Ecofeminism:*** This strand of ecofeminism advocated women's liberation through the overturning of the existing socio-economic hierarchies that translates all aspects of life to a market system. This ideology is based on the concept of *'social ecology'* that views the domination of nature by humans as an outcome of the domination of humans by humans.

It envisioned of a society constituted of decentralized units that would transcend the public-private dichotomy characteristic of the capitalist system.

- ***Socialist Ecofeminism:*** This variant of ecofeminism blurs into the previous strand and, provides a critique of the dialectical relations between production and reproduction, and between production and ecology that is typical of capitalist patriarchy.

Ecofeminism has been criticized for being too idealistic in its standpoint. It overemphasized on the mystical connection between nature and women rather than highlighting the actual conditions of women. The entire onus of environmental conservation rested with women and completely undermined the role of men in maintaining environmental sustainability. Moreover, it also failed to differentiate women across space with different social background. Nevertheless, it provided a sort of platform for the achievement of *sustainable development.*

Feminist geography essentially employs the feminist philosophy in addressing several issues of social and spatial relations and the role of men and women in such social and spatial frameworks. There is a political element inherent in it that adds some sort of *'radicalism'* to this field of geography. The basic premise of *'feminist geography'* is that advocates that inequality exists between the sexes over space which should be done away with to establish an egalitarian society.

UNIT - 11: POSTMODERNISM AND POSTMODERN GEOGRAPHY

INTRODUCTION

It is currently fashionable to talk of modernism and post-modernism in social discourse though precise meanings of the two terms are not always clear; so that one commentator described them as "the buzzwords of the 1980s". Important issues are, however, at stake so that the debate on postmodernism cannot be dismissed simply as "the product of disciplinary politicking, in which ascendant academics are once again jostling for new stalls in the intellectual marketplace", or as "little more than the cultural logic of late capitalism, the designer label on the baggage of flexible accumulation". But,

if these twin objections are superimposed, a different reading becomes possible and post-modernism can be seen to mark an unprecedented crisis of intellectual activity within the contemporary crisis of modernity. The debate then has the liveliest of implications for the project of a critical human geography (and critical theory more generally) because it raises acute questions about the very possibility of critique itself (Gregory, 1989, p. 348).

The two terms are closely interrelated so that it is not possible to understand post-modernism without due reference to the history and nature of modernity. Following the *Dictionary of Human Geography* (1994), *modernity* may be described as "A particular constellation of power, knowledge and social practices which first emerged in Europe in the sixteenth and seventeenth centuries, and the forms and structures of which changed over time and extended themselves over space until, by the middle of the seventeenth century, they constituted the dominant social order on the planet". Toward the end of the eighteenth century "the idea of being modem came to be associated not only with novelty but, more particularly, with looking forward rather than backward: with the so-called 'Enlightenment project' of reason, rationality and progress towards truth, beauty and just life".

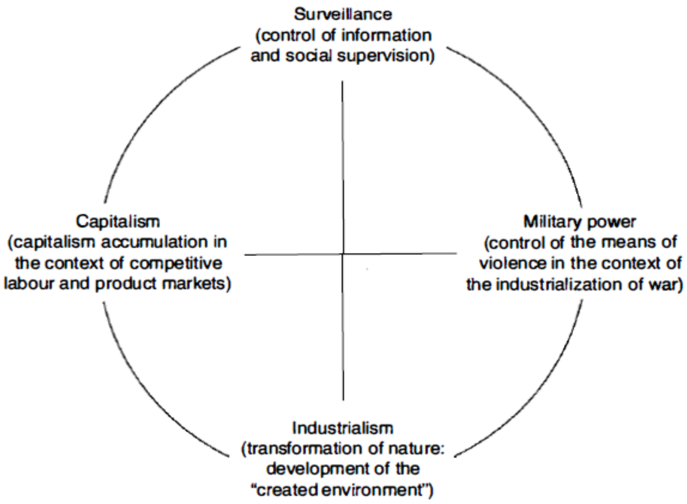
In the middle of the nineteenth century, Marx and Engels found an essential connection between modernity and the revolutionary dynamics of capitalism under the impact of which "All fixed, fast-frozen relations, with their train of ancient and venerable prejudices, are swept away, all new-formed ones become antiquated before they can ossify. All that is solid melts into air", wrote Marx and Engels. The beginnings of the modern world system are traced back to the sixteenth century, beginning in Europe, and organized around market exchange at international scale. But the process was greatly intensified through the nineteenth and the twentieth centuries through global scale of economic operations, so that modernity became identified as the system of market-based political-economic structure and organizations. Thus, by the second half of the twentieth century, a profound crisis of identity and representation had begun to be felt all around-in the arts, humanities, and the social sciences-particularly in the Western metropolises and the colonial capitals, owing to the accelerated pace of time-space compression. By the end of the 1950s, the modern way of viewing phenomena in ter11,s of Cartesian coordinates had begun to be questioned, separation between place and space began to be drawn, and the idea of bounded cultures through which one could journey out and analyse other cultures appeared no longer relevant (Clifford, 1988, p. 2). All this called for new approaches to explanation of social reality-approaches focused on space specificity, that is, approaches beyond the modernist tradition of universalism. This was the beginning of post-modernism in social thought. There are contrasting perspectives on the plus and minus sides of modernity, as defined above. One set of commentators holds that:

... modernity was an unqualified human good which promised to banish ignorance, misery and despotism; to free human beings from myth and superstition, disease and hunger, oppression and arbitrary rule. In the middle years of the twentieth century these assumptions culminated in models of modernization, and development programmes of which sought to remake the so-called traditional world in the image of the West (Gregory, 1994a).

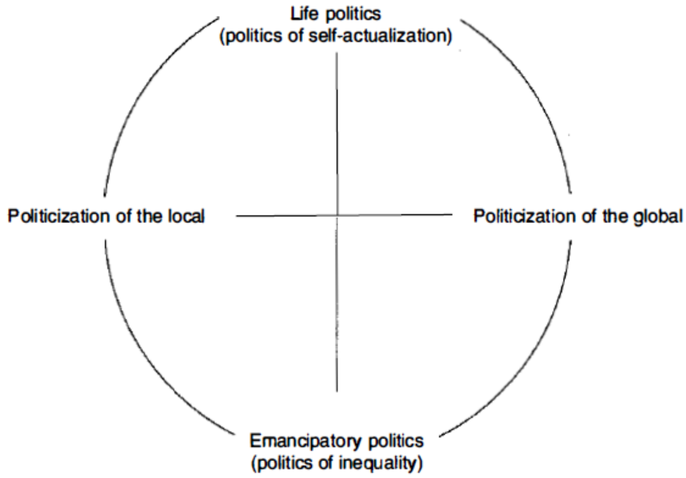
But as the critics have noted, these blessings were a mixed lot. Modernism was closely intertwined with the processes and objectives of colonialism and repressive imperialism, besides ethnocentrism rooted in European cultural values and perspectives:

To other commentators, therefore, modernity has always had its dark side, many of them argued (in different ways) that European modernity installed new grids of power and surveillance which, in Max Weber's well-known image, confined human agency, consciousness and creativity within an "iron cage" of· bureaucracy and regulation (Gregory, *op.* cit.).

Giddens (1990) has drawn attention to four major institutional dimensions of modernity: surveillance, industrialism, capitalism, and military power. He has also proposed an equivalent nexus of political strategies wherein the intertwining of the local and the global scales and of life politics and emancipatory politics are underlined. Figure 14.1 illustrates the proposed relationships.



(a) The institutionalized politics of modernity (Giddens, 1990)



(b) The politics of modernity

Figure 14.1: Politics and modernity (After Giddens, 1990)

MODERNITY, MODERNIZATION, AND MODERNISM: THE SOCIAL SCIENCE CONTEXT

In a social, scientific context Berman (1982) refers to *modernity* as "a mode of vital experience", a collective sharing of a particularized sense of "self and others", and of "life's possibilities and perils". "Modernity is thus comprised of both context and conjecture, the *specificity* of being alive in the world at a particular time and place, a vital sense of what is *contemporary.* As such, it becomes a useful general term to capture the specific and changing meaning of the three most basic and formative dimensions of human existence: space, time, and being; the spatial, temporal, and social orders of human life" (Soja, 1989a, pp. 320-321). The concept of modernity is intimately related to the multiple reconfigurations of social life that have punctuated the historical geography of capitalism since the sixteenth century. In this context, *modernization* refers to "the many different processes of structural change associated with the ability of capitalism to develop and survive, to reproduce successfully its fundamental social relations of production and distinctive divisions of labour". As Soja *(op. cit.)* points out, "these restructuring processes are continuous, but they become especially critical and accelerated during periods of deep and systematic global crisis" like the one we have entered since around late 1960s. Such periodically intensified modernizing episodes of capitalism are, in the words of Berman (1982, p. 16), shaped by:

... the industrialization of production, which transforms scientific knowledge into technology, creates new environments and destroys old ones, speeds up the whole tempo of life, generates new forms of corporate power and class struggle; immense demographic upheaval, severing millions of people from their ancestral habitats, hurtling them halfway across the world into new lives; rapid and often cataclysmic urban growth; systems of mass communications, dynamic in their development, enveloping and binding together the most diverse people and societies; increasingly powerful national states, bureaucratically structured and operated, constantly striving to expand their powers; mass movements of people and peoples challenging their political and economic rulers, striving to gain control over their lives; finally, bearing and driving all these people and institutions along, an ever expanding, drastically fluctuating capitalist world market.

In broad terms, *modernism* "refers to the cultural, ideological, reflective and ... theory-forming reactions to modernization and restructuring .... Modernism is thus the explicitly evaluative, culture-shaping and situated consciousness of modernity and is itself roughly able to be split into periods in conjunction with the historical rhythms of intensified capitalist crisis, restructuring and modernization" (Soja, 1989a, p. 322, *also see* Soja, 1989b).

THE RISE OF MODERN HUMAN GEOGRAPHY

The discipline of geography as practised over the past two centuries is essentially an "European science" which, according to Stoddart (1986), dates back to 1796, the year in which Captain James Cook first entered the Pacific. Cook's was pre-eminently a scientific voyage carrying on board illustrators, collectors and scientists whose work displayed, according to Stoddart, three features of decisive significance for the transformation of geography into an "objective science" focused on concern for realism in description; systematic classification in collection of data; and the comparative method of explanation. Stoddart maintains that it was the extension of the scientific methods of observation, classification and comparison to the domain of peoples and societies that made geography, as practised for the past two hundred years, possible. (It may be recalled here that it was George Forstera member of the scientific team that had accompanied Captain Cook-to whom Humboldt owed his interest in the study of geography.) Following Foucault (1979, 1980), Gregory (1989) has underlined that this was far from a simple "extension" of scientific principles because these scientific principles had the most radical consequences for the constitution of the human sciences and their conception of the human subject. Besides the logic of this extension to peoples and societies "was powered by more than the ship of reason since these nominally scientific advances were at the same time the spearheads of colonialism and conquest: of subjection", so that "its trajectory cannot be separated from what Wolf (1982) once called 'the people without history'. It follows that some of the most seminal cross-fertilizations during this period were those between anthropology and geography, loosely defined, and what later became anthropogeography has to be acknowledged as a tradition of basic importance to the modern discipline".

This thread of relationship is indelibly present in the writings of Humboldt, Ritter, and of course, Ratzel and Sauer. Towards the end of the nineteenth century, a second thread of relationship between sociology and geography had emerged: "the interactions between the two disciplines", Gregory comments, "were, in part, a way of clarifying their different destinations". While sociology, according to Durkheim, was to subsume the study of the spatial structure of *society-morphology sociale-the* Vidalian school insisted on the independent course of human geography predicated, on the one hand, on horizontal (spatial) relations and, on the other, on vertical relations between society and nature (Berdoulay, 1978). Around the same time, Richard Park at Chicago was airing similar views regarding sociology subsuming the study of *morphology sociale,* and he described geography as a *concrete science* wherein interest in spatial relations was restricted to the study of the individual and the irregular (unique), that is, to the idiographic, and not concerned with the elucidation of the general processes and universal principles, as in the case of history.

Against the "concrete science" nature of geography, sociology and human ecology were described as *abstract sciences* which, even though their foundations were laid on history and geography, were essentially theory-seeking in their orientation. Thus the encounter between the two disciplines had ended in geography and sociology ultimately developing in isolation from each other (Further details on pp. 177-178). By mid-twentieth century, geography had developed close links with economics alongside the existing frontiers with anthropology and sociology. Until the late 1940s, like the rest of human geography, the human geography economics frontier was pursued in a resolutely empirical manner; but the emergence of the "new" geography in the 1950s and the 1960s gave it a pronouncedly theoretical orientation so that human geography now assumed the characteristics of the abstract sciences, in the sense that like sociology and economics, it also became a nomothetic science but (in general) it continued to eschew the study of the social element.

Thus, few of the contributions in Chorley and Haggett's *Models in Geography* (1967) were addressed to social and cultural questions in any manner, and "the geometric order which they supposed to be immanent in the human landscape was derived, for the most part, from the economy: from trans codings of an abstract calculus of profit and loss into an equally abstract 'friction of distance'" (Gregory, 1989), so that human geography was transformed into spatial science. This short resume of the historiography of geography, according to Gregory, underlines three distinctive features of modern geography. These are:

1. All the three strands were bound into a naturalism-quite naturally so for a discipline that combines the study of both the human and the physical, but the other branches of the social sciences did not escape the impact of evolutionary biological thought either. "It is scarcely surprising, in consequence, that identical links should have been forged in human geography; from Vidal de la Blache's seemingly simple organicism to Hagerstrand's vision of time-geography as a 'situational ecology'; from essentially Newtonian models of spatial interaction to technically more sophisticated models of dissipative structures within abstract landscapes".

2. All the three strands were woven into a conception of science as *totalization.* Modem science holds that explanation resides in the disclosure of a systematic order whose logic imposes coherence on the apparent chaos observed by the investigator. In the study of man, this scientific tenet involved identifying a *centre* around which social life revolves. The concepts of "social system" or social structure, on the one hand, and the alternative perspective around human subjectivity on the other, served this purpose. As Gregory wrote, both these can be read in human geography, but the formalization of geography as spatial science in the 1960s accorded central place to geometry, i.e., to the search for abstract order in mathematical spaces: "Spatial science was part of an economy of explanation which put a high premium on the *simplicity* of its explanation", so that whatever be the subject of investigation, the same explanatory mould was applied.

3. If the three strands in geography since the beginning of the nineteenth century to mid-twentieth century are joined together, then human geography as spatial science would appear as a natural growth, rather than as representing any break from the past (as it was projected at the time). Geography as spatial science was "a continuation and *culmination* of modem geography", in the sense that it helped complete the search for scientific identity of the discipline as a nomothetic science. (Figure 14.2 depicts the sequential construction of human geography as a *modern* discipline.

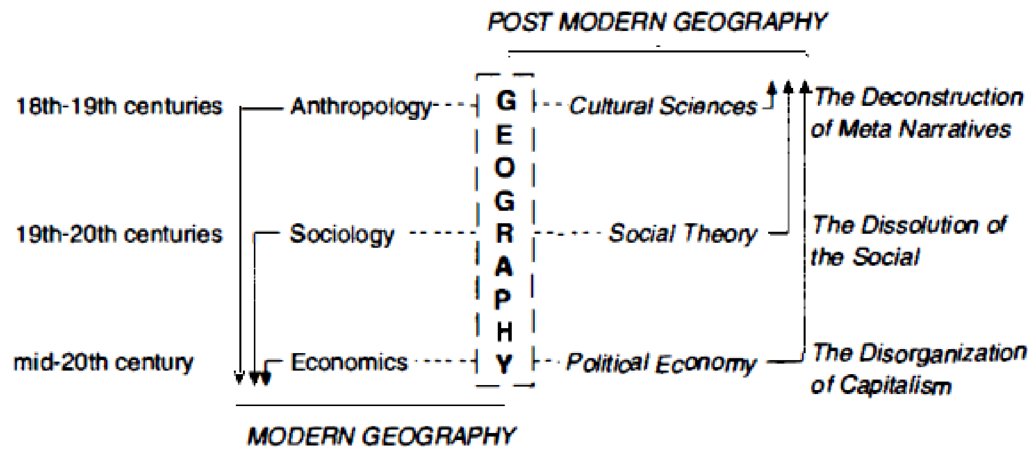


Fig. 14.2 The making and unmaking of modern human geography (*Source:* Gregory, 1989, p. 350).

THE CHANGEOVER TO POST-MODERN GEOGRAPHY

Post-modernism is a recent movement in philosophy, the arts and social sciences. Its distinguishing characteristics (according to Ley, 1994) are: Scepticism towards the grand claims and grand theory of the modem era, and their claim to intellectual superiority. Post-modernism (as contrasted to modernism) stresses openness to a range of opinions in social enquiry, artistic experimentation and political empowerr11ent. Dear (1986) classified modernism into three components: style, method, and epoch. *Style* is intimately implicated in the constitution of meaning and identity of the phenomenon being represented. Styles may, according to Hebd.ige (1979), either be supportive of dominant ideologies, or they may offer a ritual resistance to movements which are socially and politically opposed to the dominant view. Ley' and Mills (1992) note that the same is true about the socially constructed environment whose variations over the earth surface are of special concern to human geography. The post-modern *methodology* is based essentially on the strategy of deconstruction. Deconstruction is defined as a mode of critical interpretation which seeks to demonstrate how the multiple positioning of an observer (an author or a reader) in terms of culture, class and gender etc. has influenced his observation (his writing or reading) of the *text* (the thing being observed or read or written about). Thus, "Deconstruction is essentially a destabilizing method, throwing into doubt the authoritative claims of preceding traditions, and seeking to prise loose alternative readings of texts" (Ley, 1994).

Also, post-modernism may be viewed as an *epoch*, that is, a historic era in which changes in culture and philosophy are sought to be located in the very process of the evolution of economy and politics at the global level. Thus some commentators regard post-modernism as the culture of late capitalism Uameson, 1984), an integral part of the new phase of post-Fordist flexible accumulation (Harvey, 1989). This, also, according to Soja (1989), is part of the attack on the predominance of *historicism* in modem system of thought, which highlighted the importance of individual and collective biography of diverse communities in the explanation of social relations but neglected the role of spatiality, so that the role of geographical imagination was marginalized. This subordination of space to time obscured the role of geographical interpretation in the changeability of the social world. As a consequence, geography failed to attain its rightful position in the development of theory in the social sciences (see: pp. 176-179). Historical materialistic interpretation assumed that society everywhere (irrespective of its spatial context) was characterized by similar historical pulsations in social change.

Thus the social theories of historical materialism were deficient and partial in that they ignored the basic fact that the spatial context of history and culture makes a basic difference in how societies interpret particular episodes in history, and develop ways and means of adjusting to them. Consequently, the historical flow of social change was not the same everywhere. Post-modernism represented a response to modernism as a homogenizing force and to its totalizing theoreticism. Instead, postmodernism lays stress on discontinuities and disjunctures characteristic of everyday life in the real world. The homogeneity in the built-up landscape of the era of "organized capitalism" -or the Fordist era, as it is sometimes called-is contrasted with the heterogeneity of economic, social and political life in the current phase of "disorganized capitalism" (or of flexible accumulation of the post-Fordist era) (Harvey and Scott, 1989).

The post-modernist emphasis on heterogeneity, particularity and uniqueness of phenomena in differing contexts of time and space, would appear to hark back to the regionalist tradition in geography of the Vidalian era: Post-modern thought provides a theoretical context for the study of spatial/ areal diversity in the life-world. Viewed thus, post-modernism in geography is closely linked to the rise of a "new" regional geography referred to previously (pp. 256-259). The difference between the old style regional differentiation in geography of the Hettner-Hartshome era, and the present-day post-modernist regional geography lies in that, while the former was indifferent to everyday experience of societal relationships, there is now a declared commitment to the understanding of the condition of man in particular places, and the ways that spaces are socially constructed.

Post-modernism is also a pronouncedly anti-spatial science in terms of the spatial scientist's search for pattern laws of behaviour. Post-modernism does not aspire to generate any grand theory of universal application; its essentially heterogeneous ethos runs counter to the emphasis on consensus on theoretical perspectives (that characterized the discourse in the spatial science era). Instead of universal theory, the emphasis now is on what Dear (1988) called, the contemporaneity of social process over time and space, so that geography, like history, is preeminently a contextual discipline, engaged in the time-space reconstruction of social life.

The movement toward post-modernism in geography (i.e., unmaking of modern human geography, Fig. 14.2) started in the 1970s with the introduction of the political economy perspective in human geography perspective that reoriented human geography to the condition of man in society and gave rise to the "radical" or "relevance" revolution. In the words of Gregory, the cutting edge of the political economy perspective was the *logic of capital* scrawling its signs on the surface of the landscape-eas against the *logic of space* (i.e., geometry) that dominated spatial science geography until the 1960s. Thus, the focus of post-modem geography became a focus on uneven development in the spatial context. There was an inevitable confusion in conceptual design in the beginning owing to the fact that political economy derived its perspective from predominantly Marxist thought, and Marx's writing, howsoever liberally interpreted, contained little about space. The rise of the political economy perspective in geography almost coincided with the emergence of what Harvey (1987) called "disorganized capitalism". Commentators now began drawing attention to the fact that the Marxist view of capitalism was deficient in that production of space is focal to the development of capitalism in the twentieth century, and the new regime of "flexible accumulation" depends upon the production of "a tense and turbulent landscape" with disjointed space economies. Such developments in space economy are fragments of a still larger mosaic:

. .. of Castell's "new relationship" between space and society, formed by the hypermobility of capital and information cascading through a vast world system, where "space is dissolved into flows", "cities become shadows" and places are emptied of their local meanings; of Jameson's "post-modem hyperspace", part of the cultural logic of late capitalism, whose putative "abolition of distance" renders us all but incapable of comprehending--of rnapping-the decentred communication networks whose global webs enmesh our daily lives (Castells, 1983, p. 314; Jameson, 1984) (Gregory, 1989, p. 353).

Towards the end of the 1970s, working out the implications of historical materialism thus moved beyond Marx's original thesis to fashion a p o s t Marxist theory of socio-spatial structuration which emphasizes that *people not only make history but also geography-which* is to say that time-space relations are intrinsic to the constitution of societies (Gregory, 1978). Such a view cuts into the sociological concept of *societies* as totalities with clearcut boundaries, and renders them extremely unstable, since under the new concept, society becomes the product of the varying spans of time and space.

This ends the earlier view of dualism between society and space, between agency and structure. Such a (supposedly) Kantian view of division between the social and the spatial was made all the more distinct in the spatial science episode in human geography. This duality (born essentially out of the urge to apply the methodology of the natural sciences to the study of sociological phenomena) had long been attacked by scholars of hermeneutics (the study and interpretation of meaning) from the time of F. Schleiermacher (1768-1834) and Wilhelm Dilthey (1833-1911). Dilthey had argued that the human sciences, owing to the peculiarity of their subject matter, required a special methodology very different from the empirical methodology of the natural sciences. This special methodology was hermeneutics that tried to search for meaning in all kinds of activities and objects, including tools and landscapes as well as individual biographies. Hermeneutics went beyond mere physical appearances-physical and physiological processes, and surface patterns-to seek deeper meanings and causes by paying attention to the meaning and motivation behind actions. But hermeneutics was not a barrier to objective knowledge. Hermeneutics was introduced in human geography in the mid-1970s through the writings of Buttimer (1974) and Tuan (1974), and was in late 1970s visible under two different titles: humanistic geography, and critical theory.

The latter was proposed by Gregory (1978) with a view to linking the hermeneutical approach to a critique of traditional historical and regional geography. Thus by around 1980, the supposed dualism between space and society was finally broken, since with such a linguistic/hermeneutic, the problem of geographical description became much more than a mere matter of abstract logic: It became "a question of reading and decoding", of ***"making sense*** of different places and different people". Thus, as in the case of ethnography (Geertz, 1983), in geography also, the stress was now once again on local knowledge, which "requires us to attend not only to the theoretical categories which inform our accounts but also to the textual strategies through which they are conveyed".

The linguistic turn to the study of human geography through the introduction of hermeneutics now made geography open to experimentation with different modes of representation, so that the discipline reopened its doors to cultural studies through greater interaction across the frontier with other branches of cultural studies. Thus was completed the postmodernist journey in human geography, beginning in the 1970s, through political economy of the disorganization of capital (and thereby highlighting the relationship between history and geography, structure and agency), followed by the dissolution of the barrier between the social and the spatial, arrived at through revival of the frontier between geography and sociology by giving greater attention to social theory. This was followed by opening of the frontier with other disciplines focused on the study of cultures. The ***post-modern*** journey thus followed a route in an order that represented the reversal of the route charted by ***modern*** geography (Fig. 14.2), but the reverse journey was imposed by the turn of developments in the contemporary capitalistic system-a consequence of history.

It represents essentially a process of evolution so that post-modernism must not be seen as a negation of everything that earlier human geography represented, but as "a commentary upon it" (Gregory, 1989). The crisis of "modern" geography until the mid-1970s lay in that geography had cut itself loose from mainstream philosophy, humanities and the social sciences so that the postmodernist response lay in a realignment with these other streams of knowledge for intellectual leads, particularly with social theory, but asserting at the same time that there is no scope for grand theory in human geography. Indeed, post-modernism focuses on the creative tension between the different theoretical formulations, so that post-modernism is essentially polyphonic: It combines a number of individual but harmonious melodies, i.e., points of view or perspectives, with a view to arriving at more comprehensive explanation of social reality.

Gregory (1989) has identified three distinguishing features of postmodernist thought in geography: First, space-time specificity in social explanation. This implies that post-modernist thought in social science insists on the understanding of "a world which is meaningful for the people who live within it". Reflexiveness (implying focus on the role of the human agent's action upon himself or the identity between object and subject) is intrinsic to post-modernist work in the social sciences but this does not imply complete repudiation of the scientific principles of naturalism though the search for universal theories is eschewed *(see:* Gregory, 1991).

Secondly, post-modernism insists on distancing itself from the totalization (i.e., the concept of the society as a totality following a universal process of historical change irrespective of space context). Post-modernist thought holds that: "The ebb and flow of human history is not reduced to the marionette movements of a single structural principle, whatever its location, and the differences which make up human geographies are not explained by some central generating mechanism". Thirdly, post-modem human geography represents a *critique* of spatial science it is *not a continuation* of it-but at the same time, *it does not represent a break with the past.* Post-modem geography focuses on the essential *spatiality* of social life (rather than the geometry of social behaviour) (Dikshit 2018).

UNIT - 12: SUBALTERN STUDIES IN GEOGRAPHY

INTRODUCTION

The Subaltern Studies collective was forged by a small group of young historians based in Britain who, in the late 1970s, met regularly to discuss their increasing frustration with South Asian historiography. The group, who were collectively dissatisfied with historical interpretations of the Indian Freedom movement, included Shahid Amin, David Arnold, Partha Chatterjee, David Hardiman and Gyanendra Pandey, and they found intellectual leadership in the figure of Ranajit Guha. The problem with Indian historiographical scholarship as they saw it was that extant historical narratives of the making of the Indian nation celebrated elite political contributions while denying those contributions made by ordinary people. In 1982, the group formally established itself as the Subaltern Studies collective with the release of the first of what became an eponymous series of edited volumes. *Subaltern Studies I* was published by Oxford University Press, edited by Ranajit Guha, and contained contributions from all of the collective's founding members.

The collective's adoption of the term 'subaltern' of course referenced the early twentieth-century Marxist and founding member of the Italian communist party, Antonio Gramsci's writings on class and state in the *Prison Notebooks*. Gramsci's use of the term was a way of avoiding censorship while imprisoned by Mussolini (see Gidwani, 2009: 66), and it literally referred to lower ranking military personnel. Gramsci (1971: 55) used the term, however, to refer to social groups in fascist Italy who were 'always subject to the activity of ruling groups'. Subalternity named a proletariat whose voices were excluded from a ruling class, the Catholic Church, whose own interests forged hegemony in civil society and thus constituted domination in and of the state. But for Gramsci, subaltern organization fostered by the work of organic intellectuals should aim toward the reconstitution of hegemony by incorporating subaltern voices into a dynamic 'national-popular consciousness' (Hoare & Nowell Smith,1971: 3–4). In other words, and in terms that become relevant for the closing sections of this paper, subalternity and hegemony existed in dialetical relation to one another.

The Subaltern Studies collective seized upon Gramsci's political agenda to work through their own critiques of the dominant narrative of Indian nationalism, which at the time was written as an achievement of colonial India's political classes. This elitist historiography was to be expected given that the colonial archive was simply not able to retain the utterances and actions of illiterate peasant movements, of which there were few textual records. Nonetheless, these were the subaltern forms of agency that the collective wanted to recover in order to interpret, as Guha put it, 'the contribution made by the people *on their own*, that is, *independently of the elite* to the making and development of this nationalism' (2000:2, emphasis original). From the outset then, the Subaltern Studies project was driven by a desire to write histories from below. They reconsidered India's subaltern classes, not as ‘pre-political’, voiceless subjects, but as ‘[political] subjects in the making of their own history’ (Chaturvedi, 2000a: viii).

Guha's seminal 1983 monograph, *Elementary Aspects of Peasant Insurgency in Colonial India*, was a magisterial lesson on how practically to go about this kind of history writing from below in the absence of any archives of peasant voice. As he showed, insofar as a physical archive of subaltern expression cannot exist, the only available historical resources to glimpse peasant agency and insurgency were in fact those colonial archives that recorded the counter-insurgency measures of the ruling classes, their armies and the colonial police (Chakrabarty, 2002: 16). Guha argued, therefore, that writing history from below entails reading these archives *against their own grain*, being attentive to the ways that peasant agency is inferred, effaced and can be delineated from the voices of power that structure the textual properties of the colonial archive (Chakrabarty, 2002: 16). In other words, authentic peasant agency and subjectivity would be recovered only if the historian reads for lack, absence and effacement in the colonial archive.

HISTORICAL ORIGINS: INSURGENCY, NATIONALISM, AND SOCIAL THEORY

In the last forty years, scholars have produced countless studies of societies, histories, and cultures "from below" which have dispersed terms, methods, and bits of theory used in Subaltern Studies among countless academic sites. Reflecting this trend, the 1993 edition of The New Shorter Oxford English Dictionary included "history" for the first time as a context for defining “subaltern.” The word has a long past. In late-medieval English, it applied to vassals and peasants. By 1700, it denoted lower ranks in the military, suggesting peasant origins. By 1800, authors writing “from a subaltern perspective” published novels and histories about military campaigns in India and America; and G.R.Gleig (1796-1888), who wrote biographies of Robert Clive, Warren Hastings, and Thomas Munro, mastered this genre. The Great War provoked popular accounts of subaltern life in published memoirs and diaries; and soon after the Russian Revolution, Antonio Gramsci (1891- 1937) began to weave ideas about subaltern identity into theories of class struggle. Gramsci was not influential in the English-reading world, however, until Raymond Williams promoted his theory in 1977, well after translations of The Modern Prince (1957) and Prison Notebooks (1966) had appeared.14 By 1982, Gramsci’s ideas were in wide circulation. Ironically, though Gramsci himself was a Communist activist whose prison notes were smuggled to Moscow for publication and translation, scholars outside or opposed to Communist parties (and to Marxism) have most ardently embraced his English books (as well as those of the Frankfurt School).

Subaltern Studies deployed some of Gramsci's ideas at a critical juncture in historical studies. By the late 1970s, a rapid decline in state-centred historical research had already occurred and social history "from below" was flourishing. E.P.Thompson’s 1963 book, The making of the English working class, is often cited as an inspiration for the growing number of "bottom up" studies of people whose history had been previously ignored. By 1979, women's history was popular enough in the U.S. to merit source books and guides to research. In 1982, Eric Wolf published what can be called the first global history from below. In South Asia, the history of subaltern groups was thriving, though they were not called that then. In the seventies, two new journals featuring studies of South Asian peasants had begun publishing in the US and UK. Hundreds of titles on rural history had appeared. In 1976, Eric Stokes announced the "return of the peasant" to colonial history. Guides to sources promoted more local research.

Insurgency attracted special attention. In India, the 1857 centenary had stimulated new histories of rebellion, some directly inspired by rebels like Kattabomman Nayakkar, whose epic of resistance to British rule had been reproduced in many popular media, including cinema. Romantic heroism attached to old rebel histories, but in addition, the sixties and seventies raised concern about revolution in the present. Even the Indian Home Ministry feared revolution. In this context, more scholars took up studies of insurrection. N.G.Ranga and L.Natarajan pioneered this field, decades before, and elements of its intellectual history go back to the twenties, when early Indian studies of Indian rebels sought to recuperate insurgent mentalities. Indigenous Indian theories of peasant revolt had emerged in the thirties, among communists and in the Kisan Sabha, but in the sixties, the academic study of insurrection came into its own, when Hamza Alavi theorised peasant revolution, Stephen Fuchs explored tribal messianism, J.C.Jha studied Kol rebellions, and Muin-ud-din Ahmad Khan studied early Fara'Idi rebels in Bengal. In the seventies, the upward trend in research on popular insurgency accelerated: highlights include work by K.K.Sengupta, B.B.Chaudhuri, and S.K.Sen on rebels in Bengal; V.Raghavaiah's work on tribal revolts (published by the Andhra Rastra Adimajati Sevak Sangh); Ghanshyam Shah's early studies of Gujarat; a flurry of work on Mappillai revolts in Malabar; Kathleen Gough and Hari Sharma's path breaking Imperialism and Revolution in South Asia4 and A.R.Desai's masterful collection, Peasant Struggles in India.

Historians were dividing along schisms in social theory into opposing schools that separated society and culture from state institutions and political economy. Subaltern Studies dramatized this division. So did Benedict Anderson's book, first published in 1983, Imagined communities: reflections on the origin and spread of nationalism, which abandoned class analysis, ignored state politics, and argued that cultural forces produced national identity and passion. By 1983, scholars were writing two kinds of national history: one, a people's history filled with native culture and popular insurgency; the other, an official history filled with elites and political parties. Nations and states were separating like oil and water. So were culture and political economy. A new kind of nationality was coalescing in a separate domain of popular experience, which was becoming increasingly isolated from state institutions and national elites.

But even so, when Ranajit Guha announced, in 1982, that "the politics of the people … [form] … an autonomous domain," even those who agreed with him -- like Sumit Sarkar, who soon joined the project -- still assumed that diverging domains of nationality were connected. After all, this connection sustained the possibility of radical change, even revolution. In the seventies, this possibility had become a serious problem, because state institutions had remained substantially unchanged despite many decades of popular insurgency, nationalist agitation, and tumultuous independence not only in 1947 (India and Pakistan) and 1948 (Sri Lanka) but also in the 1971 Bangladesh Liberation War. Modern states did not prevent rebellion, but insurgency had not become revolution. Why did nationalism provoke revolution in China and Vietnam, but not India? How do oppressed people take over governments? How do nations redesign states? Why not revolution in South Asia? These were pressing questions.

SHIFTING GROUND: NATIONS, POLITICS, AND GLOBALISATION

Subaltern Studies joined debates about insurgency and nationality at the breach between popular unrest and state power. The breach was widening at the time, in part because, despite rampant crises, dominant state institutions had managed to survive as though secure inside a mountain fortress high above the plains. Looking back from 1980 into the decades before 1947, historians were busy exploring dis-connections between official nationalism and popular movements. Muslims had acquired a separate political history that became more prominent in the context of Hindu majoritarianism. Regional movements became prominent – and most thoroughly studied by Cambridge historians -- after the 1956 reorganisation of Indian states along linguistic lines. But communalism and regionalism did not attract Subaltern Studies, which instead focused on the separation of political strata. D.N.Dhanagare, Majid Siddiqi, and Gyanendra Pandey had already published books on splits between the Indian National Congress and peasant movements. Sumit Sarkar's Modern India gave workers' and peasants' movements more autonomous political space than any history text had ever done before. Ranajit Guha's Elementary Aspects of Peasant Insurgency depicted tribal revolts as completely separate from nationalism, inside a subaltern space, below. Subaltern Studies entered the academic scene by asserting the complete autonomy of lower class insurgency.

The breach between popular and national history then expanded to vast proportions in the eighties and nineties. This changing intellectual climate has yet to be adequately historicized and can only be outlined here in the sketchiest manner. One key feature stands out when we recall that histories "from below" had originally emerged inside an intellectual fusion of historical research and national politics. Books like A.R.Desai's Peasant Struggles in India (1979) and Agrarian Struggles in India After Independence (1986) not only promoted the study of agrarian upheavals in the past, they also opposed the technocratic developmentalism of the Green Revolution and the status quo politics of cultural traditionalism. In South Asia, this kind of scholarship goes back to the 1870s, when a nationalist academic critique of empire inspired national politics and history at the same time. It is easy to forget how radical the intellectual work of the early Indian nationalists was in its day. A third generation of nationalists, including Jawaharlal Nehru and Mohandas Gandhi, built upon a long legacy of critical scholarship. Nehru used history to inform his politics, the way Gandhi used philosophy, and in 1930, when he became President of the All-India Congress Committee, Nehru announced an enduring theme in historical research by saying, "the great poverty and misery of the Indian People are due not only to foreign exploitation in India but also to the economic structure of society, which the alien rulers support so that their exploitation may continue." Such pronouncements at the apex of nationalism stimulated many histories from below, which engaged the past to inform national debates about land reform, planning, local democracy, farm finance, industrialisation, and other topics of hot dispute. In this intellectual environment, history "below" embraced history "above." Gaps and failures separating levels and types of national activity seemed to be conjunctural problems to be overcome within a unified national history.

After 1980, an expanding gulf between the histories of peoples and states ripped many old bonds between academics and politics. Scholars who claimed to speak for people who had been left out of nationalism marched away from scholars who continued to fuse popular history with national politics. Ranajit Guha accounts for his own alienation from nationalism by citing the early seventies' "drama of Naxalite clashes with the organs of the state and the violence of counterinsurgency measures." But more importantly for many others, Indira Gandhi's Emergency in 1975 made the Indian state blatantly dictatorial. As new popular movements arose from many quarters in India -- communal, regional, and expressing radical aspirations among women, peasants, workers, and tribal groups -- old nationalism lost legitimacy and the Left and the Right fought for its legacy. Popular resistance to state power became a prominent academic theme in the eighties. In 1986, James C. Scott's Weapons of the Weak: Everyday Forms of Peasant Resistance announced a broad move away from studies of revolution into the analysis of localised, personal resistance to the power of elites and states. Foucault's influence was spreading. By the nineties, an array of scholars inside and outside Subaltern Studies had made everyday resistance a basic feature of life in South Asia.

INVENTING ORIGINALITY: REJECTION, CROSSROADS, AND NEW DEPARTURES

The original substance of Subaltern Studies emerged from work-in-progress in the late seventies. Eleven authors in the first three volumes -- Shahid Amin, David Arnold, Gautam Bhadra, Dipesh Chakrabarty, N. K. Chandra, Partha Chatterjee, Arvind N. Das, David Hardiman, Stephen Henningham, Gyanendra Pandey, and Sumit Sarkar -- were doing close empirical work in social, economic, and political history. The leader of the project, Ranajit Guha, was different. A "difference of generations," he says, "sets me apart ... by at least twenty-five years," but four other collaborators had also published books before 1982. His academic work sets him apart as sharply. His first book, A Rule of Property for Bengal: An Essay on the Idea of Permanent Settlement, was an intellectual history of colonial land policy. His published work in the seventies concerned intellectual trends surrounding one nineteenth century text, and his second monograph, Elementary Aspects of Peasant Insurgency distilled data from studies of peasant revolts in the colonial period to evoke a theory of subaltern resistance. Since 1982, his major publications have appeared first in Subaltern Studies, with which he is most personally identified. In his accumulated writings, colonialism appears to be a single, unified, discursive structure of power inside a vast ethnographic present; and state institutions, texts, personnel, and discourse, including those of the nationalist movement, stand in stark opposition to subaltern India and its indigenous culture from the first day of British rule down to the rupture of Subaltern Studies. Ranajit Guha might be said to be the Louis Dumont of colonialism, which in his writing attains a comprehensive power like that of caste in Homo Hierarchicus.

Even readers who applauded Subaltern Studies found two features troubling. First and foremost, the new substance of subalternity emerged only on the underside of a rigid theoretical barrier between "elite" and "subaltern," which resembles a concrete slab separating upper and lower space in a two-storey building. This hard dichotomy alienated subalternity from social histories that include more than two storeys or which move among them; and not only histories rendered through the lens of class analysis, because subaltern social mobility disappeared along with class differentiation. Secondly, because subaltern politics was confined theoretically to the lower storey, it could not threaten a political structure. This alienated subalternity from political histories of popular movements and alienated subaltern groups from organized, transformative politics, in the past and in the present. Not surprisingly, a rift soon opened between Subaltern Studies and Indian scholars committed to class analysis, political action, and popular histories of nationalism. Some critical responses appear in the first four reprints in this volume.

The meaning of subalternity in Subaltern Studies shifted as the framework of study increasingly stressed the clash of unequal cultures under colonialism and the dominance of colonial modernity over India's resistant, indigenous culture. Subalterns in India became fragments of a nation; their identity and consciousness reflected India's colonial subjugation. This approach has organized an impressive collection of enduring scholarship on colonial texts, vernacular resistance, bureaucracy, police, factories, communalism, ethnography, prisons, medicine, ethnography, science, and related topics. It has also enabled Subaltern Studies to speak as India's subaltern voice. Methodologically, recuperating subaltern subjectivity entails the analytical and rhetorical liberation of Indian culture from its domination by the colonial archive and by modernity. Ingenious methods for uncovering fragments of subaltern nationality became the project's particular speciality. Critical readings of colonial texts, oral histories, and ethnographic techniques are employed to reveal India's cultural roots in subaltern subjectivity. Subaltern Studies thus becomes a post-colonial critique of modern, European, and Enlightenment epistemologies. A new kind of cultural essence for India is found in iconic residues of hidden identities, expressions of difference, and misunderstood mentalities.

1.7 SELF-ASSESSMENT TEST

Discuss about the contribution of the Greeks and Romans during ancient times.

Discuss about the dualism and dichotomies in geography.

Discuss about the impact of quantitative revolution in geography

Discuss about the Critical Revolution in geography.

1.8 SUMMARIES AND KEY POINTS

Following points have been discussed in this section:

- Contributions of Greek, Roman and Indian scholars during the ancient period and Arab scholars during the medieval period
- Contributions of Humboldt and Ritter in Geography
- Social Darwinism and its importance in Geography; Morphology of cultural landscape (Carl O. Sauer)
- Major paradigms in Geography and their shift
- Dualism and Dichotomies in Geography: Physical and Human Geography, Regional and Systematic Geography, Ideographic and Nomothetic approach
- Positivism and Quantitative revolution in Geography
- System approach in Geography
- Critical revolution in Geography; Humanistic Geography; Radical Geography; Behavioural Geography
- Welfare Geography
- Feminism and Feminist Geography
- Postmodernism and Postmodern Geography
- Subaltern studies in Geography

1.9 STUDY TIPS

Fundamentals of Geographical Thought: Sudeepta Adhikari

Geographical Thought: A Contextual History of Ideas: R.D.Dikshit

Evolution of Geographical Thought: Majid Husain

Art and Science of Geography: Integrated Readings: R.D.Dikshit

Explanation in Geography: David Harvey

Disclaimer

This self-learning material is based on different books, journals and web sources.