

DISTANCE LEARNING MATERIAL



BINCY BADAL DWESH BHAWAN
DIRECTORATE OF DISTANCE EDUCATION



VIDYASAGAR UNIVERSITY

DIRECTORATE OF DISTANCE EDUCATION

MIDNAPORE - 721 102

M. A. / M. Sc. in Geography

PART - I

Paper : I • Module No. : I • Unit : 1, 2, 3, 4 & 5

Paper : I • Module No. : II • Unit : 1, 2, 3, 4 & 5

DIRECTORATE OF DISTANCE EDUCATION

VIDYASAGAR UNIVERSITY, MIDNAPORE



M.A./M. Sc. in GEOGRAPHY

PART - I

Paper - I ★ Module No. -I ★ Unit - I, 2, 3, 4 & 5

Paper - I ★ Module No. -II ★ Unit - I, 2, 3, 4 & 5

M.A. / M. Sc. In Geography

Part - I

Paper - I

Contents

<u>Paper</u>	<u>Module</u>	<u>Unit</u>	<u>Page No</u>
I	I	01	1
I	I	02	49
I	I	03	69
I	I	04	103
I	I	05	129
I	II	01	149
I	II	02	169
I	II	03	189
I	II	04	215
I	II	05	229

PART-I
Paper - I
Module - I : Unit - 01

Module Structure

Theories of the Origin of the Earth

- 1.1.0 : Introduction
- 1.1.1 : Objective
- 1.1.2 : The Origin of the Universe and the Big Bang
 - 1.1.2.1 : An Expanding Universe
 - 1.1.2.2 : The Early History of the Universe
- 1.1.3 : Formation of Stars
 - 1.1.3.1 : Evolutionary Cycles of Stars and Origin of Chemical Elements
 - 1.1.3.2 : Stellar Remnants
- 1.1.4 : Planetary Formation and Origin of the Earth
 - 1.1.4.1 : The Stages of Planetary Formation
- 1.1.5 : Summary

Study of the Interior of the Earth and the Earth's Crust

- 1.1.6 : Introduction
- 1.1.7 : Objectives
- 1.1.8 : Studies of Density Inside the Earth
- 1.1.9 : Studies of Gravity and Pressure Inside the Earth
- 1.1.10 : The Interior of the Earth as Revealed by Earthquakes
 - 1.1.10.1 : Seismic Waves
 - 1.1.10.2 : Refraction of Earthquake Waves in Earth's Interior
 - 1.1.10.3 : Determination of Earthquake Wave Velocities Inside the Earth
 - 1.1.10.4 : Some Important Seismological Observations During Early 19th Century
 - 1.1.10.5 : Refractions and Reflections of Seismic Waves in Earth's Interior
- 1.1.11 : Model of Earth's Internal Structure

Isostatic and Adjustment of the Earth

- 1.1.12 : Introduction
- 1.1.13 : Objective
- 1.1.14 : The Discovery of Isostasy
- 1.1.15 : Models of Isostasy
- 1.1.16 : Isostatic Compensation and Vertical Crustal Movements
- 1.1.17 : Summary

Doctrine of Uniformitarianism

- 1.1.18 : Introduction
- 1.1.19 : Objective
- 1.1.20 : The Doctrine and Related Concepts
- 1.1.21 : Contribution of Hutton and Lyell
- 1.1.22 : Summary
- 1.1.23 : Self Assessment Questions
- 1.1.24 : References

PART-I
Paper - I
Module - I : Unit - 01

THEORIES OF THE ORIGIN OF EARTH

1.1.0 INTRODUCTION

The earth is one of the eight planets that revolve around a rather average star, our sun. Along with the planets (Mercury, Venus, Earth- the **inner or terrestrial planets**; and Mars, Jupiter, Saturn, Uranus, and Neptune- the **outer or Jovian planets**) there is a belt of asteroids orbit the sun in the region between the paths of Mars and Jupiter. This grouping of planets and asteroids around the sun constitutes our solar system. There are also such star systems in the universe with these planets. Our solar system is a small part of a much larger aggregate of stars, planets, dusts and gases called a galaxy. Galaxies are also numerous in the universe, and their constituents are arranged into several different forms: tightly packed elliptical galaxies, irregular galaxies, and the more familiar discoidal spiral galaxies with their glowing central bulge and great curving arms. The galaxy in which our solar system is located is called the **Milky Way Galaxy**, a spiral type of galaxy. It rotates slowly in space once in every 240 million years. Our sun is located about 26,000 light years outward from the galaxy to its edge. Galaxies are grouped into clusters of 10-1,000 galaxies (galactic cluster). Elliptical galaxies predominate in regular clusters (which are tightly bound gravitationally), and spiral and irregular galaxies predominate in irregular clusters (which are loosely bound gravitationally). Galactic clusters are grouped into super-clusters (10,000-100,000 galaxies each). The super-clusters are then arranged in long, intersecting filaments, giving the universe a lacy appearance.

The whole dynamic Earth system is powered partly by the planet's own internal energy but mostly by energy from a star—the Sun—around which it revolves. Earth's internal and environmental systems are rooted in the way the planet evolved within our solar system, which itself evolved within the universe. Years of research by astronomers and physicists have produced a widely accepted theory of the origin of the universe, stars, and planets. Properly called the cosmic singularity theory, it is more commonly referred to as the theory of the "**Big Bang**."

According to the Big Bang theory, some 10 billion to 20 billion years ago an explosive cosmic event occurred from which all energy and matters are derived. Unimaginably hot, pure energy in the form of gamma rays filled the universe, cooling as it expanded and decaying to form subatomic particles of matter. [The physicist Albert Einstein (1879-1955) explained the basic equivalence of energy and matter as $E = mc^2$, in which E is the energy released by conversion of mass (m), and c is the speed of light, about 298,000 km per second.] Within a few minutes, these particles combined to form protons and neutrons, the building blocks of atomic nuclei. The number of protons in an atom determines its chemical identity, and neutrons add to the atom's mass.

The hydrogen nucleus, with only one proton and one neutron, was the first type of atomic nucleus to form. Because temperatures and densities were still extremely high, nuclear reactions occurred that fused pairs of hydrogen nuclei into single helium nuclei. By the time the universe was about three minutes old; it had expanded and cooled to the point at which nuclear fusion stopped. Of the atomic nuclei that already had formed, about 75 percent were hydrogen and 25 percent were helium. Gravity drew the hydrogen and helium together, condensing them and thereby heating them into glowing gas clouds from which, hundreds of millions of years later, galaxies of stars were born.

Within galaxies, dust and gas left over from initial star formation continuously gravitate into clouds called **nebulae** (the Latin word for "clouds"). Scientists have evidence that our Sun and solar system evolved from such a nebula. Every star, including our Sun, is composed mostly of hydrogen and helium held together by its own gravity and generating energy through nuclear fusion reactions at its center. Continued nuclear fusion and other reactions in the bodies of stars have created all the heavier elements in the universe. Nebulae condense to become stars, and planets are created as a by-product of star formation. Gravitational attraction draws the matter in a slowly rotating nebula into a flattened whirling disk, with the bulk of the matter at the center, where it develops into a star. Around this central star, the remaining materials in the nebula gravitate into larger bits of rock and metal that coalesce into planets orbiting the new star. Our solar system is believed to have formed this way about 4.6 billion years ago.

□ 1.1.1 OBJECTIVES :

From the above few paragraphs it is clear that, the phenomenon of earth's origin is basically a part of the episode of evolution of the universe. Hence, any discussion regarding the earth's origin would not be complete unless creation of the universe is addressed. The origin of the universe has been explained by Big Bang theory, which is a widely accepted theory among the scientific community. This Big Bang hypothesis led to more testable deductions. One such deduction was that the temperature in deep space today should be several degrees above absolute zero. Observations showed this deduction to be correct. In fact, the Cosmic Microwave Background Explorer (COBE) satellite launched in 1991 confirmed that the background radiation field has exactly the spectrum predicted by a Big Bang origin for the universe.

This part of the module is about the origin of the Earth, and is in three parts. The first part is about how we think the Universe formed, and the evidence for the Big Bang hypothesis for the origin of the Universe. The second part of the module is about galaxies and stars — how they form and how they are able to generate the chemical elements. In the third part of the module we consider the origin of planets and how we think our own planetary system formed.

□ 1.1.2 THE ORIGIN OF THE UNIVERSE AND THE BIG BANG

According to the Big Bang Theory, the Universe expanded to its present enormous volume from an initial miniscule starting volume. This expansion has taken place over the past 10 or so billion years. Two observations, both made during the 20th Century have profoundly shaped the way in which we think about our Universe and have lead to the Big Bang Theory for the origin of the Universe. First was the discovery by Edwin Hubble that our **Universe is expanding**. The Second was the discovery made in 1965 by Arno Penzias and Robert Wilson that our part of the Universe is filled with **microwave radiation**. This radiation has become known as the Cosmic Microwave Background

□ 1.1.2.1 An Expanding Universe

It was during the 1920's that Edwin Hubble provided the first evidence that, we live in an expanding Universe. Hubble discovered that there is a simple relationship between the distance to a remote galaxy and the **redshift** in the spectral lines from that galaxy. This redshift is known as the cosmological redshift. Hubble's observations showed that the greater the distance to a galaxy, the greater the redshift in its spectral lines. These measurements strongly indicated that galaxies appear to be moving away from us with speeds proportional to their distance. The net effect of this motion is that as time goes on the galaxies are getting further and further apart. Astronomers describe this as an expanding universe.

A very important consequence for observations which lead us to infer an expanding Universe is that at some point in the past the matter which is becoming increasingly distant must have been concentrated all in one place. Astronomers believe that at this point in time, the beginning of the Universe and the beginning of time, all the matter of the Universe was concentrated in an infinitely small volume and was in a state of infinite density. This time can be calculated by a number of different methods. The huge explosion which led to the expansion of the Universe is the event known as the Big Bang.

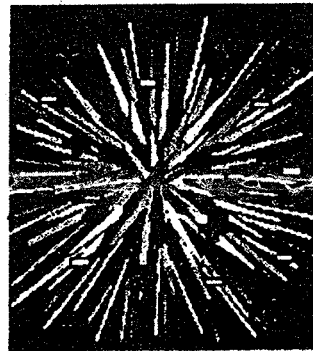
Three different techniques are currently used to calculate the age of the Universe- Velocity-distance relationships measured from galaxies, Radioactive dating methods and Models of galactic evolution and Models of stellar evolution. The three independent methods of dating all broadly agree to infer an age for the Universe of between **12 and 15 billion years**.

1.1.2.2 The Early History of the Universe

The history of the Universe may be described in the following stages :

- At the beginning of the Universe (about 13 billion years ago) all matter was in one place at a single instant. This event, in cosmological parlance, is known as a '**singularity**', a term which describes the inference that an infinitely large amount of matter is gathered at a single point in space-time, a point of extremely high, (perhaps infinite) density and extremely high (perhaps infinite) temperature. Space itself may collapse into that point. Inside that point, matter and energy will be indistinguishable.
- At the Big Bang there was a huge expansion of matter, an expansion which has continued ever since.
- For the first 3×10^{-10} s, temperature was too high for matter to be stable- only radiation would be stable. During this time temperature was infinite. The events immediately after the Big Bang blew up proceeded at a furious pace.
- The **Planckian**, the first aeon of cosmic time, which ranged from 0 to 5.390×10^{-44} s, is defined as the time it takes for light to cross the Planck length (1.616×10^{-35} m). Nothing is known about the Planckian.
- The Planckian was followed by the **Gamowian**, the second aeon of cosmic time, about which we know quite a bit. The Gamowian ranges from 5.390×10^{-44} second after time zero to 4.6 billion years ago, when the solar system formed. It is the longest aeon of cosmic time. During this time, the universe has been expanding. The rate of expansion has been either increasing, constant, or decreasing. Expansion involves the continuous creation of space between celestial objects that are not sufficiently bound to each other gravitationally, namely galactic clusters. During the earliest Gamowian, the four forces of nature were part of a single "**super force**". At the very high temperatures ($> 10^{31}$ K.) prevailing at time $t = 10^{-43}$ s, the mass of the particles was so large (because of their speed) that the gravitational force between the particles was as strong as the strong force. The other two forces, the electromagnetic and weak force, which are intermediate in strength between gravity and the strong force, follow suit. As a result, above 10^{31} K, the four forces are indistinguishable. As the universe expanded, temperature decreased, which caused the mass of the particles to decrease and gravitational interaction to weaken. Gravity, therefore, was the first force to appear as a separate force. Next to separate was the strong force, followed by the electro-weak force (which soon split into the electromagnetic force and the weak force).
- However, between 10^{-50} and 10^{-30} sec after the big bang there was a particularly rapid expansion of the Universe. This process is known as **inflation**. Later expansion slowed to a more normal rate.
- During inflation the part of the Universe that we see today expanded by a factor of 10^{60} . At this point the universe was permeated with radiation and subatomic particles (plasma/dark matter).
- Between the end of the Planckian (time $t = 5.390 \times 10^{-44}$ s, $T = 10^{32}$ K) and the beginning of inflation ($t = 10^{-33}$ s, $T = 10^{28}$ K), energy was transforming into matter + antimatter, and vice versa, because neither matter nor antimatter was stable at those high temperatures.
- When the universe inflated, temperature dropped, facilitating the transformation of energy into matter + antimatter. Following inflation, the first to stabilize were the quarks ($t = 5 \times 10^{-6}$ s; $T < 3 \times 10^{12}$ K).

- The quarks became confined and formed protons, neutrons, and mesons when the temperature dropped to about 10^{12} K.



PROTONS (positively charged particles) and ELECTRONS (negatively charged particles) existed during the early stages of evolution of the universe

- When the temperature dropped to 10^{10} K, the universe became dominantly filled with photon, electron and neutrino and their anti-particles and some protons and neutrons. As the temperature was reducing, electron and anti-electron collision was reducing because much of them were annihilated due to electron – anti-electron collision to produce photons. The collision between neutrino and anti-neutrino was very slower as they interact with each other very weakly.
- After about 100s of Big Bang ($T = 10^9$ K), protons and neutrons would have no longer sufficient energy to escape attraction of strong nuclear force. As such, they combined to produce deuterium or heavy hydrogen nuclei ($1p + 1n$). Deuterium nuclei collected another neutron and another proton to produce helium nuclei ($2p + 2n$) along with some lithium and beryllium nuclei. Thus one-fourth of the p and n were spent to produce nuclei of heavy hydrogen, helium and other elements while remaining p and n were spent to produce ordinary hydrogen.
- Thus the inflation was followed by the stabilization of ^3H and ^4He nuclei and of the electron.
- After 100,000 yrs conditions in the Universe were similar to those today inside the sun. An almost uniform plasma of electrons and hydrogen and helium ions filled the Universe. At this time the free electrons acted as a block to photons — generated from the light energy generated in the Big Bang, and prevented them escaping and rendering the early Universe opaque.
- After 300,000 yrs temperature dropped to 4500°K gave rise to atomic matter (hydrogen helium and deuterium) and because the electrons were now removed (in the atoms), radiation streamed out and the Universe became transparent.
- As the Universe began with a hot Big Bang, then the Universe is filled with electromagnetic radiation cooled from the early fireball to a temperature of around 10 degrees above absolute zero ($\sim 10\text{K}$). This is the microwave background radiation, the radiation remaining from the Big Bang that has redshifted all the way down to the microwave band. The spectrum of the microwave background radiation is that of blackbody radiation, at a temperature of 2.735 K.
- As the universe continues to expand the radiation will appear to be derived from a much cooler body, so that today the Cosmic Background radiation is 3.73 deg above absolute zero.

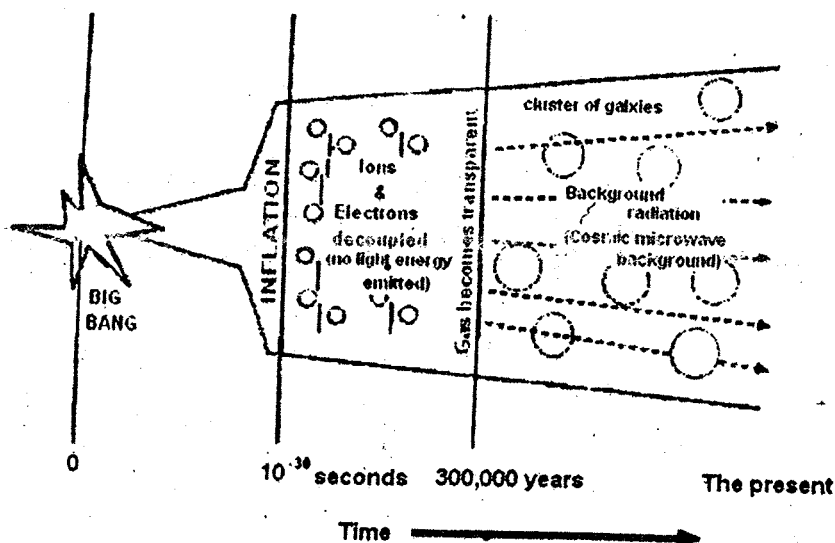


Fig-1: The origin, inflation and expansion of the universe, from Big Bang to the present

1.1.3 FORMATION OF STARS :

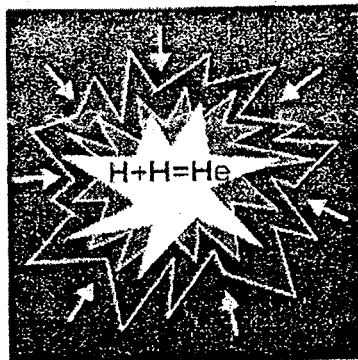
As the universe expanded, according to current scientific understanding, matter collected into clouds that began to condense and rotate, forming the forerunners of galaxies. Within galaxies, including our own Milky Way galaxy, changes in pressure caused gas and dust to form distinct clouds. In some of these clouds, where there was sufficient mass and the right forces, gravitational attraction caused the cloud to collapse. If the mass of material in the cloud was sufficiently compressed, nuclear reactions began and a star was born.

Lumpiness of the universe: Matter and energy were not distributed evenly in the very early Universe. This unevenness in the distribution of matter and energy has been described as the 'lumpiness' of the early Universe. The more scientific terminology is that the Universe shows 'anisotropy'. It is thought that these early inhomogeneities subsequently developed into the regions in the present Universe where there is matter (galaxies and galaxy clusters) and the regions from which matter is absent (space). The early inhomogeneous distribution of matter also reflects an inhomogeneous distribution of density. These initial density differences gave rise to small differences in gravitational forces which began to draw matter together.

Gravitational collapse and the Galaxies: As the Universe began to evolve regions began to emerge in which concentrations of the gases i.e. hydrogen and helium were higher than elsewhere. These formed the basis for the development of what we now recognize as galaxies. Today galaxies are huge concentrations of stars. Galaxies range in size from 80,000 to 150,000 light years in diameter. Within this huge volume, galaxies contain very large numbers of individual stars — ca. 100 billion. There are at least 10 billion galaxies in the Universe. Galaxies formed, perhaps within the first million years of the history of the Universe, by the inward gravitational collapse of the early matter of the Universe, thus enhancing the difference between regions where there was matter and regions where there was none. Inhomogeneities within galaxies, coupled with further gravitational collapse, are the basis for star formation. A galaxy starts to form by the accumulation of hydrogen gas in a very large cloud, called a **nebula**. As matter aggregates in a growing nebula the internal gravity draws in more gases. Eventually the nebula develops localized 'lumps' of gas which continue to grow into even denser gaseous bodies — stars.

Star formation: In astronomy, stellar evolution is the process by which a star undergoes a sequence of radical changes during its lifetime. Depending on the mass of the star, this lifetime ranges from hundreds of thousands to billions of years. Stellar evolution begins with a **Giant Molecular Cloud (GMC)**, also known as a stellar nursery. As a GMC orbits the galaxy, one of several events might occur to cause its gravitational collapse.

GMCs may collide with each other, or pass through dense regions of spiral arms. A nearby supernova explosion can also be a trigger, sending shocked matter into the GMC at very high speeds. Finally, galactic collisions can trigger massive bursts of star formation as the gas clouds in each galaxy are compressed and agitated by tidal forces.



Atomic fusion reaction to produce He and energy

As it collapses, a GMC fragments breaking into smaller and smaller pieces. In each of these fragments, the collapsing gas releases gravitational potential energy as heat. As its temperature and pressure increase, the fragments condense into rotating spheres of super hot gas known as **protostars**. Very small protostars never reach temperatures high enough for nuclear fusion of hydrogen to begin. These are known as **brown dwarfs**. For a more massive protostar, the core temperature will eventually reach 10 megakelvins, initiating the **proton-proton chain** reaction and allowing hydrogen to fuse, first to deuterium and then to helium.

In stars of slightly over 1 solar mass, the **Carbon cycle** contributes a considerable portion of the energy generation. The onset of nuclear fusion leads over a relatively short time to a **hydrostatic equilibrium** in which energy released by the core exerts a "radiation pressure" balancing the weight of the star's matter, preventing further gravitational collapse. The star thus evolves rapidly to a stable state.

Energy production in a star

Proton-Proton Chain

Reaction	Energy Produced (MeV)
$^1\text{H} + ^1\text{H} \rightarrow ^2\text{H} + e + \nu_e$	1.442
$^2\text{H} + ^1\text{H} \rightarrow ^3\text{He} + \gamma$	5.493
$^3\text{He} + ^3\text{He} \rightarrow ^4\text{He} + 2\ ^1\text{H}$	12.859
OR,	
$^3\text{He} + ^4\text{He} \rightarrow ^7\text{Be} + \gamma$	1.586
$^7\text{Be} + e \rightarrow ^7\text{Li} + \nu_e$	0.861
$^7\text{Li} + ^1\text{H} \rightarrow 2\ ^4\text{He}$	17.374
OR,	
$^3\text{He} + ^4\text{He} \rightarrow ^7\text{Be} + \gamma$	1.586
$^7\text{Be} + ^1\text{H} \rightarrow ^8\text{B} + \gamma$	0.135
$^8\text{B} \rightarrow ^8\text{Be} + e^+ + \nu_e$	15.04
$^8\text{Be} \rightarrow 2\ ^4\text{He}$	3.03

Carbon Cycle:

Reaction	Energy Produced (MeV)
$^{12}\text{C} + ^1\text{H} \rightarrow ^{13}\text{N} + \gamma$	1.943
$^{13}\text{N} \rightarrow ^{13}\text{C} + e^+ + \nu_e$	2.221
$^{13}\text{C} + ^1\text{H} \rightarrow ^{14}\text{N} + \gamma$	7.551
$^{14}\text{N} + ^1\text{H} \rightarrow ^{15}\text{O} + \gamma$	7.297
$^{15}\text{O} \rightarrow ^{15}\text{N} + e^+ + \nu_e$	2.753
$^{15}\text{N} + ^1\text{H} \rightarrow ^{12}\text{C} + ^4\text{He}$	4.966

1.1.3.1 Evolutionary cycles of Stars and Origin of Chemical Elements:

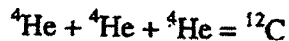
New stars come in a variety of sizes (from at least as low as 0.085 solar masses to more than 20 solar masses) and colors (from hot and blue to cool and red). A new star will fall at a specific point on the main sequence of the **Hertzsprung-Russell diagram** (H-R diagram). A mid-sized star like the Sun will remain on the main sequence for about 10 billion years. The Sun is thought to be in the middle of its lifespan: thus, it is on the main sequence. Once a star expends most of the hydrogen in its core, it moves off the main sequence. Once settled, a star spends most of its life at its place on the Main Sequence, fusing the hydrogen in its core into helium. Fusion is accompanied by a loss of mass (the Sun loses 4.257×10^6 tons of mass per second), but the loss is so slow that a star hardly moves down and to the right along the Main Sequence.

When hydrogen in the core of a star begins to be exhausted, energy production decreases, the core contracts, and temperature rises. Larger and hotter stars produce helium more rapidly than cooler and less massive ones. Eventually, the core exhausts its supply of hydrogen, and without the outward pressure generated by the fusion of hydrogen to counteract the force of gravity, it contracts until either electron degeneracy becomes sufficient to oppose gravity, or the core becomes hot enough (around 100 megakelvins) for helium fusion to begin. Which of these happens first depends upon the star's mass.

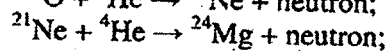
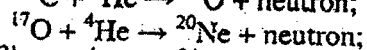
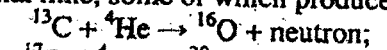
1. **Low-mass Stars:** In small stars (less than about 0.5 solar mass), the temperature rise is moderate, and the consequent energy release is slow enough for the energy to be radiated away. Their mass will never be able to fuse helium even after the core ceases hydrogen fusion. There simply is not a stellar envelope massive enough to bear down enough pressure on the core. These are the **red dwarfs**, such as Proxima Centauri. Red dwarfs of 0.1 solar masses may stay on the main sequence for almost six trillion years, and take several hundred billion more to slowly collapse into a **white dwarf**.

2. **Mid-sized Stars:** In middle-size stars, like the Sun, the temperature rises from 15 to 100 million Kelvin. The energy release is too fast for the energy to be radiated away, and the star expands 100-fold, which, in the case of the Sun, means that its surface will reach halfway between Mercury and Venus. Luminosity increases (because of the increased release of energy), but surface temperature decreases to perhaps 3,000 K (because of the expansion). The star becomes a **red giant**, moving up and to the right in the H-R diagram. When the Sun will reach the red-giant stage, an estimated 5 billion years in the future, the temperature on Earth will rise, all water will be vaporized, the atmosphere will be blown away, and all life (if there is any left by then) will be extinguished.

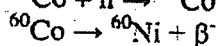
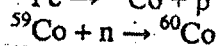
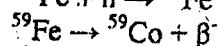
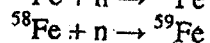
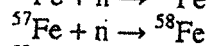
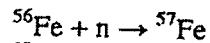
In spite of the expansion of the star in its outer part, its core collapses due to absence of sufficient hydrogen to be fused to produce energy that can resist the contraction of core. The collapse of the core raises the temperature to 10^8 K, which is high enough for two ^4He nuclei in the core to overcome their mutual repulsion and combine to form beryllium-8 (^8Be). ^8Be has a half-life of only 10^{-16} s, but in the dense state of matter in stellar cores, a third alpha particle may hit and be absorbed even within that short time. The result is a ^{12}C nucleus. This process is called the **triple-alpha process**, because it requires the practically simultaneous collision of three ^4He nuclei (which are particles). The formation of one ^{12}C nucleus from three ^4He nuclei destroys 0.00507 u of mass and produces 4.7 MeV of energy. ^{12}C is stable. In stars that are smaller than 2 solar masses, the process of helium burning begins abruptly and produces a helium flash.



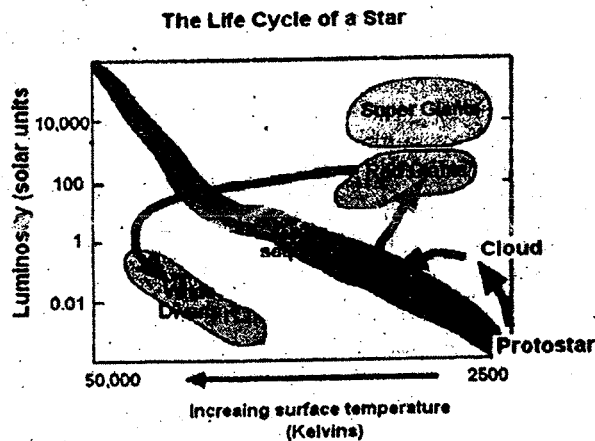
In the Sun, the fusion of helium to form carbon will last about a billion years. Many other nuclear reactions will occur during that time, some of which produce free neutrons:



These neutrons are absorbed by the nuclei of the various elements, creating heavier isotopes. If more than one or two neutrons are absorbed by a nucleus, the nucleus becomes unstable. A β^- particle is emitted and the next element up is formed.



In this way, all elements and isotopes heavier than boron, up to and including bismuth-209 (^{209}Bi , the heaviest stable isotope), are formed inside stars. This process is called the s process (s because it proceeds slowly). The process cannot continue beyond ^{209}Bi .



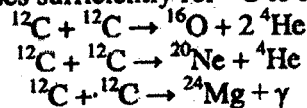
Eventually, all helium in the star's core will have fused into carbon. The star again stops producing energy, and the core again collapses. This time, a much greater amount of energy is released, causing the star to expand into a **red supergiant**, and its luminosity to increase 1,000 to 10,000 times. The Sun will expand beyond the orbit of Mars, and its mean density will decrease to $4 \times 10^{-8} \text{ g/cm}^3$. The hot solar plasma will engulf the Earth, reducing the

Earth's surface from reddish to gray. The orbital velocities of the inner planets will decrease because of drag against the solar plasma, and these planets will spiral into the expanded Sun. The frozen atmosphere of the outer planets will be vaporized and their rocky cores will be exposed. Many stars in our Galaxy are in the red-super-giant stage.

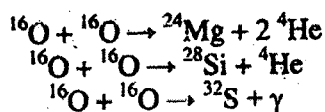
A red supergiant is so extended that it loses about half of its mass to outer space. The material is ejected as an expanding shell of gas. As long as the shell remains rather close to the mother star, UV radiation from the star excites the gas, and the gas appears luminous. From a distance, the gas is seen thickest along the periphery of the shell, forming a ring. What is left of the star is the **carbon core**, which is intensely hot, but not hot enough to fuse carbon. Nuclear reactions can no longer occur. As a result, the core shrinks, and its density rises to 10^6 g/cm^3 , forming a **white dwarf**.

3. Massive Stars: Stars with masses greater than 3 times the mass of the Sun, end up differently. Some have masses up to 100 times that of the Sun. Their cores are hotter, their nuclear reactions proceed a lot faster, and their surface temperatures may be as high as 25,000 K. or even more. The most massive ones shine a bluish light and are called **blue giants**. An example is the star Rigel.

i) When the triple-alpha process is completed in the core of a star with mass greater than 3 solar masses, the core contracts and the temperature rises sufficiently for ^{12}C to fuse and form ^{16}O , ^{20}Ne , and ^{24}Mg .



ii) If the mass of the star is greater than 8 solar masses, the temperature in the carbon core rises to 10^9 K , which is sufficient to fuse oxygen.



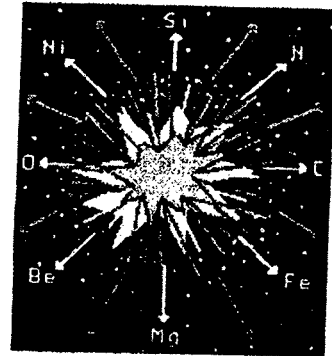


Formation of heavier elements within stars

At temperatures above 10^9 K, ^{20}Ne becomes unstable. The more energetic ^{20}Ne nuclei break up and release alpha particles, which are absorbed by other ^{20}Ne nuclei to form, in succession, magnesium-24 (^{24}Mg), silicon-28 (^{28}Si), sulfur-32 (^{32}S), argon-36 (^{36}Ar), and calcium-40 (^{40}Ca). The temperature in the core rises sufficiently to cause a number of nuclear reactions that produce a flood of alpha particles and free neutrons. These particles react with ^{40}Ca to form the iron group as the temperature in the core reaches about 4×10^9 K. The process stops with iron because; further fusion consumes energy rather than producing it. A star with a mass 25 times that of the Sun will exhaust the hydrogen in its core in a few million years, will fuse helium for half a million years, and—as the core continues contracting and the temperature continues rising—will fuse carbon for 600 years, oxygen for 6 months, and silicon for 1 day.

At the end of the process, a massive star will have a layered structure, with an iron core surrounded by concentric shells of progressively lighter elements. The iron core, like the carbon core in smaller stars, does not produce any more fusion. As a result, there is a catastrophic collapse, which takes less than one second. The amount of power released is so great that the star literally blows up in a **supernovae explosion**. Perhaps 90% of the star's original mass flies away in space, while the core is compressed by the accompanying implosion.

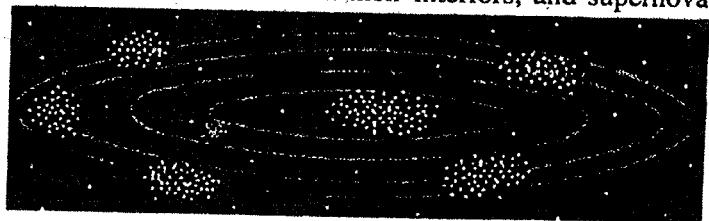
Supernovae explosion: A supernova explosion is so violent that many nuclei of the heavier elements are broken up. A lot of neutrons are freed, because the heavier elements are comparatively richer in neutrons than the lighter elements. The free neutrons are rapidly captured by lead and bismuth nuclei to form all heavier elements, up to thorium, uranium, and beyond. The process is so rapid (it is called the *r* process, where *r* stands for rapid) that the unstable elements between bismuth and thorium (polonium, astatine, radium, francium, radon, and actinium) have no time to decay. Other unstable elements, including- technetium, promethium, and the transuranic elements, are also made during supernova explosions. At the same time, spallation of heavier nuclei makes lithium, beryllium, and boron.



Elements are scattered into space by Supernova explosions

A supernova explosion has two important consequences. First, the new elements and isotopes formed in the star's interior and during the explosion are scattered in an expanding cloud of gases and dust and are mixed with the interstellar medium (which is mainly hydrogen and helium atoms). New stars, formed from this mixture, start with a complement of heavy elements, including thorium and uranium. As the process of formation of massive stars, synthesis of heavier elements in their interiors, and supernova explosions continues, the universe as a whole becomes progressively richer in heavy elements.

The Earth is rich in these elements because the solar system was born only 4.7 billion years ago, when the universe was already 10 billion years old. A second, important effect of a supernova explosion is that it can trigger the formation of new stars by send-



The scattered elements condensed into clusters. This led to the formation of new stars and planets. In the case of our own solar system this process likely began about 5 billion years ago.

ing shock waves through neighboring gas clouds.

1.1.3.2 Stellar remnants

After a star has burned out its fuel supply, and most of its mass is blown away by Supernova explosion, its remnants can take one of three forms, depending on the mass during its lifetime.

- i) **White dwarfs:** For a star of 1 solar mass, the resulting white dwarf is of about 0.6 solar masses, compressed into approximately the volume of the Earth. White dwarfs are stable because the inward pull of gravity is balanced by the degeneracy pressure of the star's electrons. The Sun, after its red-supergiant stage, will form a white dwarf similar in size to the Earth. If the core is less than 1.4 solar masses, the collapse will be stopped by electron pressure. This value, 1.4 solar masses for a carbon core, is called the **Chandrasekhar limit**, after the Indian astrophysicist Subrahmanian Chandrasekhar (b. 1910), who has studied the physics of stellar interiors. There are perhaps 10 billion white dwarfs in our Galaxy. White dwarfs keep radiating away their residual energy for many billions of years. They keep cooling off, becoming **brown dwarfs**, and eventually **dark dwarfs**.

- ii) **Neutron Stars:** In massive stars with cores between 1.1 (the Chandrasekhar limit for an iron core) and 3 solar masses, the implosion will force the electrons in the core to combine with protons and form neutrons. The core becomes a compact **neutron star**, 5 to 15km in radius, and with a density similar to that of nuclear matter (10^{14} g/cm^3). The implosion conserves angular momentum, and so a young neutron star rotates very rapidly. But the momentum gradually decreases due to interaction between the magnetic field of the neutron star and the ejected matter, which decelerates the rotation of the neutron star while accelerating the ionized particles in the ejected matter. As such, an old neutron star rotates so slowly that it becomes difficult or impossible to be detected. For all practical purposes it is dead.



Neutron Star

- iii) **Black holes:** If the mass of the core of a massive star is greater than 3 solar masses, the implosion forces matter to collapse into a single point at the center, called a **singularity**. There, density is infinite, and space and time are infinitely distorted. The gravitational field in the vicinity of a singularity is so strong that the escape velocity is greater than the speed of light. Not even light can escape. For this reason, the object is called a **black hole**. Only beyond a certain distance from the singularity can light begin to escape. The distance is called the **Schwarzschild radius**, after the German astronomer Karl Schwarzschild (1873-1916). A black hole consists of a singularity at the center, where mass is concentrated, and a spherical region of empty space around it from which light cannot escape. The surface of this sphere, whose radius is the Schwarzschild radius, is called the **event horizon**. Because light cannot escape from within a black hole, nothing can be known about it. There are ways by which black holes can be detected, however. Matter, in fact, can enter a black hole, simply by falling into it. The black hole is revealed by the radiation emitted by such matter as it falls in. If a black hole has a companion star, the presence of the black hole is revealed by the gravitational perturbation of the companion star. So far, only three objects in the sky have been tentatively identified as black holes. According to Stephen Hawking (b. 1942), the English physicist who developed this theory, black holes tend to "evaporate".

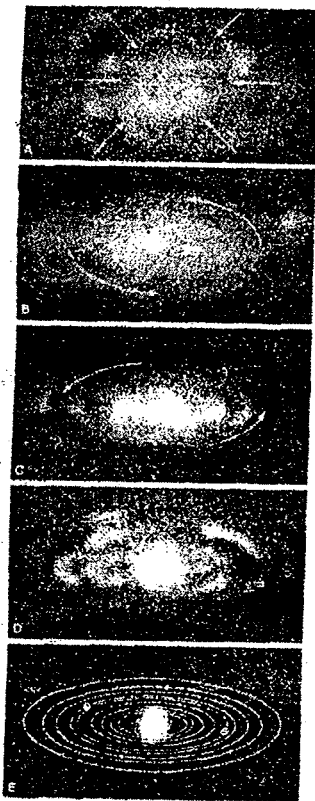
1.1.4 PLANETARY FORMATION AND ORIGIN OF THE EARTH :

In the previous section it was shown that matter in the form of the chemical elements formed during the process of star formation. In this section we discuss the processes whereby this early-formed matter is assembled into planetary bodies. A key concept here is the process of accretion — the progressive growth of larger bodies from smaller fragments. Thus the 'story' of planetary formation is one in which a very large number of small particles becoming progressively transformed into a small number of larger bodies, which eventually become planets.

The solar system is located about 25,000 light-years from the center of the Galaxy, near the inner edge of one of the spiral arms; it revolves around the galactic center at a speed of 225km/s. Formation of the solar system includes following stages.

1.1.4.1 The Stages of Planetary Formation:

1. **Gravitational collapse of cloud:** The Sun formed 4.7×10^9 y ago from a cloud of gas and dust (Nebula) whose collapse was triggered by a supernova explosion. The condensing gas and dust that went to form the Sun contained, all its original elements plus the elements formed during the supernova explosion. Those heavier than uranium, from plutonium to californium and beyond, rapidly fissioned into lighter elements or decayed into uranium or thorium. Uranium and thorium are also unstable, but their half-lives are so long that they survived in considerable amounts to the present day. Many unstable isotopes of the lighter elements were also formed by the supernova explosion (^3H , ^{26}Al , ^{36}Cl , ^{60}Fe , etc.). They also decayed rapidly into stable isotopes ($^3\text{H} \rightarrow ^3\text{He}$, $^{26}\text{Al} \rightarrow ^{26}\text{Mg}$, $^{36}\text{Cl} \rightarrow ^{36}\text{Ar}$, $^{60}\text{Fe} \rightarrow ^{60}\text{Co}$). These early radioactive decay processes contributed to raise temperatures within the Sun and the planetary bodies that were forming.



An illustration of the solar nebula hypothesis for the origin of the solar system. (A) The solar nebula takes form. (B) It begins to rotate, flatten, and contract. (C) The nebula is now distinctly discoidal and has a central concentration of matter that will become the sun. (D) Thermonuclear reactions begin in the primordial sun as material in the thinner area of the disk condenses and accretes to form protoplanets. (E) The late stage in the evolution of the solar system in which planets are fully formed, and materials not incorporated in planets are largely swept away by solar wind and radiation.

2. **Formation of planetary rings:** The cloud whose collapse formed the solar system probably had a very small net angular momentum. As the cloud began collapsing, conservation of angular momentum required it to increase its angular velocity. As a result, the collapse produced a central body (the early Sun) and a planetary ring. In addition to the usual large amounts of H and He, the planetary ring contained all the other elements, lots of molecules, and fine grains of solid matter.

Solid matter in interstellar space today — and most probably also in the gas cloud that formed the solar system—consists of microscopic crystals of frozen gases (including methane, ammonia, water, etc.), iron-magnesium silicates [i.e., minerals rich in silica (SiO_2) as well as in iron and magnesium], and metals (mainly iron and nickel). Most of the material appears to have crystallized into only a few major types of minerals- i) Olivine (a green mineral consisting of iron-magnesium silicate); ii) Pyroxenes (silicate minerals containing varying amounts of magnesium, iron, and calcium); iii) Kamacite (a body-centered cubic mineral of iron); iv) Taenite (a face-centered cubic mineral of iron and nickel).

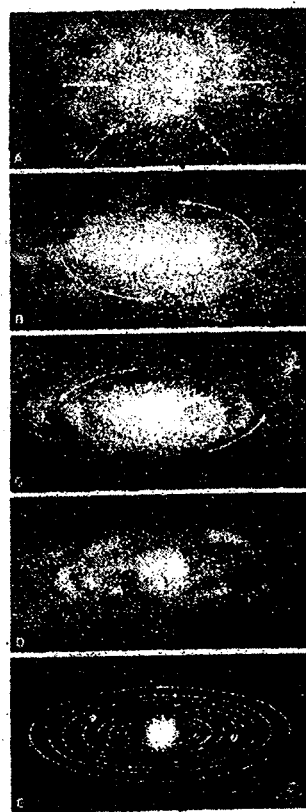
3. **Separation of Concentric rings:** Collapse of the cloud to form the Sun transformed a lot of gravitational energy into heat, most of which radiated away. The strong solar radiation ionized the gases in the planetary ring. The sweep of the Sun's magnetic field

speeded up the revolution of the ionized gases around the Sun, while slowing down the rotation of the Sun itself. During that period, the solar ring broke down into 10 concentric rings at specific radial distances from the Sun. The spacing of the rings has been summarized in a numerical formula called the Titius-Bode law which is given by the formula

$$r = 0.4 + 0.3 \times 2^n$$

Where, $n = -\infty$ for Mercury, 0 for Venus, and 1 to 8 for the planets from Earth to Pluto (?), including the asteroids.

4. **Planetesimals formation:** In addition to ionizing the planetary gases, the strong solar radiation raised the temperature within the planetary ring system. According to a model first proposed by Harold Urey in 1952. Water, methane, and ammonia liquefied and wetted the solid particles, which became stuck together and formed mud balls one kilometer or more across, called Planetesimals. Strong solar radiation blew all the gases that had not been trapped inside the planetesimals away from the inner rings toward the outer ones. In a few million years, the swarm of planetesimals within each ring agglomerated into a



An illustration of the solar nebula hypothesis for the origin of the solar system. (A) The solar nebula takes form. (B) It begins to rotate, flatten, and contract. (C) The nebula is now distinctly discoidal and has a central concentration of matter that will become the sun. (D) Thermonuclear reactions begin in the primordial sun as material in the thinner area of the disk condenses and accretes to form protoplanets. (E) The late stage in the evolution of the solar system in which planets are fully formed, and materials not incorporated in planets are largely swept away by solar wind and radiation.

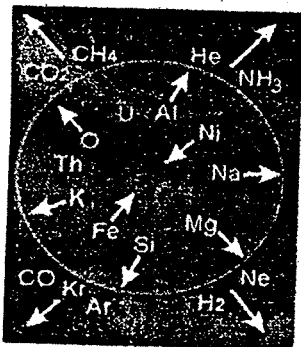
single body, with the largest planetesimal within each ring exerting the strongest gravitational attraction. The docking of the first few planetesimals did not develop much heat, because the gravitational field of the masses involved was small. As agglomerations proceeded, however, the gravitational field increased proportionally, causing agglomeration to become increasingly more rapid and also more violent. In a few million years, the swarm of planetesimals within each ring agglomerated into a single body, with the largest planetesimal within each ring exerting the strongest gravitational attraction. The docking of the first few planetesimals did not develop much heat, because the gravitational field of the masses involved was small. As agglomerations proceeded, however, the gravitational field increased proportionally, causing agglomeration to become increasingly more rapid and also more violent. Agglomeration soon became catastrophic, and in a short time, perhaps a few million years, the four inner rings formed the four inner planets (Mercury, Venus, Earth, Mars), and the five outer rings formed the five outer planets (Jupiter, Saturn, Uranus, Neptune and Pluto?). The fifth ring from the center failed to form a planet, but left as asteroids.

5. Differentiation of Earth:

a) **Cold Accretion Model:** The young planets must have consisted of a chaotic mixture of silicates, metals, and trapped gases. Although the center of each planet probably was cold, the bulk was hot because of- (i) the decay of the short-lived radioactive isotopes and (ii) collision of debris that coalesced to form Earth. At that time, the planets were certainly plastic, and possibly liquid. The temperature of the interior of Earth rose to about 2000°C , high enough to melt iron. Between 4.5 billion and 4.3 billion years ago, the intense heat of the primordial Earth melted much, or perhaps all, of its mass, initiating a process of differentiation that formed concentric layers with different chemical compositions and physical characteristics.

i) **Chemical Differentiation:** With Earth largely molten, most of its materials moved freely. Nickel and iron, two of the heaviest elements, sank toward Earth's center to form the dense core of the hot, liquid Earth. Lighter elements, particularly oxygen, silicon, aluminum, magnesium, calcium, potassium, and sodium, floated toward Earth's surface. When Earth's surface began to cool and solidify, these elements formed a rigid crust, which is Earth's thinnest and least dense rock layer. Between the core and the crust lies the mantle, a vast layer of rock rich in oxygen, silicon, iron, and magnesium and therefore of a composition and density intermediate between the crust's and the core's.

Formation of metallic cores in the planets overturned the interior of each planet and forced out the gases that had been trapped inside. These gases formed the early planetary atmospheres.

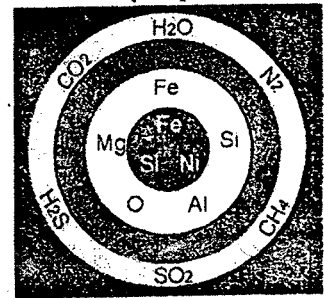


While it was still in the molten state, separation of elements occurred within the earth. Light inert gasses like Helium (He) and Krypton (Kr) were lost to outer space & heavy elements like Nickel (Ni) and Iron (Fe) sunk toward the centre.

The gases that came out from the interiors of the young planets had been in chemical high-temperature equilibrium with the metallic iron and the silicate minerals. They were, in order of abundance, hydrogen (H_2), helium (He), methane (CH_4), water (H_2O), nitrogen (N_2), ammonia (NH_3), and hydrogen sulfide (H_2S).

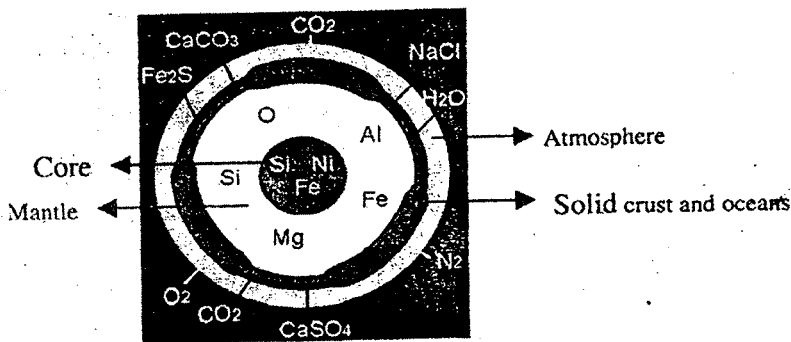
ii) **Physical Differentiation:** There are physical as well as chemical distinctions between Earth's layers, the result of differences in the structure of minerals and in the way they behave under different conditions of temperature and pressure. Heating a solid makes it less dense and more malleable. Compressing a solid makes it denser and more rigid. From Earth's surface to its inner core, both pressure and temperature increase. Below 70 km the temperature is high enough relative to the pressure of the overlying rock to cause the solid mantle to

become somewhat ductile, or plastic. At about 70 km, therefore, a boundary separates a rigid layer called the Lithosphere from the underlying, ductile layer, which is called the Asthenosphere, even though no compositional change or melting occurs. The Asthenosphere extends beneath the Lithosphere from 70 to 200 km. Lower than 200 km, increasing pressure results in a layer of dense, more rigid rock known as the Mesosphere (from the Greek word mesos for "in the middle"). Between 2891 and 5150 km, the residual temperature from when Earth melted is so high that its outer core is still molten, despite the great pressure from the overlying rock. At depths greater than 5150 km, however, the pressure is so immense that the inner core is solid, even though it is compositionally similar to the outer core and at least as hot.



Several concentric shells were formed in and around the earth, each one enriched in certain elements.

b) **Hot Accretion Model:** The idea of an initially cold birth for our planet has recently been reexamined as a result of new information emanating from the space programmes. On contrary to the cold accretion model, it is thought that the internal zonation of planets developed during, not after, accretionary accumulation by a process termed hot heterogeneous accretion. The hot heterogeneous model for accretion begins



As the earth cooled, a solid crust of silicate rocks formed more than 4 billion years ago. This sealed the hot magma inside.

with the solar nebula at a time when its gases were still very hot, possibly reaching temperatures of over 1000°C. It is possible that as it began to cool, heavy, iron-rich compounds condensed, precipitated, and accumulated in masses that would eventually become the iron-rich cores of the terrestrial planets. Next, with continued cooling, lower density silicate minerals began to condense and were swept up by the denser masses formed earlier. The silicates would have accumulated around the core as a thick covering that eventually became the mantle. During the accretionary process, the earth was hot and probably at least partially molten. In this state, chemical differentiation was able to proceed throughout the period of accumulation.

Currently, it is not possible to judge which model of differentiation is more valid. But most geologists appear to favour cold homogeneous accretion followed by melting as an explanation for the earth's internal zonation.

1.1.5 SUMMARY:

The history of the origin of the earth can be summarized as follows:

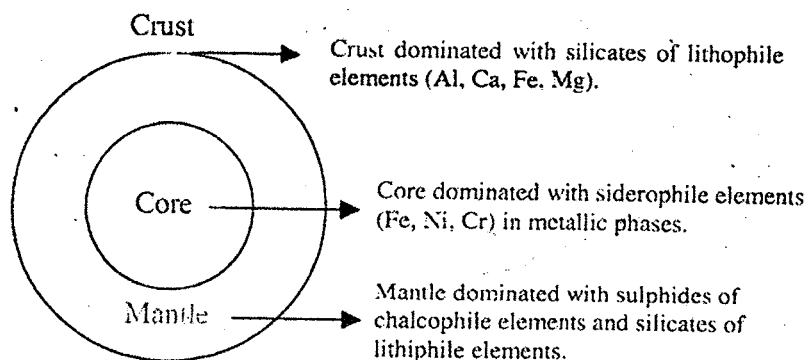
- Astronomers have developed a theory for the origin of the universe, the Big Bang, in which all matter and energy originated from a cosmic explosion 10 billion to 20 billion years ago.

- The universe consists of at least 100 billion galaxies; our **own**, the Milky Way galaxy, contains some 500 billion stars.
- Solar system, consisting of stars and orbiting planets, develop from coalescing gases and particles in a cloud of matter known as a nebula.
- Our solar system developed from a nebula about 4.6 billion years ago, and the Earth formed from bits of rocky and metallic matter that coalesced into single planet.
- Early earth heated up to about 2000° C as it formed, initiating a major period of complete or nearly complete- melting and resulting segregation of its elements into chemically distinct layers: core, mantle and crust.
- Because of changes in pressure and temperature with depth, Earth consists of layers that have distinctive physical properties: solid inner core, liquid outer core, solid mesosphere, semi soft asthenosphere and solid lithosphere.

STUDY OF THE INTERIOR OF THE EARTH AND THE EARTH'S CRUST

1.1.6 INTRODUCTION

Soon after the Earth formed, some unique processes occurred which made it different from the other planets in our Solar System. One such process was its division into metallic core, silicate mantle and crust. Way back in 1923, V. M. Goldsmith and his colleagues, analyzing vast amount of terrestrial and meteoritic mineral, found that various mineral coexisting in various types occur only as oxides, sulphides, silicates and metallic phases. Their observations also revealed that each of these groups is constituted of certain specific elements according to the cosmological affinities of those elements. The minerals contain a large number of other elements (may be < 0.1% in amount) along with the chief elements present in dominant amount. Though occurring in very small amounts, the elements showed specific chemical affinities to be concentrated in the metallic, sulphide, and silicate phases. All such elements which accompany iron (Fe) in the metallic phases are grouped as the **siderophile** elements (e.g.



Elemental differentiation in the Earth

Ni, Cr etc.) Fe occurs only as Fe^{+2} (reduced) in the metals while in the sulphides and silicates it may occur in both Fe^{+2} and Fe^{+3} forms. During the process of core formation, which should have occurred in the earliest phase of the earth's thermal evolution, all siderophile elements might have accompanied Fe^{+2} to form the metallic phase of the core. The separation of metallic phases should have taken place under reducing or hydrogen

dominated environment until enough hydrogen could be expelled outward. Then under hydrogen deficiency, the environment became rich in sulphur. Therefore, Fe combined with S to form Fe sulphide and began to be separated from the metallic phase. The elements, which accompanied Fe sulphide in such separation, in their respective sulphide phases are known as **sulphophile** or **chalcophile** elements (e.g. Mg). Chalcophile elements form covalent bonds with sulphur and thus combine readily with sulphur to

form minerals. Thus, the core formation began initially with the separation of siderophile elements and was subsequently followed by the separation of both siderophile and chalcophile elements to form the outer core. Thus most of the Fe and transition elements in the Periodic Table were consumed at this stage of earth's evolution to form metals and sulphides, leaving behind partly oxygenated and silica rich materials in the outer regions. At this stage the earth was in molten stage which had favoured the gravity separation of metallic and sulphide phases and upward displacement of the silicate phases. The elements which accompany Fe in the silicate phases are grouped as **lithophile** elements. Separation of the lithophile elements along with chalcophile elements accomplished the formation of mantle while lithophile elements formed the crust.

The above processes have given a layered structure to the earth similar to the other terrestrial planets. The Earth has an outer silicate solid crust, a highly viscous mantle, a liquid outer core that is much less viscous than the mantle, and a solid inner core. Many of the rocks now making up the Earth's crust formed less than 100 million (1×10^8) years ago; however the oldest known mineral grains are 4.4 billion (4.4×10^9) years old, indicating that the Earth has had a solid crust for at least that long.

Much of what is known about the interior of the Earth has been inferred. The force exerted by Earth's gravity is one measurement of its mass. After measuring the volume of the planet, its density can be calculated. Astronomers also have performed similar planetary measurements. Calculation of the mass and volume of the surface rocks and bodies of water allow estimation of the mass, volume and density of surface rocks. The mass which is not in the atmosphere, oceans, and surface rocks must be in deeper layers.

The structure of the Earth is separated into two categories: chemically differentiated layers and layers reflecting the strengths and density of the materials. **Chemically**, the Earth can be divided into the **crust, mantle, outer core, and inner core**. **By material strength**, the layering of the earth is categorized as **lithosphere, asthenosphere, upper mantle, lower mantle, outer core, and the inner core**. The geologic component layers of the Earth are at the following depths below the surface:

Depth (Kilometers)	Layer
0-60	Lithosphere (locally varies between 5 and 200 km)
0-35	Crust (locally varies between 5 and 70 km)
35-60	Uppermost part of mantle
35-2890	Mantle
35-660	Upper Mantle
100-200	Asthenosphere
660-2890	Lower Mantle
2890-5150	Outer core
5150-6360	Inner core

The layering of the Earth has been inferred indirectly using the time of travel of refracted and reflected seismic waves created by earthquakes. The core does not allow shear waves to pass through it, while the speed of travel (seismic velocity) is different in the other layers. The changes in the seismic velocity between the different layers cause refraction owing to Snell's law. Reflections are caused by a large increase in seismic velocity and are similar to light reflecting from a mirror.

1.1.7 OBJECTIVES :

A study of the Earth is necessary to understand many of the fundamental mechanisms underlying the Earth's dynamic surface. In depth studies of plate motions, volcanism, earthquake,

formation of the igneous rocks, tectonic landforms etc. depends on the understanding of earth's interior. The objective of this part of the module is to demonstrate the speculative indirect evidences gathered through experimental researches and from records of natural phenomena and to arrive into a reasonable model of the earth's interior. The study of the Earth's interior can be approached in two ways- i) By considering the Earth as part of the solar system and deducing indirectly such factors as its speed of rotation, density and mass, and ii) by analyzing direct information, particularly provided by seismic waves.

1.1.8 STUDIES OF DENSITY INSIDE THE EARTH

Consider a vertical prism between depths z and $z+dz$. The hydrostatic pressure increases from p at depth z to $(p+dp)$ at depth $(z+dz)$ because of the extra pressure due to the extra material in the small prism of height dz . The pressure increase dp is equal to the weight w of the prism divided by the area A of the base of the prism, over which the weight is distributed.

$$dp = \frac{w}{A} = \frac{(\text{volume} \times \rho)g}{A} = \frac{(Adz\rho)g}{A} = \rho g dz = -\rho g dr$$

The definition of bulk modulus, (K) is that, it is the ratio of the hydrostatic pressure to the dilatation (the fractional change in volume of an element, under pressure, in the limit when its surface area decreases to zero, which is given by V/dV). Then we can write,

$$K = -V \frac{dp}{dV} = \rho \frac{dp}{d\rho}$$

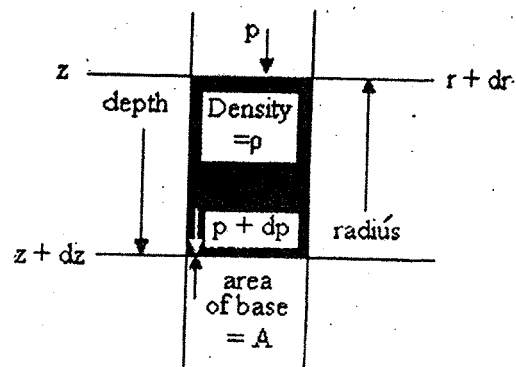
Φ is an important seismic parameter which is defined as defined as $\Phi = K/\rho$.

Combining the above equations,

$$\frac{-d\rho}{\rho g dr} = \frac{\rho}{K} = \frac{1}{\Phi}$$

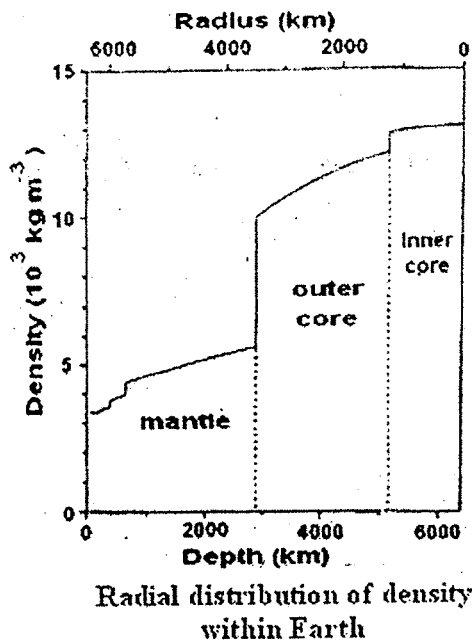
$$\frac{d\rho}{dr} = \frac{-\rho(r) g(r)}{\Phi(r)}$$

This is known as the Adams- Williamson equation. It was first applied to the estimation of density in the Earth in 1923 by E. D. Williamson and L. H. Adams. It yields density gradient at radius r , when the quantities on the right hand side are known. The seismic parameter Φ is known accurately from the velocities of the seismic waves, but the density ρ is unknown; it is in fact the goal of the calculation. The value of gravity g used in the equation must be computed separately for radius r . It is due only to the mass contained within the sphere of radius r , because external (homogeneous) shells of the Earth do not



contribute to gravitation inside them. This mass is the total mass E of the Earth minus the cumulative mass of all spherical shells external to r .

The procedure requires that a starting value for ρ be assumed at a known depth. Analyses of



crustal and upper mantle structure in isostatically balanced areas, give estimates of upper mantle density. Using this as a starting value at an initial radius r_1 , the density gradient in the uppermost mantle can be calculated from the above equation. Linear extrapolation of this gradient to a chosen greater depth gives the density ρ at radius r_2 ; the corresponding new value of g can be calculated by subtracting the mass of the shell between the two depths; together with the value of $\Phi(r_2)$ the density gradient can be computed at r_2 . This iterative type of calculation gives the variation of density with depth (or radius).

Density changes abruptly at the major seismic discontinuities showing that it is affected principally by changes in composition. Density is found to be around 4,500-5,500 kg m⁻³ in the mantle, 10,000-12,000 kg m⁻³ in the outer core and around 13,000 kg m⁻³ in the inner core. Between the major discontinuities, density increases smoothly as a result of the increases in pressure and temperature. A study of density, gravity and the Earth's dimension makes it possible to calculate the pressure within the Earth and also the temperatures that can be reached under this pressure.

1.1.9 STUDIES OF GRAVITY AND PRESSURE INSIDE THE EARTH

The radial variation of gravity can be computed from the density distribution. As stated above, the value of $g(r)$ is due only to the mass $m(r)$ contained within the sphere of radius r . Let the density at radius x ($\leq r$) be $\rho(x)$; the gravity at radius r is then given by

$$g(r) = -G \frac{m(r)}{r^2} = \frac{G}{r^2} \int_0^r 4\pi x^2 \rho(x) dx$$

A remarkable feature of the internal gravity is that it maintains a value close to 10 m s⁻² throughout the mantle, rising from 9.8 m s⁻² at the surface to 10.8 m s⁻² at the core-mantle boundary. It then decreases almost linearly to zero at the Earth's center.

Hydrostatic pressure is due to the force (N) per unit area (m²) exerted by overlying material. The SI unit of pressure is the pascal (1 Pa = 1 N m⁻²). In practice this is a small unit. The high pressures in the Earth are commonly quoted in units of gigapascal (1 GPa = 10⁹ Pa).

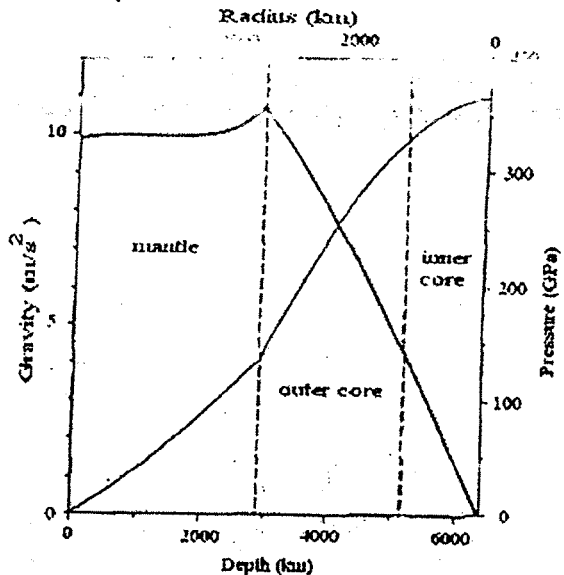
Within the Earth the hydrostatic pressure $p(r)$ at radius r is due to the weight of the overlying Earth layers between r and the Earth's surface. It can be computed by integrating the above equation using the derived distributions of density and gravity. This gives

$$p(r) = \int_r^R p(r) g(r) dr$$

Pressure increases continuously with increasing depth in the Earth. The rate of increase (pressure gradient) changes at the depths of the major seismic discontinuities. The pressure at the center of the Earth, which is 380 GPa (3.8 Mbar) at the center of the Earth, which is about 4 million times atmospheric pressure at sea level.

1.1.10 THE INTERIOR OF THE EARTH AS REVEALED BY EARTHQUAKES

The deep interior of the earth is inaccessible, and no rays of light penetrate to let us see what is below the surface. But rays of another kind penetrate and carry with them their messages from the interior. The Earth has been found to have elastic properties that allow propagation of the waves set up at the source (focus) of an earthquake to radiate into the interior and to spread over the surface. In a strong earthquake the whole of the Earth is set vibrating. At some distance from the focus, depending on the strength of the shock, the propagation of the waves is no longer perceptible, but sensitive seismographs can record the waves that emerge at the surface. The records provide data from which knowledge of the Earth's interior may be gained. Towards the end of the nineteenth century- it was realized that earthquake movements extend to great distances from the focus. Systematic recording of earthquakes began, and the very important results were gained..



Radial variation of internal gravity (thick curve) and pressure (thin curve)

1.1.10.1 Seismic Waves

The propagation of seismic disturbances through heterogeneous medium is very complex. But that can be simplified by assuming dividing it into parallel layers, in each of which homogeneous condition persists. The seismic waves propagate by elastic displacements in the medium. At or near the earthquake focus the medium is destroyed. Particles of the medium displaced permanently from their neighbours; the deformation is inelastic. But when a seismic disturbance has traveled some distance away from its source, its amplitude decreases and the medium deforms elastically to permit its passage. The particles of the medium carry out simple harmonic motions and seismic energy is transmitted as complex set of wave motions.

Types of seismic waves:

Two main types of waves are resulted from an event of earthquake. When seismic energy is released suddenly at a point near the surface of a homogeneous medium part of the energy propagates through the body of the medium as seismic **body waves**. The remaining part of the seismic energy spread out over the surface as seismic **surface waves**, analogous to the ripples on the surface a pool of water into which a stone has been thrown.

• **Seismic body waves:** When the body wave reaches a distance from its source in a homogenous medium, the wave front (defined as the surface in which all particles vibrate with the same phase)

has a spherical shape, and the wave is called Spherical wave. As the distance from the source increases, the curvature of the spherical wave front decreases. At great distances from the source, the wave front is so flat that it can be considered as plane, and the seismic wave is then called plane wave. Two types of body waves are recognized:

- i) **P-waves or Longitudinal or Compressional waves:** These are the fastest among all the waves emanated from an earthquake centre and therefore, it arrives first to any recording station. As such, they are known as **Primary waves** or P-waves. The particles of the medium oscillate **back** and forward parallel to the direction of wave propagation. This results in the medium being alternatively stretched and condensed in this direction. Thus the particle motion in P-waves transmits energy as a sequence of rarefaction and condensation and the medium is alternatively dilated and compressed. Therefore, these waves are also known as Compressional or transverse waves.

The velocity of the p-waves is given by-

$$\alpha = \sqrt{\frac{K + \frac{4}{3}\mu}{\rho}}$$

where, α = velocity of P- wave, K = bulk modulus of the medium, μ = shear modulus of the medium and ρ = density of the medium. The above equation demonstrates that, P- waves can travel through solids, liquids and gases, all of which are compressible ($K \neq 0$). Liquid and gases do not allow shear. Consequently, $\mu = 0$, and the P-wave velocity in a fluid is given by-

$$\alpha = \sqrt{\frac{K}{\rho}}$$

- ii) **S- Waves or Transverse or Shear waves:** The transverse wave motion is akin to that seen when a rope is shaken. S or 'secondary' waves are transverse or shear waves, in which the motion of each particle is at right angles to the direction of propagation. The velocity equation of these waves is as follows.

$$\beta = \sqrt{\frac{\mu}{\rho}} \quad \text{where, } \beta = \text{the velocity of S-wave.}$$

The only elastic property that determines the velocity of the shear wave is the rigidity or shear modulus μ . In liquids and gases μ is zero and the shear waves cannot propagate. In solids, a quick comparison between α and β gives-

$$\alpha^2 - \frac{4}{3}\beta^2 = \frac{K}{\rho}$$

As by definition bulk modulus K is positive, α is always greater than β . therefore, S- waves are slower than the compressional waves and reaches at an observation station later.

• Seismic surface waves:

Besides transmitting body waves P and S, two other kinds of wave are possible associated with boundaries between materials of different velocities and densities. These are responsible for the third train of signals. They were first known collectively as L (long) waves because generally they have longer periods, but now the term surface waves is invariably used. Their prominence on seismograms recorded beyond a certain distance from the source is due to the fact that they spread out and disperse

their energy over a surface nearly according to $1/d$ (d being distance from the source), whereas the body waves, P and S, disperse their energies throughout a volume according to $1/d^2$, which rapidly becomes much smaller than $1/d$ for increasing values of d .

Surface waves travel with velocities rather similar to those of the S wave in the shallow layers, but since they travel in these layers throughout their path they arrive after the S wave except at short distances.

- i) **Rayleigh waves (L_R):** One of the two kinds of surface waves was predicted by the third Lord Rayleigh in 1887, 20 years before it was recognized on seismograms. In Rayleigh waves the motion of particles is in elliptical orbits in the vertical plane of propagation. This is similar to the motion of a float on the sea as waves pass beneath it, but at the surface the orbit of a Rayleigh wave rotates in the opposite sense; they are retrograde rather than prograde.
- ii) **Love waves (L_Q):** In the second kind of surface wave the motion is horizontal and at right angles to the direction of propagation - they are like horizontally polarized S waves. These waves were recognized on seismograms before they had been accounted for. It was an Oxford mathematician, A.E.H. Love, who explained them by an extension of Rayleigh's theory, and since then they have been known as Love waves.

□ 1.1.10.2 Refraction of Earthquake Waves in Earth's Interior

The earth has a multi-layered structure with numerous thin horizontal layers, each characterized by a constant seismic velocity, which increases progressively with increasing depth. A seismic ray that leaves the surface with angle i_1 will be refracted at each interface until it is finally refracted critically. The ray that finally returns to the surface will have an emergence angle i_1 . Snell's law applies to each successive refraction (e. g. at the top surface of the n th layer, which has a velocity V_n).

$$\frac{\sin i_1}{V_1} = \frac{\sin i_2}{V_2} = \dots = \frac{\sin i_n}{V_n} = \text{constant} = p$$

The constant p is called the ray parameter, its value depend on emergence angle (i_1) and velocity at the surface layer (V_1).

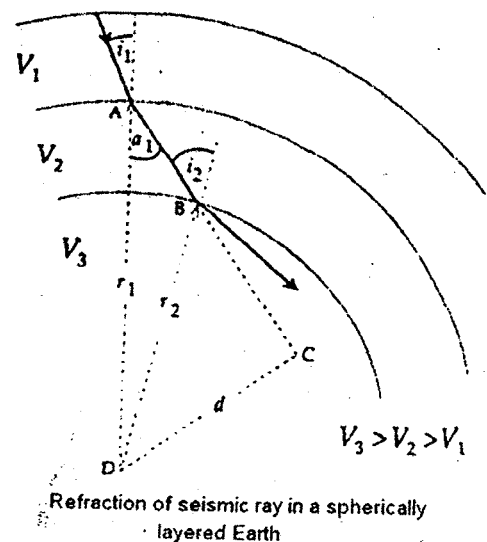
As the number of layers increases and the thickness of each layer decreases, the situation is approached in which the velocity increases constantly with increasing depth. Each ray then has a smoothly curved path. If the vertical increase of velocity is linear with depth, the curved rays are circular arcs. The vertical (radial) velocity structure can be represented by subdividing the Earth into concentric shells, each with a faster body-wave velocity than the shell above it. Snell's law of refraction applies to the interface between each pair of shells. For example, at point A we can write,

$$\frac{\sin i_1}{V_1} = \frac{\sin \alpha_1}{V_2}$$

Multiplying both sides by r_1 gives,

$$\frac{r_1 \sin i_1}{V_1} = \frac{r_1 \sin \alpha_1}{V_2}$$

In the triangles ACD and BCD, respectively, we get



$$d = r_1 \sin i_1 = r_2 \sin i_2$$

Combining the above three equations we get

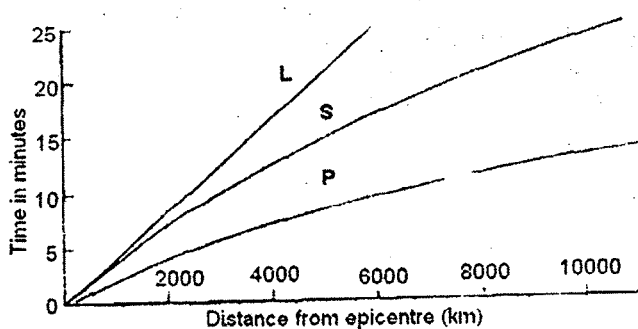
$$\frac{r_1 \sin i_1}{V_1} = \frac{r_2 \sin i_2}{V_2} = \dots = \frac{r_n \sin i_n}{V_n} = \text{constant} = p$$

Hence, if velocity increases continuously with depth, the seismic ray is refracted continuously and its shape is curved concavely upward.

1.1.10.3 Determination of Earthquake Wave Velocities inside the Earth

Access to the Earth's interior is provided by analysis of the travel-times of seismic waves that have traversed the various internal regions and emerge at the surface, where they are recorded. When the times taken by P and S waves to travel to seismological stations in various parts of the world are plotted against the respective distances from the epicentres, they fall on smooth curves which are nearly the same for all distant earthquakes originating at any particular depth, irrespective of the location of the paths between source and station. This shows that to a good approximation the wave velocities deep inside the Earth vary with depth only. The time interval between the arrival of P and S steadily increases with distance, and this is a simple consequence of the ratio of the P and S wave speeds being similar at all

Time-distance curves for P, S and L waves



depths. H. Jeffreys and K.E. Bullen measured arrival times for many earthquakes at many stations worldwide. They published their first set of travel time tables in 1935 that have been used fruitfully for the determination of earthquake locations until the present day.

When an earthquake is recorded at a distant station, measurement of the interval between the arrival times, S-P, serves to determine to a first approximation both the distance from the epicentre and the time of origin of the shock. At least three such distance

determinations at well-spaced stations are necessary. A circle is drawn on a globe around each of the three stations (e.g. Bombay, Tokyo, and Wellington), with a radius corresponding in each case to the respective distance of the epicentre. The epicentre lies at the point of intersection of the three circles. The travel-time (t) of a seismic ray to a known epicentral distance can be mathematically inverted to give the velocity of the wave, at the deepest point of the path.

1.1.10.4 Some Important Seismological Observations during Early 19th. Century

In 1897 E. Wiechert suggested that the interior of the Earth might consist of a dense metallic core, cloaked in a rocky outer cover. He called this cloak the 'Mantel', which later became anglicized to mantle. The progressive employment of seismic techniques world-wide led to the rapid development of the knowledge of earth's interior. Important results were hunted early in the 20th century.

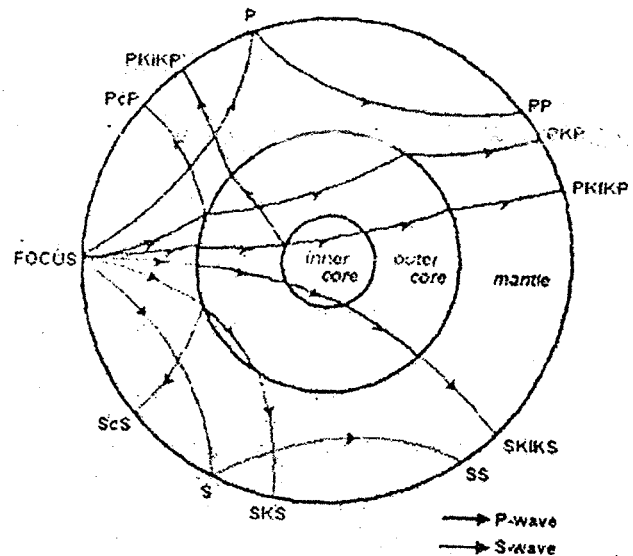
- I. The Earth's fluid core first detected seismologically in 1906 by R. D. Oldham. He observed that, if the travel-times of P-waves observed at epicentral distances of less than 100° were extrapolated to greater distances, the expected travel-times were less than those observed. This meant that the P-waves arriving at large epicentral distances were delayed in their passage through the Earth. Oldham inferred from this the existence, of a central core in which the P-wave velocity was reduced. He predicted that there would be a region of epicentral distances (a 'shadow zone') in which P-waves could not arrive. About this time it was found that P- and S-waves passed through the mantle but that no S-waves arrived beyond an epicentral distance of 105° .

II. In 1914 B. Gutenberg verified the existence of a shadow zone for P-waves in the range of epicentral distances between 105° and 143° . Gutenberg also located the depth of the core-mantle boundary with impressive accuracy at about 2900 km. A modern estimate of the radius of the core is 3485 ± 3 km, giving a mantle 2885 km thick. Gutenberg also predicted that P-waves and S-waves would be reflected from the core-mantle boundary. These waves, known today as PcP and ScS waves, were not observed until many years later. In honor of Gutenberg the core-mantle boundary is known today as the Gutenberg seismic discontinuity.

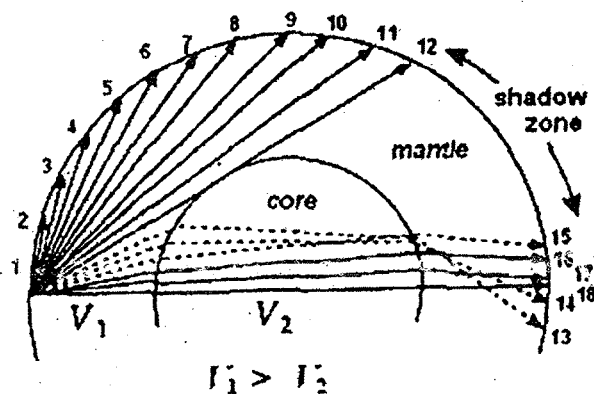
III. While studying the P-wave arrivals from an earthquake in Croatia in 1909 Andrija Mohorovicic found only a single arrival (P_g) at distances close to the epicenter. Beyond about 200 km there were two arrivals; the P_g event was overtaken by another arrival (P_n) which had evidently travelled at higher speed. Mohorovicic identified P_g as the direct wave from the earthquake and P_n as a doubly refracted wave (equivalent to a head wave) that travelled partly in the upper mantle. Mohorovicic calculated velocities of 5.6 km s^{-1} for P_g and 7.9 km s^{-1} for P_n and estimated that the sudden velocity increase occurred at a depth of 54 km. This seismic discontinuity is now called the Mohorovicic discontinuity or Moho for short. It represents the boundary between the crust and mantle. The crustal thickness is known to be very variable. It averages about 33 km, but measures as little as 5 km under oceans and as much as 60-80 km under some mountain ranges. The seismological Moho is commonly defined as the depth at which the P-wave velocity exceeds 7.6 km s^{-1} .

IV. In 1925 V. Conrad separated arrivals from a Tauern (Eastern Alps) earthquake of 1923 into P_g and S_g waves in an upper crustal layer and faster P^* and S^* waves that travelled with velocities 6.29 km s^{-1} and 3.57 km s^{-1} , respectively, in a deeper layer. Because the P^* and S^* velocities are significantly slower than corresponding upper mantle velocities, Conrad deduced the existence of a lower crustal layer. The interface separating the continental crust into an upper crustal layer and a lower crustal layer is called the Conrad discontinuity. Seismologists referred to the upper and lower crustal layers as the granitic layer and the basaltic layer, respectively.

V. The core shadow-zone and its interpretation in terms of a fluid core were well established in 1936 when Inga Lehmann, a Danish seismologist, reported weak P-wave arrivals within the shadow zone. She interpreted these in terms of an inner core with higher seismic velocity. Improved seismometer design, digital signal treatment and the setting up of seismic arrays have provided corroborating evidence.



Seismic wave paths of some important reflected and refracted P-waves and S-waves

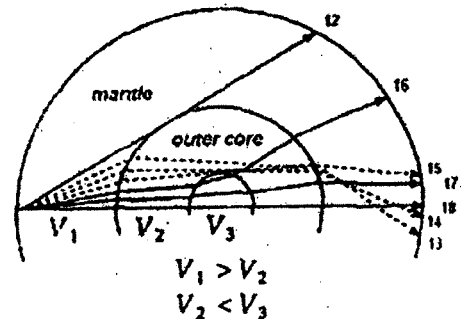


Seismic wave paths for P-waves

Rays 1-12 travel directly through the mantle as P-waves and emerge at progressively greater epicentral distances. Ray 13 is partially refracted into the core, and partially reflected (not shown in the figure). Because $V_2 < V_1$ the refracted ray is bent towards the earth's radius. This ray is further bent on leaving the core and reaches the Earth's surface as a PKP phase at an epicentral distance greater than 180° . Rays 14 and 15 impinge more directly on the core and are refracted less severely; their epicentral distances become successively smaller and their travel-times become shorter than ray 13. The paths of rays 16, 17 and 18 (which is a straight line through mantle and core and emerge at an epicentral distance of 180°) become progressively longer. No P-waves reach the surface in the gap between rays 12 (the last P-wave that just touches the core) and 15 (PKP-wave) - a **shadow zone** for P-waves. S-waves in the mantle follow the same ray paths as the P-waves. However, no direct S-waves arrive in the shadow zone, which indicates that the core must be fluid.

□ 1.1.10.5 Refractions and Reflections of Seismic Waves in Earth's Interior

The gross internal structure of the Earth is modelled as a set of concentric shells, corresponding to the inner core, outer core and mantle. An important step in understanding this layered structure has been the development of travel-time curves for seismic rays that passed through the different shells. If it possesses sufficient energy, a seismic disturbance may be refracted and reflected - or converted from a P-wave to an S-wave, or vice-versa - many times at the several seismic discontinuities within the Earth and at its free surface. As a result, the seismogram of a large earthquake contains numerous overlapping seismic signals and the identification of individual phases is a difficult task. To facilitate identification of the arrivals of these rays on seismograms a convenient shorthand notation is used. A P- or S-wave that travels from an earthquake directly to a seismometer is labelled with the appropriate letter P or S; until the margin of the core shadow-zone, the P- and

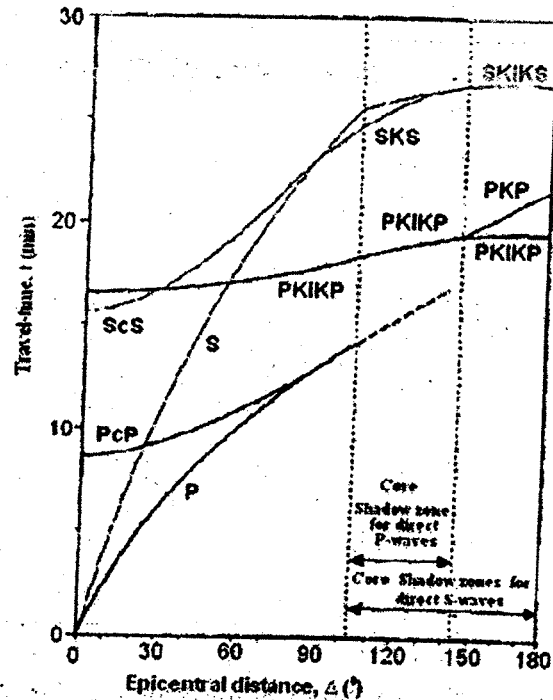


Seismic wave paths for P-waves (in Inner core)

Inner core has a constant velocity V_3 that is higher than the velocity V_2 in the outer core. The paths of rays 1-15 are the same as before, through the mantle and outer core. Ray 16 impinges on the inner core and is sharply refracted away from the Earth's radius; on returning to the outer core it is again refracted, back towards the radius. After further refraction at the core-mantle interface this ray emerges at the Earth's surface at a smaller epicentral distance than ray 15, within the P-wave shadow-zone, as a PKIKP event. Successive PKIKP rays are bent less strongly.

S-waves follow identical curved paths. A wave that reaches the seismometer after being reflected once from the crust is labelled PP (or SS), as its path consists of two identical P- or S- segments. The energy of an incident P- or S-wave is partitioned at an interface into reflected and refracted P- and S-waves.

A P-wave incident on the boundary between mantle and fluid outer core is refracted towards the normal to the interface, because the P-wave velocity drops from about 13 km s^{-1} to about 8 km s^{-1} at the boundary. After a second refraction it emerges beyond the shadow zone and is called a PKP wave (the letter K. stands for Kern, the German word for core). An S-wave incident at the same point has a lower mantle velocity of about 7 km s^{-1} . Part of the incident energy is converted to a P-wave in the outer core, which has a higher velocity of 8 km s^{-1} . The refraction is away from the normal to the interface. After a further refraction the incident S-wave reaches the surface as an SKS phase. A P-wave that travels through mantle, fluid core and inner core is labelled PKIKP. Each of these rays is refracted at an internal interface. To indicate seismic phases that are reflected at the outer core boundary the letter c is used, giving rise, for example, to PcP and ScS phases. Reflections from the inner core are designated with the letter i, as for example in the phase PKiKP.



Travel-time vs. Δ curves for some important seismic phases

□ 1.1.11 MODEL OF EARTH'S INTERNAL STRUCTURE

Once the velocity distributions of P- and S-waves inside the Earth were known the broad internal structure of the Earth - crust, mantle, inner and outer core - could be further refined. In 1940-1942 K. E. Bullen developed a model of the internal structure consisting of seven concentric shells. The boundaries between adjacent shells were located at sharp changes in the body-wave velocities or the velocity gradients. For ease of identification the layers were labelled A-G.

Regions(km)	Layer	Velocity of P-waves (km/s)	Density gm/Cm ³
A (0 - 33)	Crust:		2.7
			2.8
			2.9
	upper		Mohorovicic discontinuity
	lower		
B (33 - 410)	Upper Mantle:		
	uppermost mantle	7.9 - 8.1	3.32
	low-velocity layer	7.8	
	transitional zone	8.1	
C (410 - 1000)	Upper Mantle:		
		8.9	3.54
	transitional zones		
		9.13	3.72
D (1000 - 2900)	Lower Mantle:		
		10.27	3.99
	layer D'	10.75	4.32
	layer D"	13.71	5.57
E (2900 - 4980)	Outer Core		9.90
F (4980 - 5120)	Transitional Layer		12.2
G (5120 - 6370)	Inner Core		12.8
			13.1

The earth's layering: Compositional & Mechanical

Study of the propagation of earthquake waves reveals that the Earth has a layered structure on a broad scale, being split **compositionally** into the crust, mantle and core. Major compositional changes such as the crust-mantle and core-mantle boundaries involve large, abrupt changes in seismic velocity, leading to distinct reflected and refracted waves. However, the Earth is also layered **mechanically**, giving rise to more gradual changes in seismic velocity where an increase in pressure or temperature causes a change in the mechanical properties of the rock without changing its chemical composition. For example, the low-velocity zone is consistent with mantle material at or near its melting point (about 1300°C at pressures corresponding to the depth of the low-velocity zone), and therefore marks a major change in the mechanical behaviour of mantle material. Above this boundary the mantle is relatively cool and rigid, and can move as a single unit comprising the crust and part of the upper mantle, together known as the **lithosphere**. Below this boundary the temperatures and pressures are high enough to allow the sticky or viscous flow, of mantle material over long time periods, either completely in the solid state or with a small percentage of molten material in a form known as a partial melt. The mantle layer which can flow in this latter way is known as the **asthenosphere** and is associated with the seismic low velocity zone.

Other gradual changes in seismic velocity in the upper mantle are related to the dramatic changes in atomic structure, known as phase changes, which occur when a material is compressed too much for its previous atomic structure to bear. This results in a rearrangement of atoms within the minerals to accommodate the higher pressure, but again does not involve a change of composition. The most important phase change in the mantle are thought to be the transformation of olivine (believed to be the major component of the mantle) first to its spinel phase at a depth of about 400km, and then to its

perovskite phase at about 670 km. both of these depths are marked by slight increases in seismic velocity, more gradual than would be expected for a major compositional boundary.

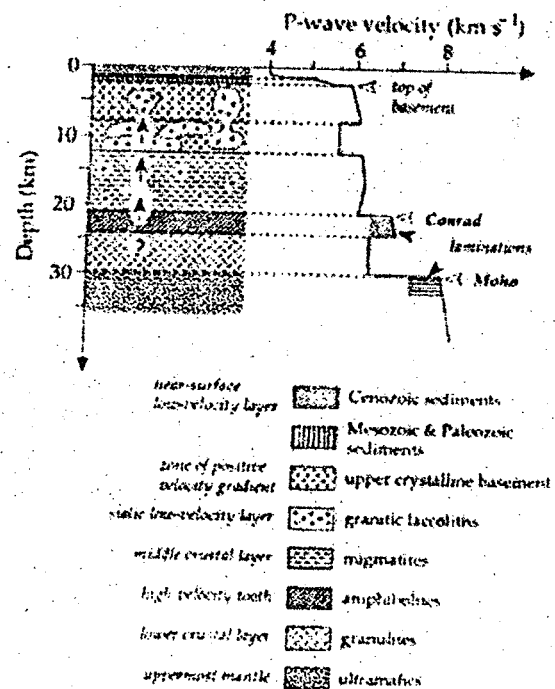
I. The Crust :

The crust makes up only 0.5 % of the Earth's total mass and can be subdivided into two main parts, continental and oceanic. Both differ in thickness, density and composition. Although oceanic crust covers approximately 61 % of the Earth's surface, it only comprises some 30 % of the crustal mass, as the continental crust is much thicker.

The Earth's crust corresponds to Bullen's region A. The structures of the crust and upper mantle are complex and show strong lateral variations. The most reliable information on crustal seismic structure comes from seismic refraction profiles and deep crustal reflection sounding. Ancient continental shield domains have different vertical velocity profiles than younger continental or oceanic domains. In view of this variability any generalized petrological model of crustal structure is bound to be an oversimplification. However, with this reservation, it is still possible to summarize some general features of crustal structure, and the corresponding petrological layering.

a) Continental crust: from the surface of the Earth down to 30 km - 50 km. The exposed parts of the continental crust are less dense than oceanic crust (this is because oceanic crust contains minerals rich in heavier elements such as iron and magnesium). However, continental crust appears to be stratified (layered) and becomes denser with depth. It is largely composed of volcanic, sedimentary and granitic rocks, although the older areas are dominated by metamorphic rocks.

• **Structure:** The vertical structure of continental crust is more complicated than that of oceanic crust, and the structure under indent shield areas differs from that under younger basins. The most striking difference is that the continental crust is much thicker than oceanic crust. Under stable continental areas the crust is 35-40 km thick and under young mountain ranges it is often 50-60 km thick. The continental Moho is not always a sharp boundary. In some places the transition from crust to mantle may be gradual, with a layered structure. The Conrad seismic discontinuity is believed to separate an upper crustal (granitic) layer from a lower crustal (basaltic) layer. However, the Conrad discontinuity is not found in all places and there is some doubt as to its real nature. Crustal velocity studies have defined two anomalous zones that often disrupt the otherwise progressive increase of velocity with depth. A low-velocity layer within the middle crust is thought to be due to intruded granitic laccoliths; it is called the **sialic low-velocity layer**. It is underlain by a middle crustal layer composed of **migmatites**. Below this layer the velocity rises sharply, forming a 'tooth' in the velocity profile. This tooth and the layer beneath it often make up a thinly layered lower crust. Refractions and reflections at the top of the tooth are thought to explain the Conrad discontinuity.



The continental crust can be divided into:

1. **Stable continental regions** (e.g. Precambrian cratons, platforms of undeformed sediments on

P-wave velocity - depth profile of continental crust

crystalline basements)

Thickness: 35 km - 45 km

Density: 2.69 gm/cm^3 - 2.74 gm/cm^3 for the upper crust and 3.0 gm/cm^3 - 3.25 gm/cm^3 for the lower crust.

- **Composition**

- Upper Crust (sial): felsic igneous and metamorphic rocks. Average composition close to that of either a mafic granodiorite or quartz diorite.
- Lower Crust (sima): anhydrous (no water) - quartz andesite or andesite; hydrous - amphibolite or diorite

2. **Tectonically active regions (e.g. the Andes, Himalayas)**

Thickness: 55 km - 70 km

- **Composition:** varies from region to region but may consist of any of the following:

1. mafic rocks
2. amphibolites
3. ultramafic, mafic and felsic-intermediate rocks

b) Oceanic crust:

Oceanic crust is only 5-10 km thick under a mean water-depth of about 4.5 km the top part of the oceanic crust consists of a layer of sediments that increases in thickness away from the oceanic ridges. The igneous oceanic basement consists of a thin (0.5 km) upper layer of superposed basaltic lava flows underlain by a complex of basaltic intrusions, the sheeted dike complex. Below this the oceanic crust consists of gabbroic rocks.

- **Structure:** Oceanic crust can be divided into:

1. **Ocean basins:** water depth exceeds 4 km, layered and very uniform and consists of three distinctive layers:

Layer 1: The uppermost layer is sedimentary with a thickness ranging from zero up to several kilometres. The thickness varies according to distance from continental margins. The thickest sediments occur close to the foot of continental slopes, and where large rivers have built sedimentary fans and cones that extend hundreds of kilometres out to sea. A relatively thick sedimentary layer also characterizes the equatorial belt which is a rich breeding ground.

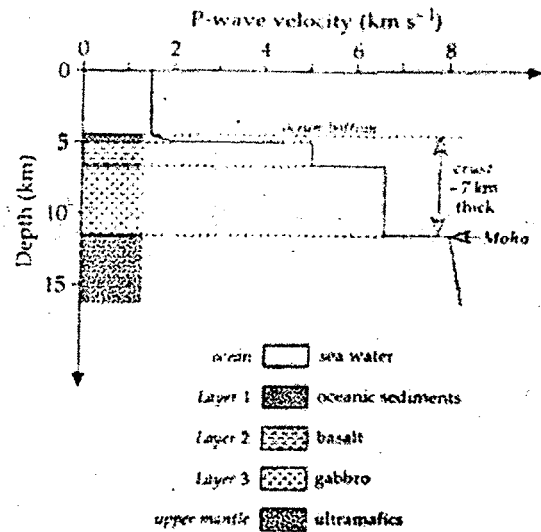
- Thickness of < 1 km; deep water sediments in various stages of lithification (turning into rock), commonly foraminiferal ooze, chert and mudstone.

Layer 2: Layer 2 is about 1 to 1.5km thick with P velocities that would fit many kinds of sedimentary as well as basaltic rocks. However, wherever samples of this layer have been brought up by dredging or drilling, or are exposed on land in an ophiolite sequence, it has been found to be basaltic. The layer is also highly fractured, allowing the penetration of seawater deep into the oceanic crust. Seismic profiles show layer 2 to have a rough upper surface, and consists of pillow lava, which forms when molten lava cools suddenly as it extrudes in 'blobs' directly into contact with seawater. The pillow lavas grade into a complex of vertical sheeted dykes of dolerite at the base of layer 2. These dykes are each approximately 0.5m wide, and are in direct contact, side by side.

- Thickness of 1.6 km - 2 km; basalts and dolerites with pillow lavas, dykes and sills (i.e. volcanic extrusive and intrusive rocks).

Layer 3: Layer 3, the lowest and main oceanic crustal layer has an average thickness of about 5 km, with P velocities indicative of a basic composition. The upper part is comprised of relatively coarse grained gabbro, with a sequence of inter-layered olivine gabbro and plagioclase peridotite forming at the base of layer 3. The base is much more similar in composition to the mantle than to continental crust, so the oceanic Moho is more gradational rather than a sharp discontinuity.

- Thickness of 3.0 km - 5.7 km; the main crustal layer; dolerite, gabbro and amphibolite (i.e. intrusive mafic rocks).
- **Layer 3B:** thickness is variable but up to 3 km; not always present; metagabbros and pyroxenites (i.e. intrusive mafic and ultramafic rocks).



P-wave velocity and depth profile of oceanic crust

2. Mid-ocean ridges

In mid-ocean ridges (such as the Mid-Atlantic Ridge), Layer 1 is absent; Layer 2 crops out on the surface and is thicker than normal; Layer 3 is thinner than normal and passes transitionally into the upper mantle.

3. Island arcs

The structure beneath island arcs (such as the Indonesian Arc) is very complex and the composition of the crust in these regions is very heterogeneous. The most common rock types in surface exposures are volcanic andesites (from explosive volcanoes) and deep-sea sediments.

III. The Moho :

The boundary between the crust and mantle is known as the Mohorovicic Discontinuity (Moho). It marks a significant change in chemical composition and is the boundary between the crust and mantle. It is a sharp boundary at which both P- and S-wave speeds rise from crustal to higher values. This change occurs over a distance of 0.1 km to 0.5 km in oceanic and stable continental regions. In tectonically active regions, the change is far more gradual and generally ill-defined - in many cases; there may not even be a Moho. At present there are two main hypotheses for what the Moho represents:

- A phase transition: The Moho may mark an isochemical phase transformation. (i.e. change in mineralogy but not chemistry) from a gabbroic lower crust to an eclogitic upper mantle. However, this phase transformation appears to be a gradual process in oceanic and stable continental regions.
- A chemical discontinuity: The most widely accepted theory for the Moho is that it represents a chemical change from intermediate and mafic crustal rocks, to an ultramafic mantle.

IV. The Mantle :

The mantle is thought to be primarily composed of ultrabasic rocks (rocks rich in magnesium and iron, and poor in silica; mostly peridotites). It is predominantly solid and overlies the Earth's iron-rich core, which occupies about 30% of Earth's volume. The mantle differs substantially from the crust in its mechanical characteristics and its chemical composition. The distinction between crust and mantle is based on chemistry, rock types, rheology and seismic characteristics. The crust is, in fact, a product of mantle melting. Partial melting of mantle material is believed to cause incompatible elements to separate from the mantle rock, with less dense material floating upward through pore spaces, cracks, or fissures, to

cool and freeze at the surface. As seismic velocity change through the mantle, we can subdivide it into an **upper mantle**, **transition zone**, and **lower mantle**. There is also a zone in the upper mantle that we call the **low velocity zone**.

a) The upper mantle (From Moho to 1000 km.):

In his 1942 model of the Earth's interior (see table 3.4) Bullen made a distinction between the upper mantle (layers B and C) and the lower mantle (layer D). The upper mantle is characterized by several discontinuities of body-wave velocities and steep velocity gradients.

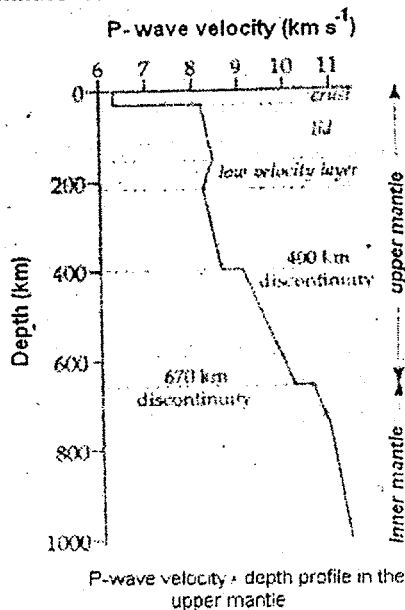
• Structure:

Uppermost Mantle: The top of the mantle is defined by the Mohorovicic discontinuity (Moho), below which the P-wave velocity exceeds 7.6 km s^{-1} . The Moho depth is very variable, with a global mean value around 30-40 km. The uppermost mantle between the Moho and a depth of 80-120 km is rigid, with increasing P- and S-wave velocities. This layer is sometimes called the *lid* of the underlying low-velocity layer. Together with the crust, the lid forms the lithosphere.

The Low Velocity Zone (LVZ): An abrupt increase of P- and S-wave velocities by 3-4% has been observed at around $220 \pm 30 \text{ km}$ depth; it is called the **Lehmann discontinuity**. It is not found everywhere and its true meaning is in question. Between the lid and the Lehmann discontinuity in the depth range 100-200 km, body-wave velocity gradients are weakly negative, i.e., the velocities decrease with increasing depth. The layer is called the low-velocity layer, (LVZ). Its nature cannot be evaluated from body waves, because they do not bottom in the layer and its lower

Composition of Earth's mantle in weight percent

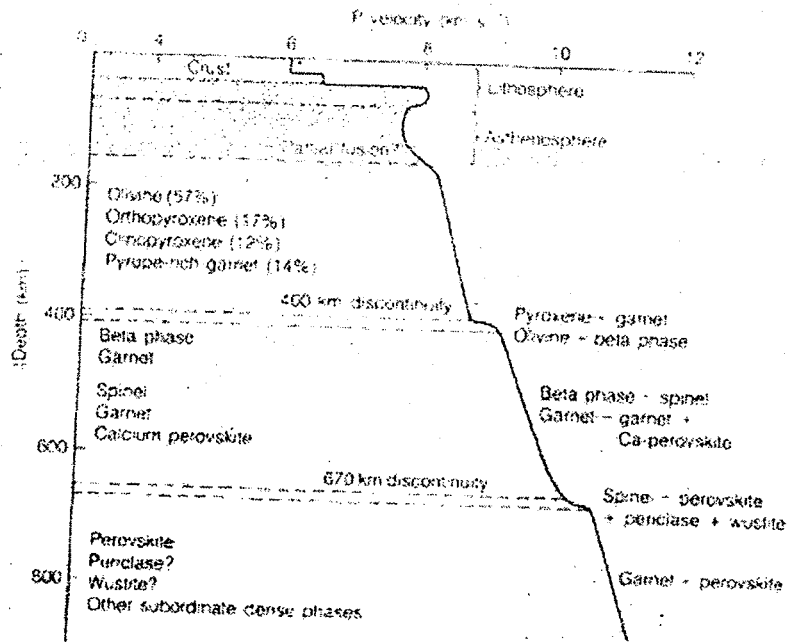
Element	Amount	Compound	Amount
O	44.8		
Si	21.5	SiO_2	46
Mg	22.8	MgO	37.8
Fe	5.3	FeO	7.5
Al	2.2	Al_2O_3	4.2
Ca	2.3	CaO	3.2
Na	0.3	Na_2O	0.4
K	0.03	K_2O	0.04
Sum	99.7	Sum	99.1



boundary is not sharp. (located at depths of 70 km to 250 km) where S-wave speeds decreases rapidly to a minimum and then gradually increases again. We believe that most magmas (molten rock) are generated in this zone. The top and bottom of the LVZ are not sharply defined. The LVZ is usually associated with the asthenosphere, which also plays an important role in plate tectonic theory. The decreases in seismic velocities are attributed to reduced rigidity in this layer. Over geological time-intervals the mantle reacts like a viscous medium, with a viscosity that depends on temperature and composition. From the point of view of plate tectonics the asthenosphere is a viscous layer that decouples the lithosphere from the deeper mantle, by allowing slow convection; it permits or promotes the relative motions of the global plates. The LVZ and asthenosphere reflect changes of rheological properties of the upper mantle material.

- **Density:** 3.25 gm/cm^3 to 3.40 gm/cm^3

- **Composition:** The composition of the upper mantle is generally taken to be peridotitic, with olivine ($(\text{Mg,Fe})_2\text{SiO}_4$) as the dominant mineral. With increasing depth the hydrostatic pressure increases and eventually causes high-pressure transformation of the silicate minerals. This is reflected in the seismic properties. Travel-time- Δ curves of body waves show a distinct change in slope at epicentral distances of about 20° . This is attributed to a discontinuity in mantle velocities at a depth of around 400 km. The 400 km discontinuity (or 20° discontinuity) is interpreted as due to a petrological change from an olivine-type lattice to a more closely packed spinel-type lattice. Thus the chemical composition of the upper mantle is characterized by Peridotite (e.g. olivine + pyroxene) along with plagioclase (< 30 km depth), spinel (30 km - 70 km depth) and garnet (> 70 km depth). In tectonically active regions, eclogite (amphibole + garnet) is a major component. Although it is thought that almost all basalts are derived from the upper mantle, experiments have shown that their compositions cannot be formed by the partial melting of peridotite alone. The upper mantle therefore is probably composed of two main zones: an **upper peridotite zone** and an underlying **primitive mantle or pyrolite zone** composed of pyroxene, olivine, and/or garnet and/or plagioclase, where basaltic magmas are generated.



Phase transition of constituents within earth's interior

Transition zone (400 km - 1000 km below the Earth's surface): It contains 17 % of the Earth's total mass. The lid, low-velocity layer and the zone down to the 400 km discontinuity together correspond to layer B in Bullen's earth Model A. A further seismic discontinuity occurs at a depth of 650-670 km. This is a major feature of mantle structure that has been observed world-wide.

- **Composition:** In the transition between the 400 km and 670 km discontinuities there is a further change in structure from β -spinel to γ -spinel. But this is not accompanied by appreciable changes in physical properties. At 670 km the spinel transforms [breakdown of olivine into its constituent oxide components of **periclase** (MgO) and **stishovite** (SiO_2) and garnet is broken-down to its component oxides] to perovskite. This transition zone corresponds to the upper part of Bullen's layer C.

V. The Lower Mantle:

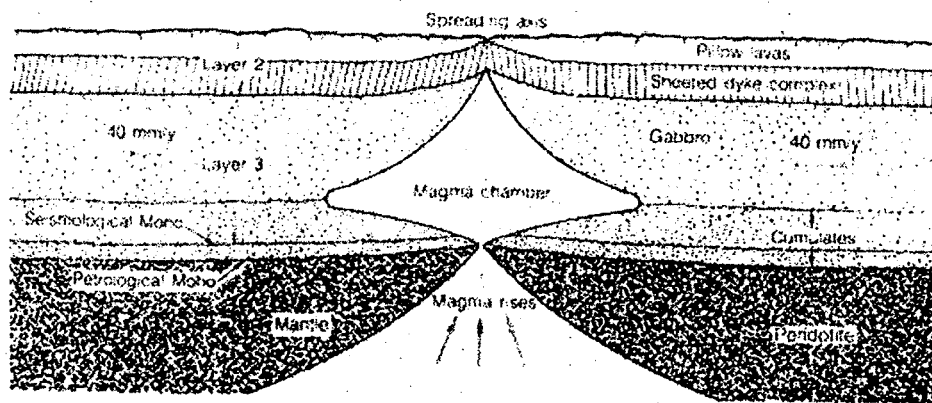
The lower mantle is now classified as the part below the important seismic discontinuity at 670 km.

- **Composition:** Its composition is rather poorly known, but it is thought to consist of oxides of iron and magnesium as well as iron-magnesium silicates with a perovskite structure.

• Structure

The D' Layer: The uppermost part of the lower mantle between 670 and 770 km depth has a high positive velocity gradient and corresponds to the lower part of Bullen's layer C. Beneath it lies Bullen's layer D', which represents a great thickness of normal mantle, characterized by smooth velocity gradients and the absence of seismic discontinuities.

The D" Layer (Core-Mantle Boundary/CMB Or Gutenberg Discontinuity): Just above the core-mantle boundary an anomalous layer, approximately 150-200 km thick, has been identified in which body-wave velocity gradients are very small and may even be negative. Although part of the lower mantle, it serves as a boundary layer between the mantle and core. Models of the internal structure of D" have been proposed with positive velocity gradients, others with negative velocity gradients, and some with small velocity discontinuities. The latter possibility is important because it would imply some stratification within D". The most interesting aspect of D' is the presence—revealed by seismic tomographic imaging—of velocity variations of several percent that take place over lateral distances comparable in size to the continents and oceans in Earth's crust. The term 'grand structure' has been coined for these regions; the thicker parts have also been termed 'crypto-



Cross-section of the structure of the crust and topmost mantle beneath an ocean ridge crest.

continents' and the thinner parts 'crypto-oceans'. Moreover, the seismically fast regions (in which temperatures are presumed to be cooler than normal) lie beneath present subduction zones. This suggests that cold subducted lithosphere may eventually sink to the bottom of the mantle where it is colder and more rigid than the surrounding mantle and hence has higher body-wave velocities. A large low-velocity (hot) region underlies the Pacific basin, in which many centers of volcanism and locally high heat flow 'hotspots' are located. The D' layer is suspected of being the source of the mantle plumes that cause these anomalies. Exceptionally hot material from D" rises in thin pipe-like mantle plumes to the 670 km discontinuity, which opposes further upward motion in the same way that it resists the deeper subduction of cold lithospheric slabs. Occasionally a hot plume is able to break through the 670 km barrier, producing a surface hotspot.

The D' layer plays an important role in geodynamic and geothermal behavior. On the one hand, D' serves as the source of material for the mantle plumes that give rise to hotspots, which are important in plate tectonics. On the other hand, the thermal properties of D" could influence the outward transport of heat from the Earth's core.

IV. The Core:

Early in Earth's history, dense metallic elements are thought to have settled towards the Earth's center, forming the core, while lighter silicates ascended and solidified to form the mantle.

• **Composition:** Studies of the meteorites give idea about crust's composition. It consists mainly of iron, with perhaps up to 10% nickel. The observed pressure-density relationships suggest that some less-dense non-metallic elements (Si, S, O) may be present in the outer core. It is not known if small amounts of the more common radioactive elements (^{40}K , ^{132}Th , ^{235}U and ^{238}U) are present in large enough abundances to contribute to the heat supply of the core.

• **Structure:**

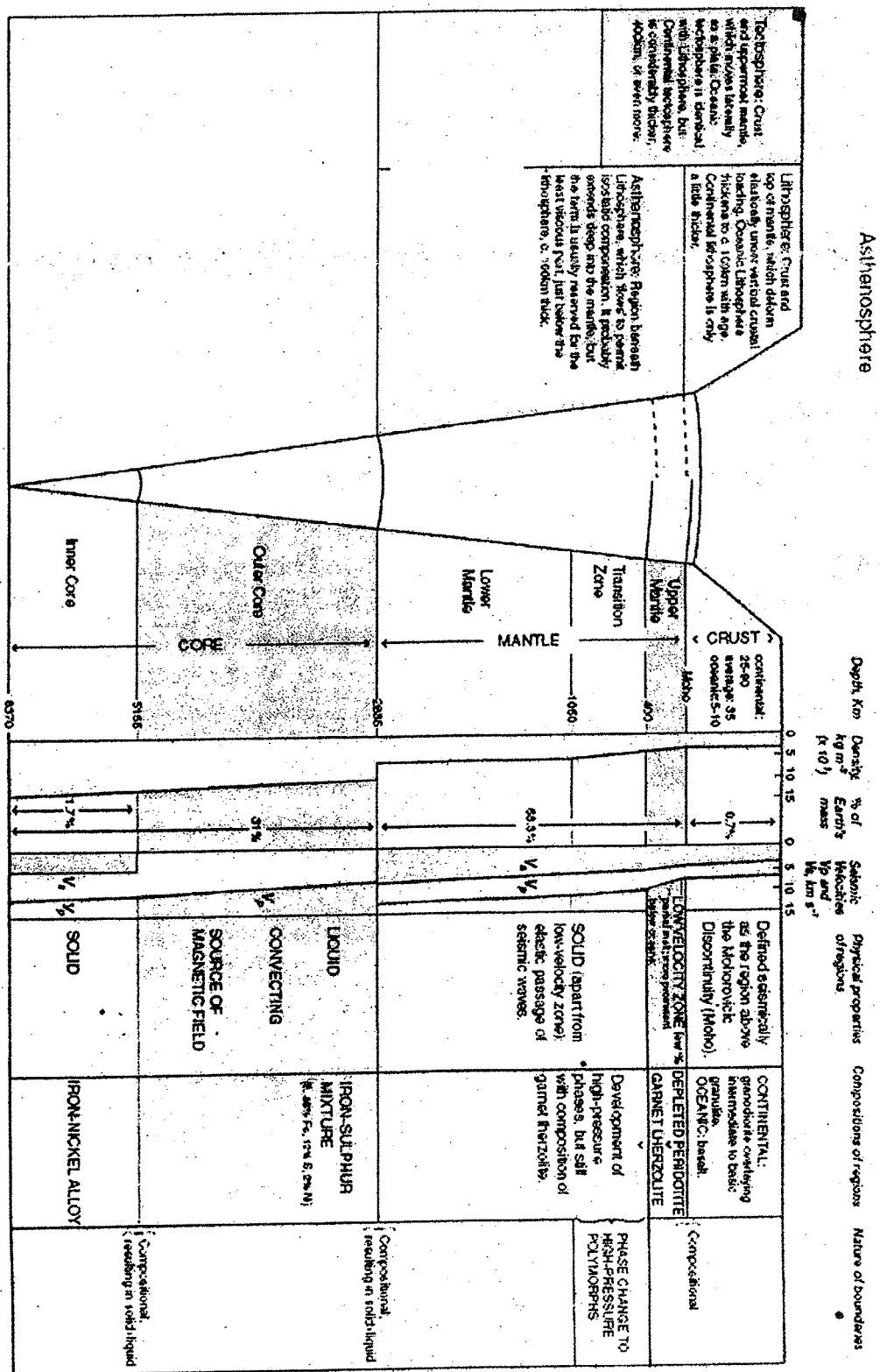
The core has a radius of 3480 km and consists of a solid inner core (layer G in Bullen's Earth Model A) surrounded by a liquid outer core (Bullen layer E) that is 1220 km thick. The transitional layer (F) was modelled by Bullen as a zone in which the P-wave velocity gradient is negative (i.e., a decrease with increasing depth).

Outer core: The outer core is fluid, with viscosity similar to that of water. It is assumed to be homogeneous and its thermal state is supposed to be adiabatic. These are the conditions to be expected in a fluid that has been well mixed, in the earth's case by convection and differential rotation.

Inner core: The inner core-outer core boundary is sharp, 5 km in thickness. Evidence supports to infer that the inner core is not completely solid, but it is a mixture of solid and liquid phases at a temperature close to the solidification temperature. It has been opined that the iron-rich inner core is solidifying from the fluid outer core. These constitute a less dense, gravitationally buoyant, fluid which rises through the denser overlying liquid.

2.7 SUMMARY :

The following diagram gives the summary of the modern scientific understanding about the earth's interior as discussed above.

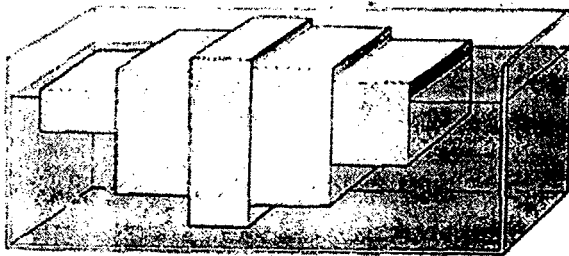


ISOSTATIC ADJUSTMENT OF THE EARTH

1.1.12 INTRODUCTION

The concept of isostasy is developed by the application of Archimedes' principle to the layered structure of the Earth. The elevated topography of Earth is roughly equivalent to an iceberg that floats in the surrounding, denser water. Just as an iceberg extends beneath the exposed ice, the concept of isostasy proposes that topography is supported, or compensated, by a deep root. The buoyant outer shell of the Earth, the crust, displaces the denser, viscous mantle in proportion to the surface elevation. Isostasy implies the existence of a level surface of constant pressure within the mantle, the depth of compensation. Above this surface the mass of any vertical column is equal. Equal pressure at depth can also be achieved by varying density structure or by the regional deflection of the lithosphere.

The term isostasy (Gr. *isostasios*, 'in equipoise') was proposed by Dutton, an American geologist, in 1889. The idea may be grasped by thinking of a series of wooden blocks of different heights floating in water. The blocks emerge by amounts which are proportional to their respective heights; they



Wooden blocks of different heights floating in water, an analogy to mechanism of Isostatic adjustment

are said to be in a state of hydrostatic balance. Isostasy is the corresponding state of balance between extensive blocks of the Earth's crust which rise to different levels and appear at the surface as mountain ranges, plateaus, plains or ocean floors. The idea implies that there is a certain minimum depth below sea-level where the pressure due to the weight of the overlying material in each unit column is everywhere the same. In Figure 2.9 this level of uniform pressure is that of the base of the highest block. The Earth's major relief is said to be compensated by the

underlying differences of density, and the level where the compensation is estimated to be complete - i.e. the level of uniform pressure - is often referred to as the level of compensation. Naturally, individual peaks and valleys are not separately balanced or compensated in this way; the minor relief features of the surface are easily maintained by the strength of the crustal rocks. Perfect isostasy is rarely attained, because of the restlessness of our globe; however, in regions which have remained free from geological disturbances for long periods of time there is generally a remarkably close approach to a state of equilibrium.

1.1.13 OBJECTIVE

The important point to be drawn from this discussion of isostasy is that the lithosphere acts as if it were "floating" on the asthenosphere. Sometimes gravity measurements suggest that a mountain has been pushed up so rapidly that it is top-heavy and has too little root of low density rock to counterbalance its upper mass. Sometimes, as in the seafloor trenches, it is observed that low-density crust has been dragged down so rapidly that it forms a root without a mountain mass above it. These and many other situations lead to local gravity anomalies. That the anomalies do not seem to become very large suggests that the Earth is always moving toward an Isostatic balance. Indeed, isostasy is the principal explanation for vertical motions of the Earth's surface, just as plate tectonics is the principal explanation for lateral motions.

□ 1.1.14 THE DISCOVERY OF ISOSTASY :

Isostasy was first recognized in the 18th century when a party of French geodesists was measuring the length of a degree of latitude in Peru in an attempt to determine if the shape of the Earth corresponds to an oblate or a prolate ellipsoid. Plumb lines were used as a vertical reference in the surveying and it was recognized that a correction would have to be applied for the horizontal deflection caused by the gravitational attraction of the Andes. When this correction, based on the mass of the Andes above sea level, was applied, however, it was found that the actual vertical deflection was less than predicted.

Similar discrepancies were met with during the survey of the Indo-Gangetic plain, south of the Himalayas, carried out by Sir George Everest, Surveyor-General of India, over a century ago. The difference in latitude between Kalianpur and Kaliana (603 km due north) was determined astronomically, and also by direct triangulation on the ground. The two results differed by 5.23 seconds of arc, corresponding to a distance on the ground of 168m. The discrepancy was ascribed to the attraction exerted by the enormous mass of the Himalayas and Tibet on the bob of the plumb line used for levelling the astronomical instruments. The error so introduced does not arise in the triangulation method. A few years later (1855) Archdeacon Pratt made a minimum estimate of the mass of the mountains and calculated what the corresponding gravitational effects would be at the two places on the plains to the south. His estimates for the deflections of the plumb line towards the mountains were 27.853" at Kaliana, and 11.968" at Kalianpur. The difference, 15.885", was more than three times the observed deflection, 5.23".

This phenomenon was attributed to the existence of a negative mass anomaly beneath the Andes which compensates, that is to say, supports, the positive mass of the mountains. In the 19th century similar observations were made in the vicinity of the Himalaya and it was recognized that the compensation of surface loading at depth is a widespread phenomenon.

The presence of subsurface compensation is confirmed by the variation in the Earth's gravitational field over broad regions. Bouguer anomalies are generally negative over elevated continental areas and positive over ocean basins. These observations confirm that the positive topography of continents and negative topography of oceans is compensated by regions at depth with density contrasts which are, respectively, negative and positive and whose mass anomaly approximates that of the surface features.

The principle of isostasy is that beneath a certain depth, known as the depth of compensation, the pressures generated by all overlying materials are everywhere equal; that is, the weights of vertical columns of unit cross-section, although internally variable, are identical at the depth of compensation if the region is in isostatic equilibrium.

□ 1.1.15 MODELS OF ISOSTASY :

Separate explanations of the anomalous plumb-line deflections were put forward by G. B. Airy in 1855 and J. H. Pratt in 1859. Their hypotheses have in common the compensation of the extra mass of a mountain above sea-level by a less-dense region (or root) below sea-level, but they differ in the way the compensation is achieved. In the Airy model, when isostatic compensation is complete, the mass deficiency of the root equals the excess load on the surface. At and below a certain *compensation depth* the pressure exerted by all overlying vertical columns of crustal material is then equal. The pressure is then hydrostatic as if the interior acted like a fluid. Hence, isostatic compensation is equivalent to applying Archimedes principle to the uppermost layers of the Earth.

In 1909-10, J. F. Hayford in the United States derived a mathematical model to describe the Pratt hypothesis. As a result, this theory of isostasy is often called the Pratt-Hayford scheme of compensation. Between 1924 and 1938 W. A. Heiskanen derived sets of tables for calculating isostatic corrections based on the Airy model. This concept of isostatic compensation has since been referred to as the Airy-Heiskanen scheme.

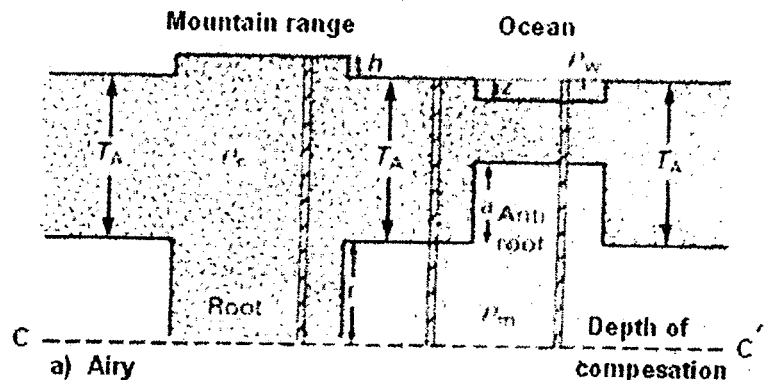
In 1931 F. A. Vening Meinesz, a Dutch geo-physicist, proposed a third model, in which the crust acts as an elastic plate. As in the other models, the crust floats buoyantly on a substratum, but its inherent rigidity spreads topographic loads over a broader region.

a) The Airy-Heiskanen Model

According to the Airy-Heiskanen model of isostatic compensation an upper layer of the Earth floats on a denser magma-like substratum, just as icebergs float in water. The upper layer is equated with the crust and the substratum with the mantle. The height of a mountain above sea-level is much less than the thickness of the crust underneath it, just as the visible tip of an iceberg is much smaller than the subsurface part. The densities of the crust and mantle are assumed to be constant: the thickness of the root-zone varies in proportion to the elevation of the topography.

The analogy to an iceberg is not exact, because under land at sea-level the normal crust is already about 30-35 km thick; the compensating root-zone of a mountain lies below this depth. Oceanic crust is only about 10 km thick, thinner than the normal crust. The mantle between the base of the oceanic crust and the normal crustal depth is sometimes called the *anti-root* of the ocean basin.

The Airy-Heiskanen model assumes local isostatic compensation, i.e. the root-zone of a mountain lies directly under it. Isostasy is assumed to be complete, so that hydrostatic equilibrium exists at the compensation depth, which is equivalent to the base of the deepest mountain root. The pressure at this level is due to the weight of the rock material in the overlying vertical column (of basal area one square meter) extending to the earth's surface.



The vertical column for the mountain of height h in Fig. contains only crustal rocks of density ρ_c . The pressure at CC' due to the mountain, 'normal' crust of thickness T_A and a root zone of thickness r amounts to $(h+T_A+r)\rho_c$. The vertical column below the 'normal' crust contains a thickness T_A of crustal rocks and thickness r of mantle rocks; it exerts a pressure $(T_A\rho_c + r\rho_m)$. For hydrostatic equilibrium the pressures are equal. Equating, then we get,

$$r = \frac{\rho_c}{\rho_c - \rho_m} h$$

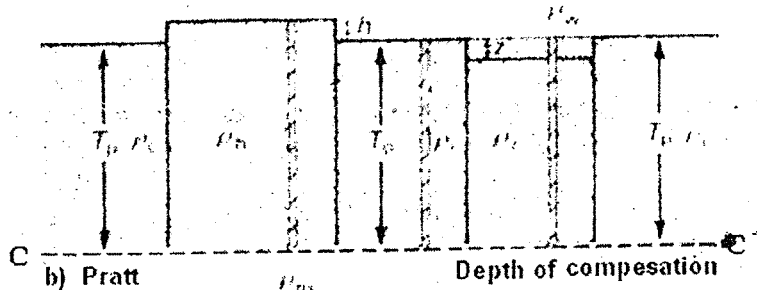
with a similar expression for the anti-root of the oceanic crust under an ocean basin of water depth z and density ρ_w is given by

$$a = \frac{\rho_c - \rho_w}{\rho_m - \rho_c} z$$

The Airy Heiskanen model assumes an upper layer of constant density floating on a more-dense substratum. It has root-zones of variable thickness proportional to the overlying topography. This scenario agrees broadly with seismic evidence for the thickness of the Earth's crust. The continental crust is much thicker than the oceanic crust. Its thickness is very variable, being largest below mountain chains, although the greatest thickness is not always under the highest topography. Airy-type compensation suggests hydrostatic balance between the crust and the mantle.

b) The Pratt-Hayford Model

The Pratt-Hayford isostatic model incorporates an outer layer of the Earth that rests on a weak magmatic substratum. Differential expansion of the material in vertical columns of the outer layer accounts for the surface topography, so that the higher the column above a common base the lower the



mean density of rocks in it. The vertical columns have constant density from the surface to their base at depth T_p below sea-level. If the rock beneath a mountain of height h has density ρ_h the pressure at CC' is $(T_p + h) \rho_h$. Under an ocean basin the pressure at CC' is due to water of depth z , and density ρ_w on top of a rock column of thickness T_p and density ρ_z , it is equal to $\rho_w z + \rho_z (T_p - z)$.

Equating pressures, we get

$$\rho_h = \frac{T_p \rho_c}{(T_p + h)}$$

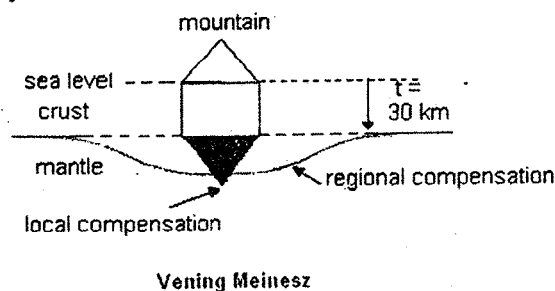
for ocean,

$$\rho_z = \frac{(T_p \rho_c - z \rho_w)}{(T_p - z)}$$

The Pratt Hayford and Airy Heiskanen models represent local isostatic compensation, in which each column exerts equal pressure at the compensation level.

c) Vening Meinesz elastic Plate Model

In the 1920s F. A. Vening Meinesz made extensive gravity surveys at sea. His measurements were made in a submarine to avoid the disturbances of wave motions. He studied relationship between topography and gravity anomalies over prominent topographic features. In 1931 he proposed a model of regional isostatic compensation which, like the Pratt-Hayford and Airy-Heiskanen models, envisages a light upper layer that floats on a denser fluid substratum. However, in the Vening Meinesz model the upper layer



behaves like an elastic plate overlying weak fluid. The strength of the plates distributes load of a surface feature (e.g. an island or seamount) over a horizontal distance wider than the feature. The topographic load bends the plate downward into the fluid substratum, which is pushed aside. The buoyancy of the displaced fluid forces it upward, giving support to the bent plate at distances well away from the central depression. The bending of the plate which accounts for the regional compensation in the Vening Meinesz model

depends on the elastic properties of the lithosphere.

1.1.16 ISOSTATIC COMPENSATION AND VERTICAL CRUSTAL MOVEMENTS :

In the Pratt-Hayford and Airy -Heiskanen models the lighter crust floats freely on the denser mantle. The system is in hydrostatic equilibrium. The additional depths of the root-zones below the local

isostatic compensation level are exactly proportional to the elevations of the topography above sea-level. The topography is then completely compensated. However, isostatic compensation is often incomplete. The geodynamic imbalance leads to vertical crustal movements.

Mountains are subject to erosion, which can disturb isostatic compensation. If the eroded mountains are no longer high enough to justify their deep root-zones the topography is isostatically *over compensated* causing the mountains to uplift. The opposite scenario is also possible. When visible topography has roots that are too small the topography is isostatically *under compensated* causing subsidence of the region.

• Isostatic effects of deposition and erosion

When large amounts of sediment are deposited on a particular region, the immense weight of the new sediment may cause the crust below to sink. Similarly, when large amounts of material are eroded away from a region, the land may rise to compensate. Therefore, as a mountain range is eroded down, the (reduced) range rebounds upwards (to a certain extent) to be eroded further. Some of the rock strata now visible at the ground surface may have spent much of their history at great depths below the surface buried under other strata, to be eventually exposed as those other strata are eroded away and the lower layers rebound upwards again.

An analogy may be made with an iceberg - it always floats with a certain proportion of its mass below the surface of the water. If more ice is added to the top of the iceberg, the iceberg will sink lower in the water. If a layer of ice is somehow sliced off the top of the iceberg, the remaining iceberg will rise. Similarly, the Earth's lithosphere "floats" in the asthenosphere.

• Isostatic effects of plate tectonics

When continents collide, the continental crust may thicken at their edges in the collision. If this happens, much of the thickened crust may move downwards rather than up as with the iceberg analogy. The idea of continental collisions building mountains "up" is therefore rather a simplification.

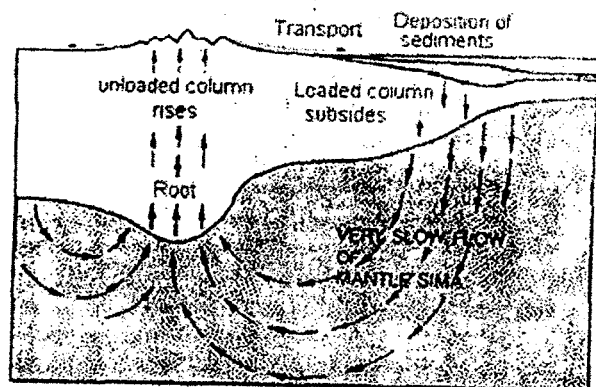
Instead, the crust thickens and the upper part of the thickened crust may become a mountain range. However, some continental collisions are far more complex than this, and the region may not be in isostatic equilibrium, so this subject has to be treated with caution.

• Isostatic effects of ice-sheets

The formation of ice-sheets can cause the Earth's surface to sink. Conversely, isostatic post-glacial rebound is observed in areas once covered by ice-sheets which have now melted, such as around the Baltic Sea and Hudson Bay. As the ice retreats, the load on the lithosphere and asthenosphere is reduced and they *rebound* back towards their equilibrium levels. In this way, it is possible to find former sea-cliffs and associated wave-cut platforms hundreds of metres above present-day sea-level. The rebound movements are so slow that the uplift caused by the ending of the last Ice Age is still continuing.

• Eustasy and relative sea level change

Eustasy is another cause of relative sea level change quite different from isostatic causes. The term "eustasy" or "eustatic" refers to changes in the amount of water in the oceans, usually due to global climatic changes. When the Earth's climate cools, a greater proportion of the earth's water is stored on land masses in the form of Glaciers, snow, etc. These results in a relative fall in global sea levels (relative to a stable



Isostatic readjustment by sub-crustal flow

land mass). The refilling of ocean basins by glacier melt water at the end of ice ages is an example of eustatic sea level rise.

A second significant cause of eustatic sea level rise is thermal expansion of sea water, when the Earth's mean temperature increases. Current estimates of global eustatic rise vary from as little as zero to +2 mm/a, its determination being difficult due to the need to separate out local vertical crustal movements.

When the term "relative" is used in context with "sea level change", the implication is that both eustasy and isostasy are at work, or that the author does not know which cause to invoke.

□ 1.1.17 SUMMARY

Isostasy is a concept used to describe how the principle of buoyancy applies to blocks of the earth's crust as they rest on the mantle. When British engineers were attempting to make a map of northern India near the Himalayan foothills during the mid-1800s, they discovered that their plumb bobs did not hang straight down but were deflected towards the Himalayas. Using the same observations, two engineers developed slightly different models to explain the plumb bob deflection in India. Pratt hypothesized that elevation is inversely proportional to density. Therefore, the higher the mountain, the lower is its density (i.e., light rocks "float" higher). Airy hypothesized that mountains have "roots" which extend down into the mantle. Therefore, elevation is proportional to the depth of the underlying "root". For many years, no one knew which model is a more accurate depiction of reality. But as geologists learned about the earth's interior (through studying shock waves created by earthquakes and nuclear explosions), it became apparent that Airy's model is a better one for most continental mountain ranges. However, Pratt's model is better at explaining why mid-oceanic ridges rise above the surrounding ocean floor, so both models have application to the "real" world. The principle of isostasy suggests that the earth's crust should adjust to any changes in mass that occur at the earth's surface (we call these "isostatic adjustments"). There are basically two types of responses:

1. SUBSIDENCE

- a. Definition: the slow, sinking of the earth's crust
- b. Cause: the addition of mass to the crust
- c. Example: the advance of glacial ice sheets

2. REBOUND

- a. Definition: the slow, vertical rise in earth's crust
- b. Cause: the removal of mass from the earth's crust
- c. Example: post-glacial rebound (Scandinavia)

DOCTRINE OF UNIFORMITARIANISM

□ 1.1.18 INTRODUCTION

Among the many important questions that have faced geologists is one that concerns the importance of small, slow changes like erosion caused by a single rainstorm, as opposed to massive, catastrophic changes, like earthquakes and floods that are infrequent and sporadic but cause rapid, dramatic changes to the landscape. During the seventeenth and eighteenth centuries, before geology became the scientific discipline it is today, people believed that all the Earth's features—mountains, valleys, and oceans—had been produced by a few great catastrophes. The catastrophes were thought to be so huge they could not be explained by ordinary processes; the supernatural had to be called upon. This concept came to be called

catastrophism. Not only were the catastrophes thought to be gigantic and sudden, but they were also thought to have occurred relatively recently and to fit a chronology of catastrophic events recorded in the Bible.

During the late eighteenth century, the concept of catastrophism was reexamined, compared with geological evidence, and found wanting. The person who assembled much of the evidence and proposed a counter theory was James Hutton (1726-1797). He was intrigued by what he saw in the environment around him, especially in Edinburgh where he studied and lived. He wrote about his observations and used logical arguments to draw conclusions based on evidence. In short, he investigated his environment using the scientific method. Hutton is widely regarded today as the father of modern scientific geology. In 1795 he published a two-volume work titled *Theory of the Earth with proofs and Illustrations* in which he introduced his counter theory to catastrophism.

□ 1.1.19 OBJECTIVE

All through its history, geology, as a scientific discourse, has tried to gain knowledge about the evolution of the features on the landscape and to interpret it. The most important technique that has been used to improve our understanding regarding the face of the earth is field observation. Interpretation of the field observations have led the geologists to develop concepts, theories and laws that can be successfully employed to explain the evolution of the landform features and landscapes. Ideas developed by one have been tested by others in the laboratory of the nature to collect evidences in support or against it. Accordingly, the history of geology/geomorphology has been found to be frequented with continued debates that have strengthened the scientific foundation of the discipline and augmented its growth as a separate branch of studies. One such incident was the discord between Catastrophism and Uniformitarianism. The objective of this part of the module is to discuss the basic premises and main contributors of the concept of uniformitarianism and to review its importance in the present context.

□ 1.1.20 THE DOCTRINE AND RELATED CONCEPTS

Uniformitarianism is the great underlying principle of modern geology. It was first enunciated by Hutton in 1785 beautifully restated Playfair in 1802 and popularized by Lyell in the numerous editions of his *Principles of Geology*. Hutton taught that the "present is the key to the past", but he applied this principle some what too rigidly and argued that geologic processes operated throughout geologic time with the same intensity as now. We know now that this is not true. Glaciers were much more significant during the Pleistocene and during other periods of geologic time than now, world climate have not always been distributed as they now are, and, thus, regions that are now humid have been desert and areas now desert have been humid; periods of crustal instability seem to have separated periods of relative crustal stability although there are some who doubt this; and there were times when volcanism was more important than now. Numerous other examples could be cited to show that the intensity of various geologic processes have varied through geologic time but there is no reason to believe that streams did not cut valleys in the past as they do now; that the more numerous and more extensive valley glaciers of the Pleistocene behaved any differently from existing glaciers that the winds which deposited the Navajo sandstone during Jurassic times obeyed any laws from those which control wind movements today. Groundwater opened up solutional passage ways in limestone and other soluble

rocks and formed surface depressions which we now call sinkholes during the Permian and Pennsylvanian periods as it does today in many parts of the world. Without the principle of uniformitarianism there could hardly be a science of geology that was more than pure description.

Uniformitarianism is one of the most basic principles of modern geology, the observation that fundamentally the same geological processes that operate today also operated in the distant past. It exists in contrast with catastrophism, which states that Earth surface features originated suddenly in the past, by geological processes radically different to those currently occurring. Note, however, that many "catastrophic" events are perfectly compatible with uniformitarianism. For example, Charles Lyell thought that ordinary geological processes would cause Niagara Falls to move upstream to Lake Erie within 10,000 years, leading to catastrophic flooding of a large part of North America. Uniformitarianism is a generalization of the principle of actualism, which states that present day-processes (astronomical, geological, paleontological,...) can be used to interpret past patterns. The principle of actualism is the cornerstone of paleoecology.

- **Uniformitarianism and rates of geological processes (an example- rock cycle):**

During the nineteenth century, geologists tried to estimate the duration of the rock cycle by estimating the thickness of all sediment laid down through geological time. They assumed that the Principle of uniformitarianism applied to rates of processes as well as the processes themselves, and hence that rates of deposition have always been constant and equal to today's rates. Thus, they thought, it would be a simple calculation to estimate the time needed to produce all the sediments. The results, we now know, were greatly in error. One of the reasons for the error was the assumption of the constancy of geological rates.

The more we learn of the Earth's history and the more accurately we determine the timing of past events through radiometric dating, the clearer it becomes that cycle rates have not always been the same. The evidence is strongly against constancy; some rates were once more rapid, others much slower.

One reason the rate of the rock cycle has changed through time is that the Earth is very slowly cooling down as its internal heat is lost. The Earth's internal temperature is maintained, in part, by natural radioactivity. Early in the Earth's history, more radioactive atoms were present than there are today, and so more heat must have been produced than is produced today. Internal processes, which are all driven by the Earth's internal heat, must have been more rapid than they are today. It is possible that a billion years ago oceanic crust was created at a faster rate than it is now, and that continental crust was uplifted and eroded at a faster rate. Either or both actions would cause the rock cycle to speed up.

At the same time, the rates of external processes have also varied. Long-term changes in the rates have occurred because of slow increases in the heat output of the sun, and also because of the gradual slowing in the rate of rotation of the Earth on its axis—scientists estimate that 600 million years ago there were 400 days in the year, and 2 billion years ago there were 450 days a year, for instance. Short-term effects on external process have also arisen because of changes in the orientation of the Earth's axis of rotation, and changes in the Earth-Sun distance. It is clear, therefore, that even though the cycles have been continuous, none has maintained a constant rate through time.

- **Role of weathering:** The slow reduction of solid rock to a particulate state by physical and chemical weathering before its removal by erosional agents is so important that it might be expected to be cornerstone of the uniformitarian edifice. By that time the details of the processes were unknown. For the catastrophists weathering was irrelevant to the creation of topography since it was a contemporary process. To the uniformitarians the fundamental forces for shaping the landform were much larger like fluvial or marine. Therefore, the processes of weathering were not seen as crucial by any of the opposing schools. But during late nineteenth century geomorphologists like Powell (1875), Gilbert (1877) recognized the importance of water as a solvent and thus the weathering processes were given due importance in shaping the landscape.

Similarly, evidences were also being collected by many scientists (Stanley, 1876 etc.) to document effects of thunderstorm, or sudden cooling of sun-heated rock surfaces in sudden changes in landforms.

□ 1.1.21 CONTRIBUTIONS OF HUTTON AND LYELL

Contributions of Hutton: The geologist James Hutton was a pioneer of the principle. Hutton, a Scot, was a physician and gentleman farmer. In the 18th and 19th centuries, the debate between the two theories- catastrophism and uniformitarianism was intense, since uniformitarianism seemed hard to reconcile with the prevailing religious beliefs of the time. Hutton observed the slow but steady effects of erosion: the transport of rock particles by running water and their ultimate deposition in the sea. He reasoned that mountains must slowly but surely be eroded away, that new rocks must form from the debris of erosion, and that the new rocks in turn must be slowly thrust up to form new mountains. Hutton didn't know the source of the energy that causes mountains to be thrust up, but everything, he argued, moved slowly along in a repetitive, continuous cycle. His ideas evolved into what we now call the Principle of Uniformitarianism, which states that the same external and internal processes we recognize in action locally have been operating throughout the Earth's history.

Contributions of Charles Lyell: Charles Lyell (1797 – 1875) dominated nineteenth century geology. He had travelled widely abroad and by this time he published the first volume of the 'Principles of Geology' in 1830. He made an attempt to explain former changes of the earth's surface by reference to the causes now in operation. He gave an explanation of the Playfair's key concept what Lyell termed the 'uniformity of physical laws'. Lyell stated 'that all former changes of the organic and inorganic creation are referrible to one uninterrupted succession of physical events, governed by the laws now in operation'. After publication of his book of Lyell one of the serious criticisms was leveled by his friend Scrope (1830) wrote- 'the value of all geological evidence... must depend entirely on the degree of confidence which we feel in regard to the permanency of the laws of nature. Their immutable constancy alone enables us to reason from analogy, by the strictest rules of induction, respecting the events of former age'. When addressing the Geological Society Adam Sedgwick told in 1831 that, 'the earth's surface ought to present an indefinite succession of similar phenomena, if Lyell's doctrine was true. In spite all such criticisms

Lyell continued to stick to his concept of uniformity in laws over time and lectured in many of the European countries to convince the geologist community.

Should we abandon Catastrophism?

Geologists who followed Hutton have been able to explain the Earth's features in a logical manner by using the Principle of uniformitarianism. But in so doing they have made an outstanding discovery—the Earth is incredibly old. It is clear that most processes of erosion are exceedingly slow. An enormously long time is needed to erode a mountain range down, for instance, or for huge quantities of sand and mud to be transported by streams, deposited in the ocean, then cemented into new rocks, and the new rocks deformed and uplifted to form a new mountain. Slow though it is the cycle of erosion, formation of new rock, uplift, and more erosion has been repeated many times during the Earth's long history.

Uniformitarianism is a powerful principle, but should we abandon catastrophism as a totally incorrect hypothesis? Recent discoveries of thin but very unusual rock layers at many places around the world suggest that random, catastrophic events did cause massive changes in the geological record. These were not the catastrophes perceived by seventeenth-century biblical scholars, but they do seem to have caused sudden massive changes in the earth's external appearance.

The kind of catastrophe suggested by the unusual rock layers that have led geologists to think again about the validity of catastrophism is the sudden impact of a huge meteorite somewhere on the Earth. Many meteorites have hit the Earth during human history, but fortunately there have not been any really big impacts. The peculiar rock layers mentioned above are rich in the uncommon metal iridium, which is much more abundant in meteorites than in the Earth's common rocks. The unusual iridium-rich rocks have been discovered in Italy, Denmark, and other places around the world. They suggest that a catastrophic impact did occur about 66 million years ago, and that many forms of life, including the dinosaurs, became extinct as its result.

A fascinating but frightening suggestion has been made that a disaster of a different kind may already be happening. The suggestion is that our collective human activities may be changing the Earth so rapidly and so massively, that we may be causing a change similar in magnitude.

□ 1.1.22 SUMMARY

Uniformitarianism, in geology, is a doctrine holding that changes in the earth's surface that occurred in past geologic time are referable to the same causes as changes now being produced upon the earth's surface. This doctrine, the basic concept of which was first advanced by the Scottish geologist James Hutton in his *Theory of the Earth* (1785, 1795), was further expounded by another Scotsman, John Playfair, in his *Illustrations of the Huttonian Theory* (1802). The conclusion that rates of geological processes have differed in the past from today's rates is an important one. It means that the relative importance of different geological processes have probably differed in the past.

The Principle of Uniformitarianism provides a first and very significant step in understanding the Earth's history. We can examine any rock however old, and compare its characteristics with that of similar rocks forming today in a particular environment. We can then infer that the old rock very likely formed in the same sort of environment. For example, in many deserts today we can see gigantic sand dunes formed from sand grains transported by the wind. Because of the way they form, the dunes have a

distinctive internal structure. Using the Principle of Uniformitarianism, we can safely infer that any rock composed of cemented grains of sand and having the same distinctive internal structure as modern dunes is the remains of an ancient dune.

Today, however, most if not all mainstream scientists support uniformitarianism as do most mainstream religious denominations.

□ 1.1.23 SELF ASSESSMENT QUESTIONS

1. Bring out the planetary evolution of the Earth according to our modern understanding.
2. How did the constituents of present day Earth evolve from their origin at "Big Bang"?
3. Trace out the stages of formation of the stars in the universe along with their fates.
4. Enumerate the principles of the methods employed in exploring Earth's internal structure.
5. Explain the nature of propagation of the different reflected and refracted phases of seismic waves through the earth's interior.
6. Present a generalized model of the layered structure of Earth's interior, as revealed by the studies of earthquake waves, highlighting the physico-chemical characteristics of each layer.
7. Critically examine the models put forward to address gravity anomalies.
8. How does isostatic adjustment induce vertical movements of the Earth's crust?
9. How does the Principles of Uniformitarianism help us to understand the history of a landscape?
10. Critically evaluate the field observations which emerged as apple of discord between the catastrophists and uniformitarians. How could the debate be negotiated?

□ 1.1.24 References :

Duff, P.M.D (1996): *Holme's Principles of Physical Geology*; Chapman & Hall, London, UK.

Hawking, S. (1988): *A Brief History of Time- From the Big Bang to Black Holes*; Bantam Press, London, UK.

Kearey, P. and Vine, J.F.(1990): *Global Tectonics*; Blackwell Science, Oxford, UK.

Lowrie, W. (1997): *Fundamentals of Geophysics*; Cambridge University Press, Cambridge, UK.

Merritts, J.D., Wet, A. and Menking, K. (1998): *Environmental Geology- An Earth System Science Approach*; W.H. Freeman and Company, NW, USA.

Patwardhan, A.M. (1999): *The Dynamic Earth system*; Prentice Hall of India Pvt. Ltd., New Delhi, India.

Siddhartha, K.(1999): *The Earth's Dynamic Surface*; Kisalaya Publications Pvt. Ltd., Delhi.

White, J.F.(Edt.) (1968): *Study of the Earth- Readings in Geological Science*; Prentice Hall of India Ltd., New Delhi.

PART-I
Paper - I, Module - I
GEO-TECTONICS & GEOMORPHOLOGY
Unit – 02

MODULE STRUCTURE

- 1.2.0 : Introduction
- 1.2.1 : Objective
- 1.2.2 : Plate Tectonics and Neo-Tectonics
- 1.2.3 : Plate Tectonics and Earthquakes
- 1.2.4 : Plate Tectonics and Volcanisms
- 1.2.5 : Plate Tectonics and Orogeny
- 1.2.6 : Self Assessment Questions
- 1.2.7 : References

Paper - I, Module - I
GEO-TECTONICS & GEOMORPHOLOGY
Unit – 02

□ **1.2.0 INTRODUCTION :**

Deformation of earth features by endogenetic forces is represented by the Unit of the paper. Plate tectonic theory is a revolutionary concept when it was proposed in the 1960s. Major features of the Earth's crust and interpretation of Earth's history are discussed with this part of the paper.

The structure of the Unit - 02 under Module-I includes as :

1. Plate Tectonics and Neo-tectonics
2. Plate Tectonics and Earthquakes
3. Plate Tectonics and Volcanism
4. Plate Tectonics and Orogeny

□ **1.2.1 OBJECTIVES :**

- To discuss the importance of Plate Tectonics theory for the explanation of earth's major features
- To understand about the tectonics as process in geomorphology
- To differentiate the mechanisms of plate tectonics and neo-tectonics

□ **1.2.2 PLATE TECTONICS AND NEO-TECTONICS :**

A large part of the continental blocks shows indubitable signs of recent crustal instability, and the young mountains like the Himalaya are visibly afflicted with restlessness. This is manifest in recurrent seismicity, movement of rock masses along faults and thrust planes, subsidence and rise of the ground, structural deformation and evolution of certain peculiar geomorphic features. The imperceptibly slow and secular as well as episodic crustal movements that have been taking place since the beginning of the Quaternary period some two million years ago are described as neotectonic activities. Areas involved in or subjected to the impact of continental drift and collision are particularly prone to persistent tectonic instability and attendant deformation and displacement. The Indian subcontinent is one such area of present-day geodynamic movements.

Δ **Indicators of Neo-tectonic Movements :** The significant and recognizable indicators of neotectonism, as pointed out by Dennis (1972), may broadly be categorized as follows :

- i) Historical instances of sunken or buried archaeological sites, including harbours and seaports.
- ii) Geomorphic features such as changing courses of rivers and streams, abrupt drainage swings, raised

beaches in coastal tracts, development of depositional river terraces and erosional terraced slopes in mountain belts.

- iii) Structural dislocation and deformation of subrecent and recent sedimentary deposits, including the dissection of talus cones, alluvial fans, warping or tilting of recent layered deposits, movement of ground along newly developed or reactivated older faults and thrusts.
- iv) Altimetric variation of elevation determined by geodetic measurements through precise triangulation methods or by laser-beam techniques.
- v) Recurrent seismicity, implying surface or underground episodic or spasmodic movements of rocks along faults or thrusts.

Keeping the Indian subcontinent in focus, various manifestations of present-day crustal movements are discussed in the following paragraphs.

Δ Historical Instances : Towards the close of the Mahabharat period (~3,000 years ago) the area surrounding the golden city of Dwarka on the Saurashtra coast was engulfed by coastal marshes and bogs. And the aged hero of the celebrated war, Lord Krishna, helplessly witnessed the fratricidal fighting among his Yadava men, pulling the reeds out of the marsh to exterminate each other. Soon after Krishna left for his heavenly abode, Dwarka was submerged under marine waters. This submergence indicates that the Dwarka ground had subsided. Recent archaeological finds, including painted pottery, indicate four successive occupations of the temple town between 1400 B.C. and the second century B.C. Across the gulf, the island of Bet Dwarka was likewise submerged and rebuilt, as indicated by the Late Harappan (1900-1600 B.C.) remains.

On the other side of the continent along the eastern coast of Tamilnadu and Andhra, there were famous seaports and towns of intrepid seafarers of the Chola, Chalukya and Rashtrakuta times, such as Mahabalipuram, Kaveripatnam, Korkai and Koringa (at the mouth of the Godavari). The very recently submerged port of Kaveripatnam was discovered 30 m below the surface of water, revealing remnants of the Chola kingdom and Buddhist learning centres. These places now lie several kilometres inland from the sea, obviously as a result of the emergence of coastal land. But then in the nearby Kalinokham bay on the Tirunelveli coast there is a submerged forest, which indicates sinking of the ground.

Another classical example of subsidence and uplift of the coastal tract is provided by columns ("temples") of Serapis at Pozzuoli near Naples in Italy. The floor of the market place was lowered 5 m below sea level, as indicated by the boring on the columns by intertidal organisms; and then it was elevated, but not quite upto the former level, so that the lower parts of the columns remain submerged.

Δ Rising and Subsiding Grounds : Geodetic Surveys : Isostasy is primarily responsible for imperceptibly slow and secular rise and subsidence of the ground, such as witnessed in the Fennoscandinavian region. There, the old strand lines are found located hundreds of metres above the present sea level. The uplift is due to the exponential movement decreasing with time since the period when deglaciation started. The central part of the

region has registered the highest rate of Uplift (8 cm/1000 years), while the outer part is rising slowly at the rate of 2 cm/1000 years. It is believed that the enormous load of ice during the Late Pleistocene glaciation had depressed the floor maximally in the central part. After the melting away of the ice, the ground is slowly regaining its original level. Coming to India, the Meghalaya plateau has been steadily growing in height at a rate of 3-4 cm/100 years as shown by geodetic measurements carried out between the years 1910 and 1977 (Chugh, 1973). In contrast, the adjacent Sylhet plain to the south has sunk 10 to 15 m in the last hundred years, so that the average altitude of the plain is a mere 3-4 m above sea level compared to 16-18 m elsewhere (Morgan and McIntyre, 1959). The sunken basin is characterized by a large number of lakes, marshes and buried forests. The rise of the Meghalaya plateau and the sinking of the Sylhet basin may be attributed to vertical movements along the faults parallel to the Dauki Fault delineating the steep scarp.

The Ganga basin being an area of brisk sedimentation, the growing load is causing progressive sagging of the floor of the basin, particularly adjacent to the foothills which is the belt of maximum accumulation. This is borne out from a depth of more than 6,000 m of the floor in the Sharada and Gandak depressions. Many of the archaeological sites, including Lord Ram's Ayodhya of the Ramayan period, bear testimony to the subsidence of the ground. One of the reasons why the Ghaghara, Gandak, Narayani and Kosi rivers become wild and wayward during floods, is that the beds of their channels and the adjacent flood plains are at the same level, or not much higher than that of the Ganga which they join in the southern tract. As a result, the excess waters of the tributaries are not easily drained out.

To the south of the Ganga, the Chhotanagpur plateau in southern Bihar exhibits prominent geomorphic features such as waterfalls, river terraces and mild seismicity, indicating gentle uplift of the plateau. In Keonjhar-Mayurbhanj belt, five flat levels at altitudes 130-160 m, 230-300 m, 440-500 m, 600-700 m and 940-1000 m are discernible (Ramana Rao and Vaidyanathan, 1970). These levels imply five pulses (at the rate 1 cm/1000 years) of episodic uplift presumably along the Singhbhum zone discernible between Seraikela and Sini. In the Singhbhum 'shear zone' the raised gravel beds, deeply incised terraces and other geomorphic features indicate the movements have continued along this zone in recent times and are possibly taking place even today (Banerjee, 1981).

Geodetic survey carried out in the Dehradun area showed that compared to the plains the Siwalik is rising at a rate of 0.8 mm/year (Chugh, 1973). An analysis of mineral ages by Mehta (1980) suggests that the granitic terrain of the Lesser Himalaya has been rising at an average rate of 0.77 mm/year to 0.55 mm/year in central Himachal and 0.8 mm/year in central Nepal. This faster rate of uplift is manifest also in the quickened pace of erosion in the Himalayan domain.

High up in the Himalayan domain, the Karewa sediments on the Pir Panjal sides of the Kashmir valley have been uplifted and mildly deformed in subrecent times. Kusumgar's (1981) palaeomagnetic studies demonstrate that the 5000 m thick lacustrine and glaciolacustrine sediments span the Brunhes (0.72 m.y.), Matuyama (0.72-0.47 m.y.), and Gauss (2.47-3.41 m.y.) palaeomagnetic epochs, and the radiocarbon dating indicates that the topsoil in the loessic sequence is $18,000 \pm 1000$ years old.

Δ Movements along faults : The evolution of sharp scarps, triangular facets on spurs, dissection of debris-cones and alluvial fans, and warping and tilting of layered sediments, are some of the expressions of movement along the faults plane. In the Kosi valley in Kumaun Himalay, reactivation of an important fault has caused dissection of subrecent landslide fan. The best documented example of the present-day movement along a fault is that of the San Andreas Fault of California, U.S.A. In some sections the right-lateral strike-slip horizontal movement is of the order of 10 cm/year, but the average rate over the period of a million years is 5 cm/year. The eastern block has moved about 550 km southeast with respect to the western block.

In the Indian subcontinent a fault of comparable geodynamic significance and movement is the Chaman Fault. It extends 900 km north from near the Makran coast in Baluchistan to southern Afghanistan. For more of its length, it separates greatly contrasted geological terrains, and is interpreted as the western transform boundary of the Indian plate (Lawrence and Yeats, 1979; Lawrence et. al., 1981). Its remarkably straight northern part dissects and uplifts (to the west) many alluvial fans, offsets streams and is the locus of innumerable springs. During the 1892 earthquake, a railway line was offset by 75 cm. The maximum offset discernible in Afghanistan is of the order of 200 km – a cumulative displacement over a period of several million years.

The development of new faults and the reactivation of older ones lead to displacement of the ground. This type of movement is commonly episodic, often violent. The 1819 earthquake of Kachchh, for example, gave rise to an E-W trending fault registering a vertical displacement of 5-8 m. the Okha Rann Fault is marked by a 20-40 m high cliff of the Gaj (Miocene) and Dwarka (Pleistocene) beds, the latter flexing into ENE-WSW trending gentle folds. In the southern coast of Saurashtra on the two sides of the Hiran valley, the Pleistocene meliolites occur at heights of 75 m (at Umreli) and 104 m (at Shingavada), indicating differential movement along an active fault that follows the river bed (Merh, 1980).

The South Indian Shield in its southeastern part, particularly in the Tirupattur-Mattur area, is dissected by a number of deep and long NE-SE trending faults. Some of these faults reach down to the upper mantle, and many exhibit mild seismicity ($M < 5$). Raised gravel beds and waterfalls across the fault near Hoganaikal indicate subrecent tectonic activity (Kailasam, 1979). The Eastern Ghat is depicted (from gravity data) as a horst flanked to the east by the subsiding basin in which sedimentation has proceeded since the Upper Cretaceous times.

The Nilgiri-Palni-Cardamon mountain range, which rises abruptly to a great height (2500 m) amidst deeply peneplaned terrain, is a result of block uplifting in the Late Pleistocene (Wadia, 1942) along faults that cut the land in an NE-SW direction. The spectacular waterfalls in the aged rivers such as Sharavathi, corroborate this deduction. Possibly the present elevation was achieved as a consequence of repeated uplift (Radhakrishna, 1968).

In the Narmada valley between Jabalpur and Hoshangabad the Quaternary sediments have been displaced to a level below the mean sea level, and the transverse Navagam Dam Fault has shifted the Bagh Beds by 250 m (Krishnaswamy, 1981). Small hanging valleys with short waterfalls in some of the southern tributaries of the Choral River are suggestive of progressive E-W step faulting with downthrows parallel to the narmada (Subramanyam, 1981). Likewise, in the Tapi valley the Ellichipur Fault has thrown down the Quaternary deposit by 600 m.

As already pointed out, the Meghalaya plateau is an example of continuing uplift (at a rate of 4-5 cm/year) along faults paralleling the Dauki wrench fault. Seismicity of the order $M > 6$ and the Bouguer gravity anomaly (+40 to -50 mgal compared to -180 to -250 mgal of the Brahmaputra plain to the north) show this physiographic eminence as a horts that has been rising recent times.

To the south in the flat green expanse of Bangladesh, there is an elevated, forested and dendritically dissected eastward-sloping tract comprising the Madhupur Jungles, made up of the Pleistocene sediments and demarcated against the recent flood plain by a series of fault scarps, 5 to 20 m high (Morgan and McIntyre 1959). These scarps are expressions of sub-recent movements along minor surficial en echelon faults related to a major NNW-SSE trending subsurface fracture developed on the floor of the basin that extends up to the Himalayan foothills (Sikkim-Darjeeling) and is primarily responsible for the high seismicity of the region (Valdiya, 1976, 1981). The dramatic changes of the courses of rivers and the subsidence and elevation of the ground are but manifestations of neotectonic activity of the subsurface Madhupur Fault.

Δ Movements along Thrust Planes : Examples of older Himalayan rocks advancing and riding over sub-recent and recent deposits are numerous (Valdiya, 1981). In the Lohit valley in the extreme northeast, the Palaeozoic (?) Mishmi metamorphics are thrust over the fluvial gravel and sands. In the Darjeeling Hills the Precambrian crystallines are seen resting upon the sub-recent alluvial deposits of the Tista River. Near Khewra in eastern Nepal the Middle Siwalik covers the recent alluvium. In the Ramganga valley in Kumaun, the Sarpadhuli-Dhikala Thrust has pushed the Siwalik over the recent scree and alluvial sands. Between Rishikesh and Dehradun the MB (Krol) T has caused the thrusting (by 40-75 m) of the Lower Paleozoic Chandpur phyllites over the Pleistocene Dun gravels. In the Tons valley near Chhibro the Eocene rocks have been displaced over the recent scree mantling the Siwalik. In the Hoshiarpur district the Punjab the Soan Thrust has brought the Siwalik rocks over the recent alluvial sediments. In the Vaishnodevi area in the northwest recent talus deposits lie crushed under the Precambrian Jammu Limestone. These are but a few of the multitude of instances of recent movement along the outer Himalayan thrusts.

In the Mussoorie Hills the MBT has dismembered and elevated by 150-200 m the subrecent Dun fan (Nossin, 1971). In the Nainital Hills a historically old stream draining the Manora slope in the Lesser Himalayan belt and joining the Nalena stream of the Siwalik terrain has been forced to abandon its channel as a result of the recent movement of the order of 30 m as seen near Jeolikot-Beluwakhan. The two flanks of paired terraces near Dogaon on the Balia River, southeast of Jeolikot, bear evidence of differential movement of about 90 m along the MBT that follows the bed of the river. It is obvious that the Siwalik has been uplifted compared to the Lesser Himalay and the movement is of a rotational type. Movement is taking place not only on the MBT but also on quite a few of the sympathetic faults. The Manora T, for example, has placed the Upper Palaeozoic Lower Krol upon the recent scree near the Hanumangarhi temple at Nainital.

The MBT zone shows conspicuous bends or offsets of rivers and streams of the order of 15 to 25 km. This suggests lateral movement along the thrust plane. This has been proved by actual measurement done in the Tons Valley. Ansari *et al* (1976) have shown that the country south of the thrust is moving in a southwest direction

compared to the northern block, the average southward component being 0.64 cm/year and the westward component 0.27 cm/year. The Nahan Thrust, which has developed parallel to the MBT in the Tons Valley, registered movement at the rate of 0.92 cm/year during the short period of observation (Sinvhal *et al.*, 1973). The Shali Thrust (MBT) in central Himachal Pradesh is likewise very active.

Δ Recurrent Seismicity : The occurrence of an earthquake means displacement of a body of rocks along a fault plane, whether superficial or underground. Seismicity of varied degree as that of the fault-ridden southwestern South Indian block, of the Koyna-Bahruch-Kachchh belt in western India and of the Himalayan belt in the north is a manifestation of the activeness of faults and thrusts. A majority of seismologists attribute the high seismicity of the Himalayan belt to movement along the MCT and MBT, and it has been suggested that the seismicity has progressively moved southward, so that at present the MBT is more active than the MCT. This implies that the future will witness reactivation of the now concealed Himalayan Frontal Fault demarcating the Siwalik from the Ganga Plains. However, in the author's opinion, it is the transverse faults which are more active than the MCT or MBT, and are indeed responsible for high seismicity, as would be quite obvious from the quantitative seismicity map of the Himalaya (Valdiya, 1981).

Assam has been frequently, sometimes violently, rocked by earthquakes. Epicentres of at least 40 tremors of magnitude 5 to 8.5 are located in the Brahmaputra valley. The 1950 earthquake raised the river bed by 3 m at Dibrugarh, and the area to the north of Dangori and Sai Khoa subsided and was converted into a marsh.

□ 1.2.3 PLATE TECTONICS AND EARTHQUAKES :

Plate tectonics replaced the older ideas of continental drift and sea-floor spreading when it was realized that typical plates include both continental and oceanic crust. Moreover it was realized that the sial-sima boundary, the Moho, was not a suitable surface for continents to move on, and that movement took place in the low-velocity layer, the asthenosphere. The lithosphere with its appreciable strength and rigidity includes both sial and sima, continent and sea floor so 'plates' which may likewise contain both components convey a better image of the process than continental drift or sea-floor spreading which emphasize one or the other. Plate tectonics is a name for the general idea that the earth's crust can be divided into several plates and the world's main tectonic features are related to activity at the edges of the plates. It is supposed that new crust is created at spreading sites, and that the crust is destroyed at subduction zones. Beyond this simple exposition many complications are possible, and variations from writer to writer, and from place to place are almost numberless. For the past decade plate tectonics has been the ruling theory in geology, and has been used to account for almost everything in earth science.

Firstly let us look at some distribution maps of major earth phenomena.

These follow two main sets of lines. One set, all shallow earthquakes, follows the line of the sub-oceanic ridges, the spreading sites. Another set, including most of the deep earthquakes, follows the line of subduction sites around the Pacific rim and lines of deep sea trenches. A few earthquakes follow the great fracture zones, and others the rift valleys which might be incipient spreading sites. A residue of other earthquake remains unexplained

by this hypothesis, but clearly the distribution of earthquakes, by and large, divides the earth's surface into a number of areas that correspond very well with the plates of plate tectonics. Indeed plate tectonics provides a very useful simplification in accounting for earthquakes. Earthquakes occur where the action takes place – at spreading sites and subduction zones at the borders of plates.

Spreading sites are the simplest. Here sub-ocean rises are presumably related in some way to the rise of magma. On these sites there are many volcanic eruptions, vertical dykes are intruded in great numbers (parallel to the rift), and submarine lava flows and hills are formed. Earthquakes occur in the top few kilometres of the crust, and show essentially tensional and vertical movement. The geomorphology of the rift area can be seen in fair approximation in Iceland, which may be regarded as a piece of the mid-Atlantic Ridge above sea level. The rocks get older away from the rift, the rocks are dominated by fissure eruptions, and the topography, where not primarily volcanic, is marked by many faults and fissures parallel to the main mid-ocean (or mid-Iceland) rift.

It would seem that the volcanic intrusion is permissive. That is, dyke rocks come in to fill a space created by tension. We should not imagine a thick dyke forcing its way into a crack and pushing aside a plate many thousands of kilometres across. Rather there is some force pulling the plates apart and the dykes come in to fill the space created when a crack appears.

Some Atlantic islands are on or close to the mid-ocean ridges. The active Ascension Island is virtually on the ridge. St Helena 700 km from the ridge and extinct. A very rough correlation of the age of oceanic islands with distance from the ridge led Tuzo Wilson to propose that the Atlantic islands were formed on the rift and then drifted away with the spreading of the ocean. This idea still appears roughly true, but is slightly complicated by the fact that the islands remain active for several million years and for a considerable distance from the rift. Tristan da Cunha is still active (the last eruption was in 1961), although it is 350 km from the mid-Atlantic rift.

Δ **Collision sites** : Collision zones provide more varied tectonic settings than spreading sites, including the following types :

Δ **Continental-ocean collision with subduction** : This is the commonest sort of collision envisaged in the plate tectonic theory, but can itself be divided into two :

Δ **Simple subduction (Andes type)** : This is typified by the plate junction on the west of the southern half of South America. The advancing continent overrides the sea floor, which is thrust down under the continental plate. The down going slab melts to produce andesitic magma which is intruded and erupted as andesitic volcanoes. The course of the slab's progress is recorded by a series of earthquakes that show a fairly consistent increase in depth with distance from the plate edge making an inclined earthquake zone known as a Benioff zone after its discoverer. The continental edge is uplifted to form mountains (The Andes) and the collision zone is marked by a deep ocean trench (Dewey and Bird, 1970). These ocean trenches are the deepest part of the ocean, often over 10,000 m.

Another possibility is that a subducted plate may flatten, causing perhaps a cessation of volcanism and renewed

uplift. This has been proposed for the subduction of the Nazca Plate beneath Peru by Barazangi and Isacks (1976) and elaborated by Noble and McKee (1977).

Δ **Island-arc subduction** : Around the western edge of the Pacific the map shows many islands arranged in festoons of arcs, either single lines or double rows of islands, making a quite clear pattern of curves. These line some distance off the continent and in front they have a deep trench very like that off the coast of South America. These deeps also bound a Benioff zone dipping towards the continent and marked by earthquakes reaching depths of several hundred kilometres.

Some arcs face east (Indonesian arc), some arcs have no continent behind them (South Sandwich Islands), and some lines of islands have all the attributes of arcs such as Benioff zones and trenches but lack the arcuate shape (Solomon Islands). Nevertheless the island-arc complex makes one of the distinctive landform and tectonic assemblages of the earth and despite many complexities and.

Δ **Continent-ocean collision with obduction** : In a few collision zones a slab of ocean floor has overridden rather than underridden the continent. In Papua New Guinea a slab of basic rocks with petrology, layering and structures exactly like those thought to be typical of ocean crust makes the mountainous terrain of the Papuan Ultramafic Belt. Another area is the Troodos Mountains of Cyprus where dyke-ridden oceanic rocks override continental rocks. Obduction is the term for this overriding, in contrast to subduction when the oceanic slab goes down below the continental mass.

Δ **Continent-continent collisions (Himalayan type)** : The most spectacular example of this collision types is that presumed to have occurred between India and continental Asia. Geological data show that India was once united with the southern continents of Gondwanaland, and until the Mesozoic shared their faunas, their ice-age and other features. Since then it has separated and drifted north at a rather high speed (16 cm/y) until it collided with the Asian mainland. This geological information is confirmed by the evidence of sea-floor spreading in the Indian Ocean. For a while part of a great sea, the Tethys Sea of Mesozoic and Early Tertiary times, lay between the two continental slabs, but this was crushed and uplifted as the Indian plate drove under the Asian plate. Further effects of this collision are the formation of a double thickness of continental crust. Isostatic compensation of this mass led to the formation of the Tibetan plateau, the highest in the world, fronted by the highest mountains in the world, the Himalayas. The double thickness of crust may also account for the lack of volcanoes on this plate boundary for it is too thick for magmas to penetrate. The zone is, however, marked by many earthquakes.

In some situations a continent-continent collision may continue as a continent-ocean collision. The Zagros-Makran line is regarded as a plate boundary of convergence by some (Farhudi and Karig, 1977) but there is a marked change of character from the Zagros thrust, which is a continent-continent collision, to the Makros region where oceanic crust is thought to be subducted. From south to north the Makran ranges become higher, older and more deformed. Further north still are subsiding basins and beyond them active volcanoes.

Δ **Ocean-ocean collision** : An arc may be formed, as in the South Sandwich islands (Scotia arc) where a

strip of spreading Pacific seems to be pushing a finger of oceanic crust into the oceanic crust of the Atlantic. The contact is marked by a trench, earthquake zone, active volcanicity and two sets of sea-floor spreading. The Caribbean arc is some what similar, with the Pacific pushing into the Atlantic.

The curvature of Indonesian arc may result partly from the Pacific Ocean spreading into the Indian Ocean.

Δ **Arc-arc collision** : Yet another possibility is that two arcs collide, and a possible example is provided from the Philippines (Roeder, 1977). The chief features of the region are the Manila Trench and the Philippine Trench, with convergent Benioff zones, and the collision zone is interpreted by Roeder.

□ **1.2.4 PLATE TECTONICS AND VOLCANISMS :**

Volcanoes are distributed along lines of sea-floor spreading, and on the areas bordering subduction zones. The type of volcanicity is even more diagnostic, with essentially basaltic volcanicity on spreading sites, and andesitic volcanicity at subduction sites. Alkaline volcanic activity follows some rift valleys. As with earthquake activity the plate-tectonic theory provides a satisfactory explanation for the location and type of much volcanic activity.

□ **Volcanoes and magma** : Volcanic activity produces a distinctive set of landforms built from the products of molten rock or lava that erupts from beneath the earth's crust. Magma, which differs slightly from simple molten rock because of its dissolved gases or volatiles, is derived from somewhere near the base of the crust. It is thought that the sub-crustal material which is normally solid becomes liquefied for some reason such as sudden reduction of pressure and rises as magma through the crust by way of fissures or pipes until it reaches the ground surface as lava where it erupts in some form of volcanic activity.

In the ocean basins the dominant magma is of basaltic composition. In chemical composition this is almost identical with the mantle, and could be derived from it by melting. In the Pacific an area of entirely basaltic volcanoes is separated from those areas around the edge of the Pacific by an 'andesite line', on the shoreward side of which andesite magmas are commonly erupted, though basalt magmas are also erupted. Andesitic magmas have more silica than basalt magmas and usually erupt with greater violence. Even more silica-rich (more 'acid') magmas are known, with the approximate composition of granite, and these give rise to volcanoes of rhyolitic composition or to ignimbrite pyroclastics. Small amounts of andesitic and rhyolitic magmas may be produced by various differentiation processes from basaltic magma, but the production of large volumes of acid magmas probably requires the melting of sialic continental rock in some way, and these rocks are confined to continents and their margins.

The fact that andesites are absent from the Pacific Ocean within the 'andesite line' in the Pacific at least andesites are not created by fractionation of basalt magmas, and there seems little reason to suppose they are elsewhere. Plate motion rates may be relevant to the genesis of magmas and Smith (1976) says that slip rates of about 10 cm/y (100 000 B) are commonly regarded as typical values needed to generate andesites. Liquid lava flows away from its point of eruption of distances that depend on local topography and the viscosity of the lava. Acid lavas are the most viscous and never flow freely : basalt lavas flow more readily and can form very long thin flows. The surface of a lava flow has distinctive features which have been given names. Aa lava is blocky and

slow moving, with a thick skin, broken into blocks that ride on the massive, pasty lava underneath. Pahoehoe flows are the most liquid type. Cooling forms a thin skin that may be dragged into folds by movement of the still fluid lava beneath to form ropy lava.

Volcanic eruptions that contain large amounts of gas are explosive and produce pyroclastic deposits, fragments of volcanic rock ranging from dust size to huge blocks. Lavas may be innately explosive, as are most very acid lavas. Other pyroclastics are produced when lava hits the sea or groundwater, and production of scoria seems to be common in the last stage of some basaltic eruptions, presumably when the supply of lava is running down but gas is still abundant.

Some other volcanic products indicate the environment in which they were erupted. Pillow lavas, as the names implies, look like a stack of pillows, each with a glassy edge and more crystalline interior. They are a sure sign of eruption in water. In some places, such as Sicily, the level of pillow lava can be used to determine old sea levels. Elsewhere it is clear that the pillow lavas were deposited in the deep sea. These are often rich in sodium and are commonly associated with chert and serpentine – the so-called Steinmann trinity. Pyroclastic rocks produced by eruption in shallow water (called hyaloclastics) may be found associated with some pillow lavas and volcanic breccias in volcanic deposits resulting from an eruption into an ice cap. It is the recognition of such an origin for datable volcanics in Antarctica that gives the age of the appearance of the Antarctic ice sheet.

Δ **Basalt volcanics :** Basaltic magmas are generally intruded through cracks which open through tension and become filled with magma to form dykes. These not only provide the feeder channels for surface volcanoes but they themselves add considerably to the bulk of the crust they traverse. Iceland has been spreading at a rate of 5 cm/y by dyke intrusion. In Arran, Scotland, Tertiary volcanic intrusion accounts for an extension of the crust of 1 in every 15 kilometres.

Crustal spreading on Iceland takes place along a series of rift zones, broken by several transverse fracture zones. Each lava pile appears to have built up as a series of lenticular lava units, each representing a single eruptive episode and related to its own feeder dyke swarm. Iceland has thick crust and deep seismic activity. Some workers postulate a primordial mantle plume magma source beneath Iceland.

At the ground surface eruptions from dykes give rise to fissure eruptions. These commonly erupt liquid lava which flows readily, covering wide areas and modifying the surface topography by flows, but not building greater volcanic mountains, though small scoria cones may grow along the line in the last stages of activity.

—Towards the surfaces many dykes break up, and the ascending sheet of magma turns into a series of distinct pipes. Each pipe gives rise to a volcano when it reaches the surface.

Some volcanic activity is dominantly around a single vent in contrast to that along fissures. These are the central volcanoes and include the low-angle shields of Hawaii, and steep oceanic islands like Tristan da Cunha. Activity continues at the one centre for periods of thousands or even millions of years. Sometimes lava is produced, sometimes volcanic ash and other fragmental pyroclastic deposits, but flows predominate in most basaltic volcanicity.

Scoria volcanoes tend to be late-stage eruptions and rather small.

Areal volcanism (also known as polyorifice volcanism) is characterized by monogenetic volcanoes and the absence of any tendency for eruption centres to be localized at definite points for any length of time. Individual volcanoes are short lived, ranging from a few weeks to perhaps 12 years, and seldom grow bigger than 450 m. Scoria cones, lava cones and maars are the dominant volcanic types and strato-volcanoes are absent or rare. The spatial distribution of the volcanoes is irregular, with some clustering and occasional linear groups. The petrographic composition is basaltic and remains fairly constant with an areal province. Examples of areal volcanic provinces include Victoria (Australia) and the volcanic regions of Auvergne, Armenia and Mexico.

The type of volcano, the nature of the magma and the nature of the eruption are very much affected by the tectonic setting. Mid-ocean volcanicity is largely of fissure type but may be confined at times to a few central volcanoes. The island of St Helena is now about 250 km east of the ridge, and probably had an origin like Ascension but it is now extinct and has been carried away from the active site by sea-floor spreading. It was, however, active from 15 to 8 m.y. ago, though not continuously.

Other volcanoes of the oceans appear to be related not to sub-ocean ridges but to hot spots under the crust. The hot spot periodically erupts and forms a volcano. After a period of quiescence the hot spot is reactivated, but meanwhile the crust has drifted over the hot spot and so the newer volcano is in a different position relative to the crust.

Hot spots can also lead to eruption on continents. Wellman and McDougall (1974) plotted the age of central volcanoes of eastern Australia against latitude and found a remarkable correlation. They suggest that the Australia continent has moved in a northern direction over two hot spots. The apparent southward migration rate of the centres of volcanism in eastern Australia 665 mm/1 000 y, then becomes the rate of northerly movement of the Australian plate over the asthenosphere.

Basaltic lavas are generally the most fluid, so they can flow for great distances, like the 120 km Frambuni flow from Trollaungja in Iceland. On level ground lavas can form a volcanic plain of great extent with little volume, like the 150 000 km² volcanic plain of Western Victoria. The Laki flow of 1783 in Iceland covered an area of 565 km² in seven months, with an average thickness of about 25 m.

On irregular topography lava flows down valleys, sometimes filling them and spilling over interfluvies, and generally disrupting earlier drainage systems. Sometimes lava production is on such a vast scale that even the valleys of originally mountainous regions are completely filled and a lava plateau is produced. The Columbia River Plateau, for instance, which is of Miocene age, completely buried an original topography with a relief of over 1 500 m – deeper than the Grand Canyon. The Columbia Plateau basalts average 1 000 m thick, made up of many individual flows averaging 10 m thick and they cover an area of about 130 000 km².

Rather similar, and frequently treated together with the Columbia River basalts are the Snake River basalts of Quaternary age covering 50 000 km². The Deccan traps of India are of Cretaceous to Eocene age with a

present area of 500 000 km² and thickness up to 2 000 m. The Cretaceous Parana plateau basalts of Brazil cover an area of over 750 000 km² and the Karroo basalts of South Africa, which range from Late Triassic to Early Cretaceous, cover 140 000 km² but were once probably ten times as extensive Cox, (1972).

Such vast deposits could only be derived from the mantle, with minor additions from other sources, and must relate to very major tectonic events. Cox (1972) considered that the Karroo basalts were related to a single thermo-tectonic event related to the establishment of the continental margin of southeastern Africa during the fragmentation of Gondwanaland. The waning stages of this cycle overlap with the main period of activity in South West Africa which is presumed to be associated with the development of the South Atlantic coastal margin. It is postulated that a rise of mantle material approximately equidimensional in plan, took place beneath the central part of Gondwanaland, and caused the observed surface volcanic and tectonic effects.

Windley (1977) considered that the Jurassic lavas of Gondwanaland lie on the foreland of the Mesozoic fold belt, and represents an abortive attempt at fragmentation, while the Cretaceous lavas are located along or near actual separated continental margins.

Δ Intermediate and acid volcanics : A completely different setting for volcanoes is the Pacific margin of the Americas. Here andesitic volcanoes are dominant, and they occur along distinct lines or chains. Why the single chain of southern South America gives way to a double chain at latitude 30°S is not clear, but there appear to be definite lines at depth governing the location of these huge central volcanoes. It is significant that this area is also intruded by many batholiths of andesitic composition. The line continues into North America, but is less distinct. The Cascade Range of the United States is almost entirely volcanic. It is contemporaneous with the Columbia River basalts to the east, but whereas the latter are basaltic the Cascades are andesites and granodiorite intrusives. Later stages produced the many large composite cones such as Mt Rainier, Mount Hood, Lassen Peak and Mt Mazama.

On the western side of the Pacific island arcs predominate, and here again andesitic volcanoes are present, similar in most respect to those of the Americas. One interesting feature discovered in island-arc volcanism is that the potassium content seems to vary systematically with distance from the front of the arc. Some have related this to depth of the subducted plate, but the correlation of potassium content with the assumed depth of the plate, or the depth of the Benioff zone as determined by earthquakes, does not actually provide an explanation for chemical variation.

Island arcs may have double bands of contrasting volcanism, with eruption of basic volcanic rocks, pillow lavas and breccias on the sea floor in the trench and intermediate to acid volcanics making andesitic volcanoes or ignimbrite deposits on the volcanic arc. The same relationship, if found in the past, suggests an arc situation, as in Wales where throughout the Ordovician acid volcanics appear as sub-aerial ignimbrites, while basic volcanics always occur as submarine pillow lavas and pillow-lava breccia.

A mixture of solid particles suspended in a gas can act as a liquid, a principle used in industrial 'fluidization'

for the transport of such materials as cement and coal dust through pipes. The process occurs naturally in some volcanic eruptions, when pyroclastics suspended in volcanic gases flow like extremely mobile liquid. They look like rapidly projected dust clouds, and at night they glow, and they are generally known as *nuees ardentes*, translated as 'glowing clouds'.

An eruption of this type deposits a layer of volcanic ash known as an ignimbrite (Cook, 1966) which is a rock unit term, not a petrological term. Rhyolite and andesite composition is commonest, and many ignimbrites are welded, that is gases are expelled and the hot ash sticks together to form welded tuff. Welded tuffs often look very like rhyolite, and many rocks that were identified as rhyolites in the past are now known to be ignimbrites. In Cotton's classic book of 1944 the ideas of ignimbrite had not been incorporated and he writes of rhyolite plateaus.

Pyroclastic flows are very mobile. They flow downhill and fill depressions rapidly, and when they come to rest they have a remarkably flat top. Flows have been traced for 70 km from Mt Mazama, Oregon. The Late Toba flows of Sumatra cover 25 000 km² and the Taupo-Rotorua ignimbrites have an area of 26 000 km².

Calculated volumes of ignimbrites are very great. The New Zealand field has a volume of 8 300 km³, the San Juan ignimbrites of California 9 500 km³. Individual calderas can produce large amounts such as 90 km³ from Aira calderas and 80 km³ of Aso both in Japan.

Δ **Ignimbrites erupt very quickly** : The Katmai eruption in Alaska produced 28 km³ in 60 hours and it is estimated that the 2 800 km³ of northern Queensland ignimbrites were erupted in a few days (Branch, 1966). On the other hand some ignimbrite areas had many pauses during construction, with sufficient time for forests to be established. In the Lamar River area of Yellowstone Park there are the remains of 27 distinct forests, with trees in an upright position, each buried in its turn by pyroclastics.

Δ **Global distribution patterns** : The distribution of volcanoes shows that the most volcanoes fit on plate boundaries, but there are also some such as the central Sahara mountains of Tibesti, Hoggar and Jebel Marra, where some volcanoes are over 3000 m high but located in the middle of a continental plate. In the oceans, the Hawaiian Islands are in the middle of a plate and not related to plate edges, though as we have already seen they are related to a 'hot spot'.

The global distribution of volcanic types has been studied by Suzuki (1977). He used the List of the World Active Volcanoes of Katsui (1971) which lists 830 volcanoes. Of these 699 were sufficiently well described to be classified into volcano type. Suzuki devised a volcano-morphologic formula defined as an arrangement of the abbreviations of simple volcanic landforms in order of their formation. He found 57 varieties which were then grouped into six volcano series according to the type of the oldest simple volcano (usually the main cone) and the existence of a caldera.

The six types and their global population percentages are shown below :

Strato volcano

Strato volcano with caldera	:	10
Shield volcano	:	11
Shield volcano with caldera	:	03
Caldera volcano	:	07
Monogenetic volcano	:	06

Suzuki further divided the world's volcanoes into five regional groups : island-arc volcanoes (I), Alpine zone volcanoes (A), Continental volcanoes (C), Oceanic volcanoes (O), and rift system and ridges volcanoes (R).

This correspondence may reflect the regional difference in degree of the explosivity of erupted magmas. To make this clear quantitatively the morphologic explosion index, defined as the population percentage of the pyroclastic-rich volcanoes (including strato volcanoes, stratovolcanoes with caldera and caldera volcanoes) in the total volcanoes except the monogenetic type was calculated for each regional volcano group.

The result suggests three major volcano groups based on the morphological explosion index :

Morphological explosion index (%)		
1	Island arcs	97
2	Alpine	
	Continental	82-77
3	Oceanic	
	Rift and ridge	41-38

The MEI may be compared with Rittmann's explosion index (1962) which is defined as the volume percentage of fragmentary material in total material produced and is calculated for the data between 1500 and 1914. Good agreement between the two suggests that within each regional volcano group the explosivity of erupted magmas remains unchanged for a long time – Suzuki suggests from the late Pleistocene to historic times.

Other results from Suzuki's analysis are :

1. Crater Lake (Krakatau) type calderas are about nine times as common as Kilauea-type calderas.
2. The percentage of volcanoes with parasitic or incidental volcanoes decreases in the order : caldera volcano; strato volcano with caldera; shield volcano; shield volcano with caldera; monogenetic volcano.
3. Lava domes rarely co-exist with other kinds of parasitic volcanoes.
4. It is expected that in the volcanoes to be formed in the future about half of the monogenetic volcanoes will be lava domes, while there will be hardly any lava domes amongst the strato volcanoes with caldera.

Aspects of the global distribution of volcanoes have been noted by many other workers, and some have

been pointed out in various parts of this book. Island-arc volcanicity, hot spots, and the association of carbonatites with continental volcanicity are some of the topics noted and a great deal of current work amongst petrologists is devoted to regional tectonic aspects of volcanicity. I have emphasized Suzuki's work here because it starts from morphology, while most others' work is based on theoretical petrology or tectonics.

□ 1.2.5 PLATE TECTONICS AND OROGENY :

Waterschoot van der Gracht (1931) introduced the concept of 'organic facies' into North America. Flysch was defined as 'sediments deposited previous to the major paroxysm' and molasses as sediments deposited 'during or immediately after the major diastrophism', emphasizing the early ideas of direct association of geosynclines, their sediments and orogeny. It also seems that the prevalent view of orogeny was of 'one big bang', and the idea of orogenic facies was introduced accordingly.

Petrographic classifications of geosynclines came later and certain rock types were optimistically related to parts of a geosyncline cycle, which also relates to a geomorphic and a tectonic cycle. Krynine (1941) divided geosyncline cycles into three stages :

- 1) Peneplanation (or nearly geosynclinal stage), with cyclic deposition on a fluctuating flat surface after much weathering, characterized by first-cycle quartzites.
- 2) Geosynclinal stage, with basin deposition interrupted by local vertical movement, with greywacke as the typical sediments.
- 3) Post-geosynclinal stage, with uplift after folding magmatic intrusion of the geosyncline and arkoses as the typical sediment.

The model for this sequence was the Appalachians, but in the Alps arkose is found in the 'flysch' of the geosynclinal stage and the post-orogenic molasses consists largely of rock fragments. There are too many factors affecting rock composition to allow its use as an indicator of geosynclinal or tectonic relationships. Later workers used sedimentary structures, especially cross bedding and graded bedding of turbidity currents as palaeogeographic indicators and related orogeny to geosynclinal fill through the interrelationship of sediment supply and rate of subsidence. These are not always directly related and some deep basins accumulate little sediment.

Some other rocks found in geosynclines have assumed importance in recent years, possibly relating geosynclines to plate tectonics.

Ophiolites are more or less altered basic rocks, presumed to be volcanic rocks erupted on to the sea floor at the bottom of a geosynclinal trough or perhaps parts of the sea floor itself. The fully-developed ophiolite sequence contains deep-sea sediment, pillow lavas, dykes, gabbros, peridotite and other basic and ultrabasic rocks. Since deep-sea ophiolites are hard to study considerable effort has been put into the study of ophiolites where access is easier on land.

Ophiolites are thought to be important because they may mark, in some way, the suture zone marking the

site of plate collision. Certainly many sutures bounding major tectonic units are marked by serpentine belts (metamorphosed ophiolite). The narrow zones of phiolites in the Alpine-Himalayan mountain belt are interpreted by some as slices of oceanic crust and mantle. They may mark the lines along which continents collided. Similarly, the Ural and Appalachian-Caledonian mountain belts, which lie within ancient Pangaea, have narrow phiolite zones, thought by some to mark the sites of vanished oceans. The Urals is thus interpreted first as an area of sea floor created by sea-floor spreading, consumed later when two continents collided, leaving only the ophiolite belt as a trace of the former sea floor.

Melanges are tectonically chaotic rock units, mega breccias of two types : tectonic *mélange* and sedimentary *mélange* or olistostrome. Both are found on the outer side of the trench arc gap and both are thought to be related to under thrusting tectonics at convergent plate boundaries.

Olistostromes are sedimentary stratigraphic units containing numerous exotic and local, mainly sedimentary, blocks in a fine-grained matrix. They are formed by gravity sliding, on a topographic slope such as the side of a geosyncline or trench. They are bedded mudslide deposits and have no direct indication of being subducted, though the Dunnage *mélange* of Newfoundland (Lower Palaeozoic) and the Gwna *mélange* of Anglesey, Wales (Late Precambrian) have both been attributed to subduction.

Tectonic *mélanges* are rock bodies composed of tectonically mixed blocks in a sheared shaly matrix. They may be very big, with individual blocks up to several hundred metres long. Rocks of very diverse origins may be included, such as cherts, glaucophane schist, greywacke and basalt.

A problem with these *mélanges* is that if a sedimentary olistostrome is sheared, it is difficult to distinguish it from a tectonic *mélange*. The definitions of *mélange* and olistostrome just given both depend on knowledge of assumed mechanism of deformation and often this is a mere assumption. Some prefer to use the term *mélange* in a purely descriptive manner, with no implication of the origin of the rock mass.

The tectonic significance of *mélanges* is that they might result from submarine slumping or they might be produced tectonically by scraping sediment and ocean floor from the top of a subducting plate. *Melanges* are often associated with pillow lavas, serpentinites and cherts and look like trench-bottom deposits. Some *mélanges* are associated with blueschists which indicate high-pressure and low-temperature metamorphism and could be formed in an environment such as that at the top of a subducted plate where pressure is great although the rock is so near the surface of the earth that temperature is still low. The association of blueschist with *mélange* supposedly indicates a former subduction zone.

Δ Folding in geosynclines : As seen at present the rocks of geosynclines are folded and some of them are currently mountainous regions. The idea arose that geosynclines and mountain formation were closely related and indeed in the pre-plate-tectonics day most ideas of mountain building were somehow related to geosynclines. At its simplest the idea was that a geosyncline existed as a sedimentary basin for several geological periods, but was finally compressed, thus folding the rocks and uplifting the mountains.

A number of features need further discussion. Firstly, the folding of the rocks. Geosynclines were not simply filled up and then folded for there are many periods of folding, erosion and further deposition within geosynclinal sequences. Folding was not a feature that occurred only at the end of the geosyncline's life, but frequently accompanied deposition. A squeeze at the end is not enough to make the observed folds.

The long axes of folds are of course roughly parallel to the length of the geosyncline, a result that can be obtained either by crustal compression or by sliding down the sides of the geosyncline. Study of the folds often reveals the direction of movement and it is found that some sediments moved towards the centre of the geosyncline. The term syntaphral tectonics was coined for this sort of folding and a symposium on this topic brought together many pieces of evidence for the widespread nature of gravity sliding in geosynclines (Carey, 1963b).

More often, however, the folds are away from the axis of the geosyncline. The impression is that the contents of the geosyncline were pushed towards a foreland, and the name foreland folding is given to this sort of structure. Early ideas suggested that the geosyncline is squashed, like the closing of a vice and the sediments forced out, as seen in the foreland folds. There was some concern because it seemed reasonable to find folds symmetrically arranged about the centre of the geosyncline, yet in reality fold belts are usually very asymmetrical.

The side of a geosyncline adjacent to a continent is usually well preserved and here we find that the sedimentary layers thin out over the continental foreland. The other side of the geosyncline is often so strongly deformed or covered by younger rocks that the nature of that margin at the time of deposition is obscure.

In plate-tectonics terms the deposits in trenches adjacent to island arcs of western America may be thrust towards the continent, plastered onto the inner wall of the trench and eventually be exposed as 'geosynclinal' deposits. It seems questionable whether the semi-consolidated deposits of the fill have sufficient strength to fold abruptly to form the steeply-sloping surface of the commonly 500-1 500 m high trench wall which in many areas is steeper than $15-20^\circ$ and locally may exceed 40° .

Some geosynclines appear to have been inverted by vertical movement, which raised the altered and metamorphosed rocks of the centre, eventually exposed as a 'root zone', and the more superficial rocks slid away from the uplifted area to create foreland folds by gravity tectonics. This mechanism does not explain the lack of symmetry better than any other.

The folding of geosynclinal sediments, whether syntaphral or foreland folding is not necessarily related to squeezing of the geosyncline. Nor is it necessarily related to uplift of mountain-building proportions. Folding often has nothing to do with any uplift or with mountain building. Many geosynclines finish their life as belts of sedimentation and folding not as mountain chains but as plains. Subsequently mountains may be lifted on the same site, but that is another story.

Various analogies have been suggested from modern situations to account for geosynclines. The sedimentary deposits off the eastern coast of North America provide one example. The near shore deposits form one basin, equivalent to the miogeosyncline, but a deeper trough has sediments containing many turbidity current deposits

which were thought to be analogous to a eugeosyncline, although of course there is no volcanic component to these deposits. Another actualistic model for geosynclines was proposed by Dietz (1963) who suggested that the continental-terrace sedimentary wedges are modern miogeosynclines and the subjacent continental-rise sedimentary prisms are modern eugeosynclines, together forming a couplet. Thrusting by a spreading sea floor is supposed to convert the continental-rise sediments into folded eugeosynclinal deposits. The mechanism is shown diagrammatically. This model has no role for island arcs or trenches and would be applicable perhaps to sediments on trailing edges of continents where these features do not exist.

If the Atlantic margin provided a satisfactory analogy of a geosyncline, it might be expected that pre-Mesozoic Atlantic-type continental margins would be found. Several possibilities have been suggested, but none is satisfactory (Smith, 1976, p. 273).

Other possible modern examples of non-volcanic geosynclines are provided by the sedimentary basins of California. The Los Angeles basin has 12 000 m of marine deposits; the Ventura basin has 11 000 m of marine clays; the San Joaquin basin has 6 000 m of Miocene and Pliocene sediments and 300 m of non-marine Pleistocene deposits.

The modern Bengal Deep-Sea Fan is regarded by Curray and Moore (1974) as the uppermost 4 km of the geosynclinal pile of sediments filling the Bay of Bengal in the northeast Indian Ocean. The fan post-dates the first collision of India and Asia and the uplift of the ancestral Himalayas at the end of the Palaeocene.

□ 1.2.6 SELF ASSESSMENT QUESTIONS :

- a) Why should you know about plate tectonics?
- b) Why is plate tectonics a unifying theory of geology?
- c) What are the major types of plate boundaries?
- d) How do plate tectonics theory can explain the orogenesis?
- e) Why should you know about neo-tectonics?
- f) What are hotspot and what do they tell us about plate movement?

□ 1.2.7 REFERENCE :

Paul, Ashis Kr. (2005), *Structure, Tectonics and Geomorphology* (in Bengali) Progressive Publishers, Kolkata, p. 228.

PART-I
Paper - I, Module - I
GEO-TECTONICS & GEOMORPHOLOGY
Unit – 03

MODULE STRUCTURE

- ☐ 1.3.0 Introduction
- ☐ 1.3.1 : Objective
- ☐ 1.3.2 : Evolution of Land forms by the Fluvial Process
- ☐ 1.3.2.1 : Land forms Developed by the Interruptions of the Fluvial Cycle
- ☐ 1.3.2.2 : Evolution of the Land forms Developed by Karst Process
- ☐ 1.3.3 : Evolution of Land forms by Glacial Process
- ☐ 1.3.4 : Evolution of Land forms by the Aeolian Process
- ☐ 1.3.5 : Evolution of Land forms by the Coastal Process
- ☐ 1.3.6 : Sea-level Changes and Raised Beach Features
- ☐ 1.3.7 : Self Assessment Questions
- ☐ 1.3.8 : References

Paper - I, Module - I
GEO-TECTONICS & GEOMORPHOLOGY
Unit – 03

1.3.0 INTRODUCTION :

This unit includes the topic of geomorphology. The forms, processes and evolution of various landscape are to be discussed in the form of fluvial, glacial, aeolian, karst and coastal geomorphology. Resultant land forms due to interruptions caused by climatic extremes and tectonics are also highlighted in this unit of the module - I.

1.3.1 OBJECTIVES :

- i) To discuss the forms, processes and evolution of fluvial, glacial, aeolian, karst and coastal landscapes.
- ii) To identify the interruptions of the fluvial cycle and their impact on geomorphological development.
- iii) To discuss the reason of sea level changes and to identify the coastal features developed due to the result of uplift.

1.3.2 EVOLUTION OF LANDFORMS BY THE FLUVIAL PROCESS :

Throughout most the world, runoff waters are the dominant geomorphic agency. Except for areas now covered by glaciers there are few places where rainfall does not have an opportunity to perform geomorphic work. In the most absolute deserts it rains, even though infrequently, and to a large degree, desert landforms are the work of running water. In areas that were covered by ice sheets during the Pleistocene and were veneered with glacial deposits having a unique topographic expression, streams and sheet wash have already made an impression upon the landscape and are rapidly changing it.

Δ **Surface Runoff :** Obviously we have streams because we have rainfall, but rain may fall upon the earth's surface without producing streams, immediately at least. Sheet flow of runoff proceeds concentrated or stream flow. The distance that runoff flows before it becomes concentrated into definite channels was called by Horton (1945) the *length of overland flow*. The distance of this type of flow was explained by him in terms of what he called the infiltration theory of surface runoff. This theory is based upon two concepts : (1) that there is a limiting rate at which rainfall can be absorbed by the soil, known as the infiltration capacity, and (2) when runoff does take place from a soil surface there is a definite relationship between the quantity of water which accumulates upon a soil surface and the rate of surface runoff or channel inflow.

Horton listed six and possibly seven factors which determine the infiltration capacity of the soil : (1) soil texture, (2) soil structure, (3) vegetal cover, (4) biologic structures in the soil including such features as rodent perforations and amount of humus and vegetal debris (5) amount of moisture in the soil, (6) condition of the soil surface as determined by whether it has been cultivated recently or is baked or sun-cracked, and possibly (7) the temperature of the soil.

Δ **Streams and Valleys :** Valleys are so common on the earth's surface that we seldom bother to define them.

This lack of definition leads to the application of the terms of features that really are not valleys. True valleys are present over most of the land surface of the earth. They are known by such names as gully, draw, ravine, gulch, hollow, run, arroyo, gorge, canyon or by more poetic forms such as vale, glen and dale, but the one thing that is common to all is that they were cut by running water. They are negative land forms of varying size and shape occupied by either perennial or intermittent streams.

Δ Valley development : A valley takes form through the operation of three concomitant processes, valley deepening, valley widening and valley lengthening. Valley deepening is affected by several processes. They are : (1) hydraulic action; (2) corrosion or abrasion on the floor of the valley; (3) pothole drilling along the valley floor and at the base of waterfalls (this is really a special type of corrosion or abrasion but is perhaps unique enough to merit special designation), (4) corrosion; and (5) weathering of the stream bed (in the case of an intermittent stream) plus subsequent removal of weathered material by hydraulic action.

Valley width is the linear distance between valley sides and is expressed in terms of the cross section of a valley. The upper and outer parts of many valley sides are indefinite and must be delimited arbitrarily. The width of the valley floor is more definitely determinable, and we commonly think of valley width in terms of it. Valley widening may be accomplished in a number of ways. (1) Lateral erosion or plantation by the stream in a valley may remove material from the base of the valley side through hydraulic and corrosive action. This result in local oversteepening of the valley side with attendant undercut slopes, which favors slumping of materials into the stream. This process may operate during any stage of valley development but is most noticeable during maturity and old age because then valley deepening has essentially ceased and the effects of lateral erosion are more obvious. (2) Rain-wash or sheet wash on valley sides contributes in an important way to valley widening. (3) Gulleying on valley sides is another method by which valley widening is effected. Although more spectacular than sheet wash, it is doubtful if it is an important. (4) Weathering and mass-wasting may contribute to valley widening both directly and indirectly. Some valleys have such steep sides that little mantle rock remains on them. Under such conditions weathering may loosen material which moves directly downslope into the stream channels; more commonly weathering acts indirectly through production of mantle rock which then is moved downslope by creep, slump other types of mass-wasting and by sheet wash. (5) Incoming tributaries contribute to valley widening even though they be no more than over grown gullies. Valleys commonly widen noticeably where tributaries joint them because the valley wall here is being attacked from two directions.

Valley lengthening may take place in three ways. (1) Valleys may be extended by the process of head ward erosion. This is particularly significant in the growth of lesser valleys. It would be erroneous to assume that a major stream like the Mississippi River started as a short stream heading near the sea and then extended itself by head ward erosion to its present source. No great valley was ever produced in this way, but many minor tributary valleys were so formed. The most common type of head-ward erosion involves the extension of a ravine by incoming sheet waters with attendant weathering and slumping at the ravine head. A second type of head ward erosion is effected by spring sapping. This takes place where a spring emerges at the head of a valley. The overlying rock is undermined by solution and weathering and subsequent slumping above the spring may produce an abrupt valley head. Valleys may extend head ward into or across swampy areas accompanying filling and drainage of such areas or after deepening of the valley downstream from the swamp has lowered the water table sufficiently to permit

valley incision across it. (2) Valleys may be lengthened through increase in size of their crooks or meanders. As long as a meandering stream is confined by valley walls increase in its crookedness will augment the valley length. (3) Valleys also may lengthen at their termini. Uplift of the land or lowering of sea or lake level will result in extension of the valley form across the newly exposed land. Many of the valleys across the Atlantic and Gulf coastal plains have been extended in this way as have valleys which empty into the present Great Lakes at altitudes lower than their Pleistocene predecessors. Streams may be lengthened at their termini by seaward extension of deltas. The lower Mississippi well illustrates this type of lengthening, as its mouth at the beginning of the Pleistocene was approximately 125 miles inland from its present position. It is doubtful, however, whether the channel of the Mississippi across its delta should be called a valley.

Δ **Base level and its varieties :** It is generally recognized that there is a downward limit to valley deepening. This limitation on vertical erosion is known as *base level*. When Powell (1875) proposed the concept of base level he introduced an idea which probably has been more abused than any other concept except that of the peneplain.

Davis (1902) seems to have thought likewise, for in referring to Powell's ideas he stated :

Base level, as thus defined, seems to include three ideas; first, the grand or general base level for sub-aerial erosion is the level of the sea; second, a base level is an imaginary, sloping surface which generalizes the faint inclination of the trunk and branch rivers of a region when the erosion of their channels has practically ceased; third, local and temporary base levels are those slow reaches in a river which are determined by ledges in its course farther downstream.

The more conventional viewpoint is that Powell intended to suggest that such features as hard-rock barriers or lakes athwart a stream course could act as local or temporary base levels for portions of the stream above such barriers. Johnson referred to them as "temporary local base levels" and thus combined the two into one idea. We cannot judge what Powell's intentions were, but a case can be made for the view that there are both local and temporary base levels. If these two are distinguished, *local base level* would be the downward limit of land reduction for any region as controlled by the level of streams across the region which are graded to the sea.

It is generally agreed that such features as hard-rock barriers or lakes along a stream may act as temporary base levels for the areas above them. The only question is whether they should also be considered as local base levels. It seems appropriate to describe a region as being at temporary base level whenever it is graded toward some level which in turn is not graded to the sea. Used in this sense, *temporary base level* is different from local base level.

Δ **The graded stream :** Davis, building upon Gilbert's ideas, developed the concept that a stream rather early in the geomorphic cycle attains that slope or gradient which under existing conditions of discharge and channel characteristics is just sufficient for transportation of its load. Such a stream is said to be *graded* or *at grade*. Grade is not to be confused with gradient; all streams flow over slopes and thereby possess gradients which may be expressed in degrees, per cent, or perhaps in feet per mile. A graded stream is not a stream which is loaded to capacity as often stated for streams are probably never loaded to capacity. Neither is it a stream which is neither eroding nor depositing as sometimes said for erosion in one part of a stream channel and deposition in another part are as characteristic of a graded stream as of a non-graded one. Although a stream at grade has

attained a particular gradient, it is not gradient alone that determines the graded condition. Other factors are involved, such as velocity of the stream, which is not entirely determined by gradient, channel characteristics and the caliber of the material that the stream has to transport.

The factors influencing variations in gradient along a graded stream were stated by Mackin (1948) as follows :

The longitudinal profile of a graded stream may be thought of as consisting of a number of segments, each differing from those that adjoin it but all closely related parts of one system.

Each segment has a slope that will provide the velocity required for transportation of all of the load supplied to it from above, and this slope is maintained without change as long as controlling conditions remain the same. The graded profile is a slope of transportation; it is influenced directly neither by the corrosive power of the stream nor bedrock resistance to corrosion.

Some changes from segment to segment in factors controlling the slope of the graded profile are matters of geographic circumstance that are not systematic in any way; these included the down valley increase in discharge, and down valley decrease in load relative to discharge, that characterize trunk stream flowing from highland areas through humid lowland. Other changes, as down valley decrease in caliber of load by reason of attrition may be more or less systematic between tributary junctions. Still others, as change in channel characteristics, are partly dependent on changes in load and discharge; the channel characteristics are determined chiefly by caliber of load and rate of lateral shifting of the channel and the rate of channel shifting is itself dependent upon velocity and credibility of the banks.

These changes are usually such as to decrease slope requirements in a down valley direction but, because none of them is systematic, the graded profile cannot be a simple mathematical curve in anything more than a loose or superficial sense. We can proceed towards an understanding of the graded profile, not by "curve matching", but by rigorous analysis of adequate sets of data for unit segments of natural and laboratory streams numerous and varied enough to reveal the effect of variation of each of the factors separately. An essential prerequisite for efficiency in the gathering and analysis of the data is recognition of the difference between the graded profile that is maintained without change and the upgraded profile that is being modified by up building and down cutting.

Some streams may have both graded and upgrades stretches along their courses. The breaks between them do not consist of the minor inflections in stream profile that are common along most graded streams, but rather of sharp changes in gradient marked by rapids or waterfalls. A stream profile marked by such abrupt changes in gradient is called an *interrupted profile*. The main problem which an interrupted profile presents is in determining whether it reflects the local effect of resistant bedrock along the valley or whether it has cyclical implication and was produced by diastrophic or eustatic changes in base level. The sharp inflections in an interrupted profile are called *nickpoints*. It is maintained by some that they are always explainable in terms of varying rock hardness along a stream course and hence have no cyclical significance. A temporary base level such as a hard-rock barrier or lake, may exist along a stream and cause a section of the stream above it to attain a graded condition while below the barrier the stream remains ungraded. Such a nickpoint is a temporary thing and usually has no cyclical implication. Along some streams, however, nickpoints are found which do not seem to be related to more resistant rock. They have been interpreted as the head and limits of successive periods of base leveling in multicyclic

valleys. This origin for nickpoints, found along valleys cut in homogeneous rock, would seem likely.

Δ **Cross profile of valleys :** The cross profile of a valley may shed considerable light upon its geomorphic history, as well as indicate the influence of local geologic and climatic controls. It is sometimes carelessly stated that during youth valley deepening exceeds valley widening. In a sense valley deepening is the most conspicuous effect of stream activity during youth, but it seldom exceeds valley widening in amount.

The cross profile of a valley is often asymmetrical and recognition of the causes of asymmetry is essential to a proper interpretation of valley history. There are several causes which may produce asymmetrical cross profiles.

- 1) Probably the most common one is the production of *undercut* and *slip-off slopes* attendant upon meandering and lateral erosion. The valley side on the outside of a meander often displays an over steepened and undercut slope, whereas the opposite valley side descends more gradually from the upland to the valley floor to form what is called a slip-off slope.
- 2) An asymmetrical cross profile may reflect structural controls. Valleys cut in alternating strong and weak strata having an essentially horizontal attitude will frequently exhibit *structural benches* upon the strong strata. These benches are usually not continuous for long distances and may be present on one side but lacking on the other side of a valley. They must be distinguished from alluvial terraces and strath terraces (bedrock surfaces marking former base levels of erosion) which have different implications as far as the history of the valley is concerned. Structural benches have no implication of former base level control or of partial erosion cycles, as do strath terraces, nor do they represent a stage of aggradations followed by partial excavation of the valley fill, as do alluvial terraces. Usually it is not difficult to recognize structural benches, for it will often be evident that their existence is dependent upon specific lithologic conditions.
- 3) Where valleys roughly parallel the strike of inclined strata it is commonly observed that the stream has a marked tendency to shift its position down the dip of the beds, even following the surface of a particularly resistant bed. This leads to an asymmetrical cross profile and the process responsible for it is referred to as *homoclinal shifting*.
- 4) An east-west valley may display marked asymmetry in cross profiles as a result of the direct or indirect effects of the microclimatic differences on the two sides of the valley. As yet, few quantitative studies have been made of this possible factor in valley development. The asymmetry results largely from the effects of differences in exposures upon the rates of weathering, mass-wasting and erosion.

Δ **Stream meandering and lateral erosion :** The *valley flat* is the fundamental land form produced by lateral erosion. This is a bedrock erosional surface which may go unrecognized, because it is commonly veneered with alluvium. The first stage in the development of a valley flat involves primarily the elimination of the spurs between stream meanders. This process is frequently referred to as spur trimming. This is effected by lateral erosion as a stream flows against the upstream and downstream sides of spurs in going around meanders. Continuation of this process will reduce the width of a meander spur until only a narrow *meander neck* separates the stream on the two sides of the meander. Continued impingement of a stream against the two sides of a meander neck will result in the stream cutting through it and forming a *meander cutoff*. A remnant of the meander spur will then be left as

a *meander core* and the abandoned route around it becomes a *meander scar*. The stage which immediately precedes the formation of a detached meander core is that in which a stream flows across a meander neck beneath a *natural bridge* or *arch*.

With the beginning of a valley flat we get the initiation of a *flood-plain*, the deposit of alluvium which covers a valley flat. It may be impractical to distinguish valley flat from floodplain, yet it should be kept in mind that along many valleys the floodplain is but a veneer of alluvium over a bedrock floor which has been cut by lateral erosion. The initial floodplain is different in many respects from the final floodplain of old age. Its beginnings appear in youth before the stream has ceased downcutting as crescent-shaped or slightly sinuous strips of coarse alluvium along the convex banks of stream meanders.

Opinion differs as to whether there is a limit to the extent to which lateral erosion may continue to widen valleys and reduce interstream upland tracts. Davis attributed the reduction of interstream areas more to downwasting than to lateral planation. He thought that the gently undulating topography which presumably develops toward the end of a fluvial cycle, to which he gave the name *peneplain*, was only in part a result of lateral erosion. Floodplains were incidental as compared with the more extensive interstream areas which were *worn down* rather than *planed off*. Johnson (1931) and Crickmay (1933) in particular have attached much greater significance to lateral planation in the process of peneplain development.

Δ An idealized fluvial cycle : Operation of the erosional agents leads to an orderly sequence of land form development. An attempt is made below to indicate some of the expectable changes in landscape characteristics during a somewhat idealized fluvial cycle. There is no necessary implication that progress of the cycle leads to the type of erosional surface termed by Davis as a peneplain. It should be recognized that variations in altitude, geologic structure, diastrophic history and climatic conditions may result in notable departures from the sequence of land forms described below.

Δ Youth :

- 1) There will be a few consequent trunk streams but few large tributaries. Numerous short tributaries and gullies will be extending themselves by head ward erosion and developing valley systems.
- 2) Valleys will have V-shaped cross profiles and will be shallow or deep depending upon the height of the region above sea level.
- 3) There will be a general lack of floodplain development except along trunk streams and valley sides will rise from near the streams' edges.
- 4) Interstream tracts may be extensive and poorly drained. Lakes and swamps may exist in the interstream areas if these areas are not well above the local base level.
- 5) Waterfalls or rapids may exist where stream courses cross beds of particularly resistant rock. They are most typical of early youth and will have disappeared before maturity is attained.
- 6) Stream divides will be broad and poorly defined.
- 7) Stream meandering may exist in youth, but the meanders are those on a flat and undissected initial surface

or are closely confined meanders in valleys incised below the upland surface.

Δ Maturity :

- 1) Valleys have extended themselves so that the region now has a well-integrated drainage system.
- 2) Adjustment of streams to such lithologic variations as exist may be evident in the existence of some longitudinal tributaries along belts of particularly weak rock.
- 3) Stream divides will be sharp and ridge-like resulting in a minimum of interstream uplands.
- 4) Any lakes or waterfalls that existed in youth have been eliminated.
- 5) Floodplain tracts constitute a considerable portion of the valley floors.
- 6) Meanders may be conspicuous but, in contrast to those of youth, they are free to shift their positions over the floodplains.
- 7) The widths of the valley floors do not greatly exceed the widths of the meander belts.
- 8) The maximum possible relief exists.
- 9) The topography consists not so much of valley bottoms and upland tracts as it does of slopes of hillsides and valley sides.

Δ Old age :

- 1) Tributaries to trunk streams are usually fewer in number than in maturity but more numerous than in youth.
- 2) Valleys are extremely broad and gently sloping both laterally and longitudinally.
- 3) There is marked developed of floodplains over which streams flow in broadly meandering courses.
- 4) Valley widths are considerably greater than the widths of the meander belts.
- 5) Interstream areas have been reduced in height and stream divides are not so sharp as in maturity.
- 6) Lakes, swamps and marshes may be present but they are on the floodplains and not in interstream tracts as in youth.
- 7) Mass-wasting and chemical denudations are dominant over fluvial processes.
- 8) Extensive areas are at or near the base level of erosion.

The above outline attempts to present the regional picture and not that of individual valleys. A region may be classed as having mature topography and have within it individual valleys which are still youthful and others which may have attained early old age. Usually we are more concerned with the topographic age of areas than of individual valleys and it is the picture of the total assemblage of land forms that determines its topographic age.

□ 1.3.2.1 LANDFORMS DEVELOPED BY THE INTERRUPTIONS OF THE FLUVIAL CYCLE :

Rejuvenation may result from causes which are dynamic, eustatic or static in nature.

Dynamic rejuvenation may be caused by epeirogenic uplift of a land mass, with accompanying tilting and warping. Such movements may be rather localized and associated with neighboring orogenic movements or they may be, as thought by some, worldwide in nature. Localized down tilting, warping or faulting of a drainage basin will result in a steepening of stream gradients followed by downcutting by the streams which now have transporting power in excess of that required for transport of their loads. The effects of seaward tilting would presumably be felt along the entire stream course and immediately reflected in deepening of the valley by the stream in it. Its effects will be felt immediately, now ever, only by those streams whose courses roughly parallel the direction of tilting. A stream flowing at right angles to the direction of tilting will respond to rejuvenation only after the stream which it joins has deepened its valley so as to leave the tributary out of adjustment. Even then the effect is felt only at the mouth of the tributary rather than along its entire length.

Eustatic rejuvenation results from causes that produce worldwide lowerings of sea level rather than localized changes in base level. Two types of eustatism have been recognized. *Diastrophic eustatism* is change of sea level resulting from variation in capacity of the ocean basins, whereas *glacio-eustatism* refers to changes in sea level produced by withdrawal or return of water to the oceans, accompanying the accumulation or melting of successive ice sheets. Eduard Suess was the father of the idea of diastrophic eustatism. Bauling (1935) has been one of its modern proponents. He recognized the importance of glacio-eustatism but thought there is evidence of worldwide lowerings of sea level during the Pliocene and Pleistocene of a greater magnitude than would have been possible through glacio-eustatic changes. He thought these lowerings of sea level were the result of epeirogenic movements to be associated with organic movements and a consequence of them rather than independent of orogenic movements, as thought by some geologists. Eustatic lowering of sea level will cause rejuvenation of a stream at its mouth. Regarding a stream toward the new base level will progress upvalley. The result may be an interrupted profile with the point of intersection of the old and new base levels being marked by a nickpoint which proceeds upstream as the new base level is extended headward.

Streams may exhibit signs of renewed youth from changes which involve neither uplift of the land nor eustatic lowering of sea level. This has been called *static rejuvenation* (Malott, 1920). At least three changes may produce static rejuvenation. They are decrease in load, increase in runoff because of increased rainfall, and increase in stream volume through acquisition of new drainage by stream diversion or derangement.

Δ **Topographic evidence of rejuvenation :** Rejuvenation of a region is suggested when we have with it a *topographic unconformity* or discordance. By this is meant a lack of conformity between the topographic forms of the upper parts of valleys and adjacent uplands and those of the lower parts of valleys. The upland and upper valley forms are those of maturity or old age, whereas the lower valley forms are those of youth. A common example of this is the *valley-in-valley* cross profile, also described as a *two-story valley* or *two-cycle valley*. The chief characteristic of such a cross profile is a break in slope along the valley side marked by a shoulder separating the youthful valley form below from the older valley form above. This bedrock shoulder is a remnant of the valley floor formed during a cycle of erosion preceding the present. Such bedrock terraces are not to be confused with structural benches which correlate with and are caused by resistant strata. Rather than being structurally controlled levels they mark former base levels of erosion and extend across rocks of varying lithology.

The long profile of a valley may show evidence of rejuvenation, consisting of nickpoints along the valley which cannot be attributed to hard-rock barriers. It may even be possible in this way to detect evidence of rejuvenation in valleys which have not reached a sufficiently advanced stage of development to have had valley flats along them which would form terraces after rejuvenation.

Incised or inclosed meanders are usually considered evidence of rejuvenation. Five terms, incised, entrenched, intrenched, enclosed and ingrown, have been applied to them with varying meanings. It seems to be more common practice to use incised and enclosed as generic terms to include any meanders enclosed by rock walls. We shall therefore consider *incised meanders* and *enclosed meanders* as synonymous group names for all meanders set down in bedrock, regardless of the cross profile of the valley. Two types of incised or enclosed meanders are generally recognized : (1) *entrenched* or *intrenched meanders* (the difference is only in spelling), which show little or no contrast between the slopes of the two valley sides of a meander curve, and (2) *ingrown meanders* (Rich, 1914), which exhibit pronounced asymmetry of cross profile with undercut slopes on the outside of the meander curves and slip off slopes on the inside.

Although it would seem likely that first-cycle incised meanders will usually be the ingrown type, it does not follow that all ingrown meanders are first-cycle meanders. Ingrown meanders are probably the expectable type of incised meanders, even when inherited and entrenched meanders are likely to develop under special conditions. It is not clear, however, why in one area downcutting is accompanied by significant lateral erosion and in another area not. The problem is further complicated where both entrenched and ingrown meanders are found along different stretches of the same valley. This suggests that the two types may be related to local differences in geologic structure or lithology.

Δ **Aggradation** : We have considered dynamic, eustatic and static causes which may effect rejuvenation of streams. For each type the effects are the same as those that would result from an actual lowering of base level. It will be evident that in each of the three types of change the effects may be of such a nature as to cause streams to aggrade rather than deepen their valleys. Sinking of the land may result in the downing of the portions of valleys adjacent to the sea, which will be accompanied by alleviation or valley filling upstream. Actual drowning of valley is not necessary to cause a stream to become an aggrading one. Downward tilting of the land inland will produce the same effect as soon as stream's gradient becomes too low for transport of its load.

The effect on streams of rise of sea level is the same as if the land sank. Fish (1939) has interpreted the geomorphic history of the lower Mississippi Valley as one of alternation between alleviation during interglacial high sea levels and valley trenching during glacial low sea levels. He has described four terraces which he thought represent the tops of valley fills made during times of high sea levels.

Δ **Termination of the fluvial cycle** : Because of the length of time required for the completion of a fluvial cycle there is always the possibility that a geomorphic cycle instead of merely being interrupted may actually be terminated. Davis foresaw the possibilities of this and applied the somewhat inappropriate term "accident" to the termination of a cycle. He recognized two types of accidents which might end the so-called normal cycle : volcanism and climatic change with ensuing glaciation or aridity. To these we could add submergence beneath the sea of a land mass and possibly complex diastrophism.

Δ **Slow divide migration** : It is commonly observed that, in the later stages of the fluvial cycle, stream divides for streams of similar size are somewhat uniformly spaced and stream gradients are roughly equal on opposite sides of divides. This is not always true in the earlier stages of the cycle, for then it is not unusual to find unequal stream lengths and gradients on the two sides of a divide. It was recognized by Gilbert (1877) in his classical study of the Henry Mountains of Utah that a stream flowing down the steeper slope of an asymmetrical ridge erodes its valley more rapidly than one flowing down the more gentle slope and as a result the divide line migrates away from the more actively eroding stream toward the less actively eroding one. This principle has been referred to as the *law of unequal slopes*. In a region of gently or moderately dipping weak and strong sedimentary rocks, *cuestas* or *homoclinal ridges* will form with steep front slopes facing up the dip and more gentle back slopes extending down the dip, under which conditions divides will migrate down the dip of the rock and the shorter and steeper streams down the front slope will extend their drainage at the expense of the lower-gradient streams down the back slope.

Δ **Rapid shifting of divides** : Rapid divide shifting may result from either stream diversion or derangement. *Stream diversion*, as here used, refers to the various ways by which streams may actively effect drainage changes in contrast with stream derangement in which drainage changes are accomplished by agents other than streams. Stream diversion may come about either through stream piracy or aggradation. Stream piracy involves active erosional attack by one stream upon the drainage of another with resulting diversion of part of its waters to the pirating stream. The most common cause of stream piracy is the ability of one stream to maintain and extend its valley a lower level than that of an adjacent stream. Stream piracy may take place by abstraction, head ward erosion, lateral planation and subterranean diversion. The term abstraction is often applied to the simplest type of piracy, which results from competition between adjacent consequent gullies and ravines. This is an unspectacular but important process whereby innumerable minor drainage lines become integrated into a few dominant stream courses. It takes place chiefly at the upper ends of drainage lines and can be observed in progress among gullies on almost any steep slope. It is of importance in the ultimate location of the divides between the lesser tributaries of a river system.

Stream piracy by head ward erosion may arise from either of two causes. One stream may have an advantage over another because it is cutting its valley in more easily erosible rocks, or piracy may take place because one stream has a much steeper gradient than another.

One of the best-known examples of piracy in this area is that which resulted in the formation of Snickers, Ashby and Manassas gaps. This represents an illustration of progressive stream piracy by head ward erosion. These gaps are now *wind gaps*, a term applied to former *water gaps* through which streams no longer flow, but formerly a stream flowed eastward through the Blue Ridge in each of them.

Unequal gradients along two opposed and competing streams commonly are found where there are *cuestas*, plateaus or tilted fault blocks with steep front and gentle backslopes. There need be no difference in rock resistance along the two competing streams for piracy to take place. The classic locality for this type of piracy is along the front of the Catskill Mountains where Kaaterskill Creek and other streams that flow down the steep east-facing front have captured some of the head ward drainage of the westward-flowing streams. This type of stream piracy

results fundamentally from the fact that the pirating streams have lower local base levels of erosion and consequently steeper gradients in their head ward portions at least.

Stream piracy may also result from lateral erosion or planation. This type of piracy is to be expected after streams have become graded and lateral erosion has become dominant.

Stream diversion not involving piracy may take place when, after rapid aggradation of its valley, a stream spills out of the valley into a lower route.

Δ **Stream derangement** : Stream derangement has little in common with stream piracy except that changes in stream courses are involved. Derangement may result from glaciation, wind deposition or diastrophism. Glaciation probably has been responsible more often for stream derangements than any other single cause.

Δ **Misfit Rivers** : It has been observed that some streams are not proportionate in size to the valleys that they occupy. These have been called *misfit rivers*. If a stream, or more correctly the size of the stream meanders, is too small for the size of the valley, the stream is said to be underfit (Davis, 1913); if too large, it is referred to as *overfit*. It is difficult to cite examples of overfit rivers, or streams with floodplains too small for the size of the stream. Hence there may well be a question whether overfit streams exist. The reason for their scarcity or nonexistence may be that a stream cannot long remain overfit, for an increase in volume will be accompanied by increased erosive power and rapid adjustment of valley size to river. The underfit condition can persist indefinitely; hence many examples of such streams exist.

An underfit condition commonly results from drainage changes effected by stream diversion (piracy) or derangement. Davis particularly stressed beheading of a river by stream piracy as a cause of the underfit condition.

Dury (1960, 1964) has proposed a theory regarding the cause of underfit streams which maintains that this condition is widespread in the United States and Europe. He believed that there is such widespread disparity between former river discharges and present stream volumes as to imply regional changes in regimens of runoff that can hardly be explained by any other cause than climatic change in the amount of precipitation.

Δ **River terraces and their significance** : River terraces are topographic surfaces which mark former valley flood levels. In the main they are vestiges of former floodplains, although some may have little or no alluvium on them and may thus be classed as *bedrock terraces* in contrast with *alluvial terraces*, which consist of gravel sand, and finer alluvium. As Gilbert (1877) has stressed, river terraces are largely the products of stream erosion and not of stream deposition. Although it is not unusual for the surface of a terrace to be a depositional surface, the terrace itself came into being only through the development by erosion of another valley flat below the top of the former valley floor. This is true regardless of whether the material beneath the terrace is alluvium or bedrock.

Bed rock terraces may have a thin veneer of alluvium upon them, but it is usually inconspicuous. They are essentially former valley flats which were produced in most instances through lateral erosion by graded streams.

River terraces are produced by surges of erosion along river valleys and hence reflect periods of rejuvenation which have affected streams. Cotton (1940) has suggested that river terraces may be classified as either cyclic or noncyclic. *Cyclic terraces* represent former valley floors formed during periods when valley deepening had largely stopped and lateral erosion had become dominant. The vertical distance that a terrace is above the present floodplain

or another terrace below it represents the amount of valley deepening which took place after rejuvenation. The distinguishing feature of cyclic terraces is that, for any set of terraces, remnants on opposite sides of the valley will be paired or will correspond in altitude along any particular valley stretch. This implies that there was not continuing, concomitant downcutting and lateral erosion, but that valley deepening essentially ceased for a time and allowed lateral erosion to form an extensive valley flat at the local base level of erosion. Cyclic terraces may imply partial erosion cycles related to intermittent uplift, or they may represent alternations between periods of aggradation and degradation relate to static changes. *Noncyclic terraces*, according to Cotton (1940), will be nonpaired terraces. They imply that there was continued downcutting accompanied by lateral erosion. Under these conditions the meander belt of the stream shifted back and forth over the valley and by the time that it moved from one side of the valley to the other, the valley floor had been lowered somewhat and terraces left on opposing valley sides were at different altitudes. Nonpaired terraces imply continued slow uplift of the land or other changes that would have a similar effect upon a stream.

Alluvial terraces may be better preserved in some parts of a valley than in others because of the protective effect of buried bedrock beneath an alluvial cover. These have been called *rock-defended terraces*. Alluviation may bury bedrock spurs which, when encountered during ensuing valley excavation, will deflect the stream from that part of the valley and protect the alluvial fill both upstream and downstream from the bedrock buttress from removal by lateral erosion.

Δ **Terrace surfaces and slopes :** The cross profile of a terrace is seldom flat. Numerous processes operate to destroy its flatness, if it ever possessed it. Chief among these are slope wash from the valley side with accompanying deposition of colluvium on the upper part of the terrace and deposition of alluvial fans where tributaries to the main valley debouch from the adjacent upland. Locally, landslide, slump and mudflow materials may obscure the terrace form. Even where the initial terrace surface is well-preserved there may be considerable relief. Such features as stream bars, swales representing former channels of braided streams, meander scars and natural levees give relief to the terrace surface.

Warping or tilting of an area after terrace development may in rare instances cause terraces to slope upvalley, but they will more likely lessen or increase their downvalley slopes. Warped terraces are useful in determining the amount and nature of recent deformation.

Δ **Correlation of terraces :** The problem of terrace correlation is, as Johnson (1944) has stressed, often a difficult one. Even where terraces can be established unequivocally as river terraces, there still remain two major problems. If more than one terrace level exists along a valley, there is the problem of determining to which level any particular terrace tract belongs. There is also the more difficult problem of correlating terraces from one valley to another.

Some of the more common methods of correlation, such as identification by fossils, degree of dissection, comparative decomposition of boulders, lithology of the terrace material and depth of weathering of terrace materials may not be usable because the time intervals between the formation of successive terraces are usually too short for these tests to be applicable. In that event, near identity of terrace tract altitudes in a particular valley, permissible changes in altitude going up or downvalley being taken into account, along with continuity of terrace remnants,

must remain the only safe criteria for terrace correlation.

Correlation of terraces from one valley to another presents an even more difficult problem, for absolute altitudes are not likely to be the same even in adjacent valleys because of variations in stream profiles. The only positive method of correlating terraces in different valleys is to trace them until they join each other at the junction of the valleys or join with a similar terrace in a trunk valley to which the two valleys are tributary.

□ 1.3.2.2 EVOLUTION OF THE LANDFORMS DEVELOPED BY KARST PROCESS :

The word karst is a comprehensive term applied to limestone or dolomite areas that possess a topography peculiar to and dependent upon underground solution and the diversion of surface waters to underground routes. The term comes from the narrow strip of limestone plateau in Yugoslavia and adjacent portions of Italy bordering the Adriatic Sea, where there exists a remarkable assembly of features dependent upon subsurface solution.

Δ **Conditions essential to full development of karst :** There are four conditions which contribute to the maximum development of karst. First, there must be present at or near the surface a soluble rock, preferably limestone. Dolomite may suffice but it is not as readily soluble as limestone. Chalk is soluble, but it usually lacks one of the other prerequisites. Secondly, and this is one of the most important factors, this soluble rock should be dense, highly jointed, and preferably thinly bedded. This is a point frequently lost sight of by persons who lack firsthand familiarity with karst. It is sometimes stated in textbooks that the major prerequisite for karst development is the presence of a permeable or porous limestone. As a matter of fact, permeability in the sense of mass permeability is unfavorable. Permeability as permitted by numerous joints and bedding planes is favorable. If a rock is highly porous and permeable throughout, rainfall will be absorbed en masse and move through the whole body of the rock rather than be concentrated along restricted lines of movement. This is perhaps the major reason why karst features are weakly developed in the chalk region of England and France.

A third condition which favors excellent development of karst is the existence of entrenched valleys below uplands underlain by soluble and well-jointed rocks. This favors ready downward movement of groundwater through the rock. Solution in limestones becomes significant only when they act as conduits for the movement of groundwater. Good water circulation is a prime prerequisite; moving water carries on solution, but standing water does not.

Finally, the region must be one of at least moderate rainfall. It is significant that nearly all the notable karst regions are in areas of moderate to abundant rainfall. In general, arid and semiarid regions do not exhibit marked development of karst. In such areas where karst features exist they can be most logically explained as relict features formed during more humid climatic regimes.

Δ **Features characteristic of karst regions :** It is not intended that the student suppose that all the features discussed below will be found in any one area, but an attempt will be made to describe the more important forms that may be found at one place or another. Some of the features discussed are not strictly topographic in nature, but they are so intimately associated with topographic forms that they merit description. Others may be present only here and there and still be significant.

Δ **Terra rossa :** Surface and near-surface solution by descending ground water usually leaves a residue of a

red, clayey soil mantling the surface and extending down into opened joints. This material may be lacking on steep slopes, but it is characteristically present on moderate to gentle slopes. It may vary in thickness from a few to many feet and may completely mask the rock surface. The name *terra rossa* has been aptly applied to it. Terra rossa resembles in appearance the lateritic soil of the tropics, if indeed it is not a species of it. However, it is by no means limited to tropical or subtropical locations, being found as far north as southern Europe and southern Indiana.

Δ **Lapies** : Locally, where relief is considerable, limestone surfaces are bare of terra rossa and there is exposed an etched, pitted, grooved, fluted and otherwise rugged surface to which the name *lapis* is most commonly applied. A variety of lapis in which the solution grooves are rather minute and closely spaced has been called *rill lapis* (German, *nimnenkarren*), Cvijic (1924) has described and pictured the amazing diversity of surface and form which lapis exhibits in the Dalmatian karst region. He maintained that lapis is found chiefly on outcrops of naked rock and pointed out the influences of such factors as rock composition, texture, structure, surface slope and amount of vegetation upon its formation. He further claimed that it was rare on horizontal rocks, being replaced there by sinkholes. Others have argued that it develops under a soil and vegetation cover and is exposed later by erosion accelerated perhaps as a result of deforestation.

Δ **Sinkholes and associated forms** : By far the most common and widespread topographic form in a karst terrain is the sinkhole. In any major karst area sinkholes are found by the hundreds of thousands. Some idea of their frequency is indicated by a detailed mapping of 1 square mile of sinkhole terrain that was done by Malott and Shrock (Malott, 1945) in Orange County, Indiana. In this square mile they counted and mapped 1022 sinkholes. Malott estimated that there may be as many as 300,000 sinkholes in the karst area of southern Indiana and perhaps an equal or greater number in Kentucky. Although the sinkhole may be considered the fundamental relief feature of karst terrains, there are areas where they are sparingly present or even lacking.

Topographically, a sinkhole is a depression that varies in depth from a mere indentation of a few feet to a maximum of 100 feet or more. Most of them, however, range in depth from 10 to 30 feet. In area, they range from a few square yards to several acres. The most common form is funnel-shaped and broadly open upward, but there are many variations from this.

Fundamentally, they fall into two major classes, those that are developed slowly downward by solution beneath a soil mantle without physical disturbance of the rock and those that are produced by collapse of rock above an underground void. These two types have been referred to respectively, as solution sinks and collapse sinks. The term *doline* (Serbian – *dolinas*), which comes from the karst region of Yugoslavia, will be used to designate the first type and *collapse sink* will be applied to sinkholes that exhibit steep-sided, rocky and abruptly descending forms, resulting from collapse of the roof over an underground solutional opening. Dolines are far more common than collapse sinks, even in the mature stage of karst development. Runoff, which drains into dolines, usually finds its way to underground routes by slow percolation through the soil in the bottom of the sinkhole, but surface waters enter some sinkholes through surface openings called *swallow holes*.

In a region where sinkholes are numbered by the hundreds per square mile, individual sinkholes expand in diameter and coalesce to form compound sinkholes. These usually consist of a major solutional depression covering

several acres with numerous lesser dolines or shallow holes supersposed upon it. A special type of collapse sinkhole is the karst window. This term has been applied (Malott, 1932) to an unroofed portion of an underground stream course through which may be seen a stream which flows out of a cavern at one side, across an open space, and into a cavern at the opposite side.

A somewhat special form of sinkhole is the polje. This name has been applied to depressions of various forms and origins, but the typical Bosnian polje is an elongated basin with a flat floor and steep enclosing walls which owes its existence to solutional modification of downfaulted or downfolded limestone blocks. A polje may somewhat resemble an uvala, but it differs from it in origin and extent. A typical polje is a sizable feature covering many square miles, whereas an uvala commonly covers only a few acres. The largest polje in the Western Balkans, the Livno polje, is 40 miles long and 3 to 7 miles wide. As far as possible the term polje should be restricted to these structured controlled solutional forms.

Features that resembles poljes in form but differ from them in that they are not primarily of tectonic origin are the *solution-subsidence troughs* in the Gypsum Plain area of Texas and New Mexico (Olive, 1957). These are linear features ranging in width from a few hundred feet to 1 mile and in length from about one-half mile to 10 miles. Most of them are less than 20 feet deep. They have resulted from subsidence accompanying development of solution channels along joints or faults in the Castile formation of Permian age.

Δ **Other features of a sinkhole or karst plain** : In regions of nearly horizontal or gently dipping limestone strata, there develops what is most aptly designated as a *sinkhole* or *karst plain*, which over large areas acts as a veritable regional sieve with thousands of sinkholes acting as hoppers which convey meteoric waters to underground routes. The sinkholes are so effective in collecting surface runoff that few streams are able to cross a sinkhole plain. Larger streams which rise beyond it and are entrenched below it may do so, but lesser streams lose their waters to underground routes. They are then called *sinking creeks* usually and the point at which each terminates is its *sink*. The sink of a surface stream usually takes place in an observable shallow hole, but in some streams the water disappears through alluvial materials upon the stream bed so that the exact point of disappearance is not apparent. Some sinking creeks disappear in a single swallow hole, but others may have several and use one or all of them, depending upon the volume of water in the stream. Sinking creeks constitute an important aspect of karst terrains in relation both to their surface forms and to the development of caverns beneath a sinkhole plain. The underground course followed by a sinking creek before it reappears may be several miles long. Lost River in southern Indiana flows some 8 miles underground before it again becomes a surface stream. The Reka River, back of Trieste, is said (Wray, 1922) to have a subterranean section of its course that is over 18 miles long.

The little-used or unused continuation of a valley below the point where the stream in it sinks is called its *dry bed*. Most dry beds are used only during period of heavy rainfall when swallow holes and connecting underground conduits are unable to accommodate all the storm waters. Then streams may flow for a period in their dry beds until the storm waters have subsided enough to allow the underground courses to accommodate them. It frequently happens that a sinking creek may have terminated at a swallow hole for so long that it has cut its valley upstream from its sink so much below the level of the karst plain that the valley ends at the swallow hole. Such a valley is called a *blind valley*. The surface continuation of the dry bed below the sink through long-continued disuse becomes obscured and gradually loses the characteristics of surface valley. During periods of storm waters a blind valley

may become the site of a temporary lake as a result of the inability of the swallow hole or holes to accommodate the increased volume of water.

A special type of blind valley, sufficiently distinct from that just described to merit a special name, is the *solution valley* or *karst valley* (Malott, 1939). It differs from an ordinary blind valley in that it is not part of a typical karst plain but rather is completely or nearly enclosed by clastic rocks. It is a particularly significant karst feature because its development throws light upon the process by which a limestone terrain may expand at the expense of a region that originally was drained entirely by surface drainage and had the topographic characteristics produced by surface streams. A solution or karst valley represents a transitional stage between surface drainage and underground drainage.

Return to the surface of the waters of a sinking creek is a common phenomenon in karst areas. Springs of varying sizes and volumes are thereby produced. The terms *rise* and *resurgence* have been applied to the reappearance of surface waters which have been diverted to underground routes. Some rises can be associated with specific sinking creeks but more often the source of the water is vague and undetermined. The larger rises are artesian in character and issue from sub circular *rise pits* of considerable size and depth.

An interesting and sometimes unappreciated manifestation of karst development is the subterranean cutoff, which may be defined as an underground diversion of a part or the whole of a surface stream beneath a meander spur along an entrenched valley. A subterranean cutoff consists of three parts, a swallow hole on the upstream part of the meander spur into which the water sinks, an underground route or conduit through which the water runs, often called a *natural tunnel*, and a rise or resurgence of the water on the downstream side of the meander spur.

Δ **Natural tunnels and bridges :** A spectacular example of a natural tunnel, which is either the result of a subterranean cutoff or subterranean stream piracy, is the Natural Tunnel of Virginia (Woodward, 1936). Natural Tunnel is located in Scott County, Virginia, about 3 miles north of the town of Clinchport and about 8 miles north of the Tennessee-Virginia state line. This natural tunnel is unique because it serves as a railway tunnels for a branch of the Southern Railway.

Collapse of the roof of a natural tunnel will in time reduce its length until it is better described as a *natural bridge*. Since natural bridges are developed in several ways, those resulting from solution by groundwater may well be designated as *karst bridges*. The most famous karst bridge in this country, if not in the world, is Natural Bridge, Virginia.

Three distinct ideas have been proposed to account for Natural Bridge. Walcott (1893) advanced an idea that found its way into many textbooks and is found even yet in some. According to his theory, Cedar Creek once flowed as a surface stream across what is now Natural Bridge. At the time it did so the portion of the gorge of Cedar Creek upstream from Natural Bridge was non-existent. The James River, to which Cedar Creek is tributary, underwent rejuvenation, and in turn Cedar Creek felt the effect of this rejuvenation and started cutting a gorge headward. A waterfall was believed to have existed in this gorge at a point not far below the site of Natural Bridge. Water from the stream found an underground route above the falls through which it flowed and discharged below the brink of the falls. As this passage was enlarged more and more water was diverted to the underground

route until finally all the water followed this route, leaving a natural or karst tunnel which through time has dwindled in size to the present bridge.

A second theory (Woodward, 1936) is that Natural Bridge was produced by subterranean stream piracy. According to this theory, the upper part of present Cedar Creek was once a part of the drainage of east-ward-flowing Poague Run and what is now the lower part of the valley of Cedar Creek was occupied by a short stream which was the predecessor of present Cascade Creek.

A third theory, that of Malott and Shrock (1930), is that Natural Bridge is the result of a subterranean cutoff by Cedar Creek through a meander spur. This subterranean cutoff enlarged to the point where all the water of Cedar Creek went through it developing a karst tunnel which was gradually deroofed until only a small part of its roof was left as Natural Bridge.

Δ General statement of problem : In 1930 an extended paper by W. M. Davis entitled "Origin of Limestone Caverns" challenged some of the existing ideas of cavern formation and stimulated the thinking of geologists who were interested in limestone terrains regarding the mode of development of caverns. As a result ideas regarding caverns have been considerably clarified although there still remains divergence of opinions as to their origin. Actually, prior to 1930, ideas of how caves formed were rather vague. It was taken for granted that they were produced by solution by groundwater and there the matter usually ended.

The concept of a karst cycle has perhaps been most intimately associated with Cvijic. He according to Sanders (1921), recognized four stages in the evolution of karst, which he designated as youth, maturity, late maturity and old age. Youth, according to Cvijic, begins with surface drainage on either an initial limestone surface or one that has been laid bare and is marked by a progressive expansion of underground drainage. Lapiés and scattered dolines are particularly characteristic of this stage. No large caverns exist and underground drainage is far from complete. During maturity there is maximum underground drainage. Surface drainage is limited to short sinking creeks ending in swallow holes or blind valleys. Cavern networks are characteristic of this stage. This is the time of maximum karst development. Late maturity marks the beginning of the decline of karst features. Portions of cavern streams are exposed through what we have called karst windows. These expand to form large uvalas, and detached areas of the original limestone upland have begun to stand out of hums. Old age is marked by a return to surface drainage with only a few isolate hums remaining as remnants of the original limestone terrain.

Cvijic first published a summary of his concept of the karst cycle in 1918 in an article entitled "Hydrographic souterrainage et evolution morphologique du karst". Essentially the same ideas were expressed in this country by Beede in an article appearing in 1911, but because of the obscurity of the publication in which it appeared, his ideas were overlooked by most geologists. Beede divided the karst cycle into the three conventional stages of youth, maturity and old age, rather than the four as proposed by Cvijic. Youth is marked by the beginnings of diversion of surface drainage to subterranean routes with dolines as the characteristic land form of this stage. In maturity, there is a maximum of sinkholes and underground drainage with only major entrenched streams persisting as surface streams; sinking creeks, swallow holes, blind valleys, compound sinks and dolines by the thousands typify this stage. Old age is initiated with the beginning of return to surface drainage; karst windows, karst tunnels, natural bridges and hums are the most diagnostic features.

It should be recognized that, just as a region undergoing dissection by stream erosion may exhibit the characteristics of youth in one section, maturity in another and old age in another, so in a karst region the various stages of the karst cycle may be present. Much as in the fluvial cycle, the various stages move progressively into a region accompanying and following entrenchment of major drainage lines across the terrain. Areas remote from entrenched streams are likely to be less advanced in the cycle than those adjacent to them.

□ 1.3.3 EVOLUTION OF LANDFORMS BY GLACIAL PROCESS :

Δ **Introduction :** Glaciers today, except in high latitudes and at high altitudes, are of minor importance in the present day shaping of land forms, but those that existed during the Pleistocene left their imprint upon many millions of square miles of the earth's surface.

Some of our finest mountain scenery is to a large degree a product of glacial geomorphic processes, as are many interesting lowland land forms. The effects of glaciation are particularly significant in the thickly populated sections of east-central North America and north-west Europe.

Ahlmann (1948) has given a morphological classification of glaciers which is as follows :

- A) Glaciers that are continuous sheets from which the ice may move in all directions.
 - 1) Continental glaciers or inland ice covering large areas.
 - 2) Glacier caps, which are smaller than continental glaciers.
 - 3) Highland glaciers, which cover the highest and central portion of a mountain area.
- B) Glaciers confined to more or less marked courses which direct their main movement. Included in this group are both independent glaciers and outlets of ice from glaciers of group A.
 - 4) Valley glaciers of the alpine type.
 - 5) Transection glaciers, which more or less fill a whole valley system.
 - 6) Cirque glaciers, which occupy localized niches on sides of mountains.
 - 7) Wall-sided glaciers, which cover a side of a valley or part of it.
- 8) Glacier tongues afloat.
- C) Glacier ice which spreads in large or small cake-like sheets over level ground at the foot of glaciated areas. None in this class is independent, but all connect with some other type of glacier.
 - 9) Piedmont glaciers formed by fusion of the lower parts of 4, 5 or 7 above.
 - 10) Foot glaciers, which are the lower ends and more extended portions of types 4, 5 or 7 above.
 - 11) Shelf ice.

Δ **Wastage of glaciers :** Glaciers waste chiefly by melting; evaporation or sublimation may contribute to a glaciers loss of mass, but it has become evident in recent years that this process is not as significant as once thought (Sharp, 1960). Wastage by these combined processes is designated commonly as *ablation*. Some glaciers send tongues into bodies of water and lose mass by calving of the ice into the water body.

Δ **Effectiveness of glacial erosion** : Abundant striated, polished and grooved rock surfaces give evidence of effective glacial erosion.

Although much evidence can be marshaled to support the view that ice caps are effective erosional agents, it is also possible to cite evidence of little erosion. Locally, preglacially weathered bedrock underlies glacial deposits; at numerous localities glacial till of one age is found lying upon a soil profile developed during interglacial time upon an older till. It is largely because of the preservation of the paleosols on the older drifts that it has been possible to work out a Pleistocene chronology and stratigraphy.

In most of the world's high mountains ice streams have so profoundly modified valleys that their forms are distinctly different from those produced by fluvial erosion. Although the effects of glacial erosion are most strikingly displayed, depositional features are also present and add distinctiveness to the landscape.

Δ **Major features resulting from glacial erosion - The Cirque** : The most common and probably the most striking, land form in glaciated mountains is the *cirque*. This French term is applied to amphitheater-like basins (the term amphitheater is not strictly correct, for they are not completely enclosed by walls) which are found most frequently at valley heads, but which may not connect in their distal parts with valleys. The widespread distribution of cirques is attested by the many languages which have an equivalent word : the German *kar*, the Welsh *cwm*, the Scotch *corrie* and the Scandinavian *botn* and *kjedel*.

Conditions that seem to favor maximum cirque development are : (1) rather wide spacing of preglacial valleys so as to permit expansion without intersection of adjacent cirques at an early stage; (2) snowfall sufficient to form large snowfields and glaciers but not heavy enough to form ice caps and (3) fairly homogenous rocks which permit cirque extension equally well in any direction.

Δ **Glacial troughs** : Next to cirque, the most distinctive topographic feature in glaciated mountains is the glacial trough. Most glacial troughs were originally stream-cut valleys, but glaciers usually have so altered them that neither in cross profile nor in long profile do they resemble greatly stream-carved valleys. Trough characteristics may be scanty or lacking where glaciers were not large or where valleys were not sites of glaciers during Wisconsin glaciation.

A glacial trough heads not at a cirque headwall but at the lower edge of a cirque threshold. There is usually a conspicuous drop from a cirque threshold to the floor of a glacial trough, which has been called a *trough headwall*. Trough headwalls are likely to be most conspicuous where more than one cirque contributes ice to a glacial trough. Glacial troughs are notably irregular and ungraded in their long profiles. Few exhibit the relatively smooth, concave longitudinal profiles characteristic of stream-cut valleys. Descent of trough floors takes place in series of *glacial steps* or what is sometimes called a *glacial stairway*.

Δ **Nanging valleys** : In contrast to river valleys, which usually have their tributary valleys joining them accordingly, glacial troughs so commonly have tributary troughs or valleys joining the main trough discordantly that this may be considered normal. At one time hanging valleys were almost considered *prima facie* evidence of glaciation. We understand now that hanging valleys may be produced in several ways. They may characterize stream valleys, particularly if the tributary streams are intermittent or much smaller than the main stream.

Δ **Arêtes or serrate ridges** : As cirques enlarge by the sapping process mentioned above, preglacial uplands are gradually consumed by headwall recession. Whether this exhibits a cyclical progression as postulated by Hobbs (1910a, 1910b) may be open to some question. Certainly preglacial uplands exhibit varying degrees of destruction. The somewhat extreme view (Wooldridge and Morgan, 1937) that cirques are rather static in position and that sharpening and demolition of peaks takes place through ordinary weathering processes is not widely held. At an advanced stage of cirque recession, mountain divides may be nearly consumed. All that will be left will be a sharp sawtooth-like ridge, which the French have called an *arête*. Such a ridge consists essentially of alternating sags or glacial *cols* produced by intersection of opposed cirques and pointed peaks or *horns* represent unreduced portions of the original mountain range. Well known examples of horns are the Matterhorn and Weisshorn in the Alps. Somewhat related in origin to horns, but detached from the main mountain range, are *monuments* or *tinds*, as they are called in Scandinavia.

Δ **Truncated spurs** : In general, glacial troughs are straighter than unglaciated valleys. Glacial troughs do conform to the original valley course, but ice streams may straighten their troughs by abrasion of spur ends and thereby produced *truncated* or *faceted spurs*. Different degrees of spur truncation may be seen, varying from spur ends that have been cut off completely to partially trimmed spurs still recognizable from rocky knobs or nubbins.

Δ **Fjords and piedmont lakes** : Most students of glacial land forms agree that fjords are glacial troughs eroded by ice below sea level, but a few attribute the great depths of water in them to submergence of troughs formed above sea level. Fjords are characteristic features of shorelines in high latitudes are well developed along the coasts of Norway, Greenland, British Columbia, Alaska, Chile and New Zealand. It was once held by some that fjords were largely tectonic in origin, but that idea has largely disappeared. It is possible that the plans of some fjords reflect joint or fault control, but it seems unlikely that many of them were initially grabens.

A feature of a fjord that has evoked considerable speculation is the threshold or sill which is often found at its terminus. The depth of water here is typically much less than farther seaward or headward in the fjord. Thus a fjord is really an elongate basin which would become a lake, if sea level were lowered below the top of its threshold. It has been suggested that fjord thresholds are submerged terminal moraines and in some instances this may be partly true, but it seems more likely that they result from greater glacial erosion up fjords, where the ice was thick and actively eroding, and less erosion in terminal zones, where the ice was too thin for effective erosion.

In the same class with fjords are *fiards* and *trough lakes*. Fiards differ from fjords in that they are shorter and shallower.

Δ **Depositional land forms – Glacial forms** : Glacial deposits, in contrast to those of glacially fed streams and glacial lakes are marked by heterogeneity of materials and lack of stratification. Three types of glacial deposits – end moraine, lateral moraine and ground moraine – may be distinguished, depending upon whether deposition took place at the end of, at the side of or beneath an ice stream. The term recessional moraine has been used in the past to designate end moraines back of the outermost one, the implication being that a series of end moraines marks successive pauses in the position of a retreating ice front. Actually, some may be retreatal moraines and others may mark readvances. Determining which interpretation is correct is one of the difficult problems encountered in working out the history of glaciated valley.

Slight oscillations of an ice front as it recedes may result in an irregular belt of knolls and basins, usually described as *knob and basin topography*. This type of end moraine is more commonly formed by ice caps than by ice streams.

Lateral moraines form along the sides of an ice stream chiefly from materials which are contributed from the valley sides above the glacier by weathering, snowslides, avalanches and other types of mass movement. Two lateral moraines may join to form a medial moraine, but a medial moraine is more a feature of glacier's surface than a land form, because it will not persist long after a stream occupies the trough.

Δ **Glacio-fluviatile forms** : Much of the material acquired and transported by an ice stream is ultimately got hold of, transported, and finally deposited by streams flowing on, within, beneath, and beyond the glacier. Deposits of such origin are classed as *glacio-fluviatile*. They may retain some of the characteristics of glacial debris, but they show a degree of assortment and stratification roughly proportional to the distance that they were carried by streams. The most common land forms belonging to this class found in areas of mountain glaciation are valley trains, eskers, kame terraces kames and outwash fans or deltas.

Valley trains consist mainly of outwash sand and gravel heading usually at an end moraine and extending downvalley from it. Except for those still in the process of formation, they are likely to be marked by terraces above present valley floors. Study of the valley trains extending down from the Alps led Penck and Bruckner to a recognition of four stages of glaciation.

Eskers are sinuous ridges of assorted and somewhat stratified sand and gravel which are believed to represent fillings of superglacial, englacial or subglacial stream channels. They are far more common in areas of continental glaciation than in glacial troughs, for eskers are features formed when ice stagnates and stagnant ice is rare in valley glaciers. They might be confused with lateral moraines, for any persisting in a glacial trough are likely to be found at or near its sides. In contrast to eskers, however, lateral moraines consist of unassorted materials.

Kame terraces may form in glacial troughs, but not too commonly. They, along with eskers are kames, belong to a class known as ice-contact features, so-called because the materials of which they are composed were laid down against an ice surface. Kame terraces are fillings or partial fillings of depressions between a glacier and the sides of its trough. These depressions are called *foxes*, and they owe their existence to the more rapid rate of melting which takes place here because of the added effect of heat absorbed or reflected from the valley sides.

Δ **Glacio-lacustrine features** : Lakes, present or extinct, are common features of glaciated valleys. The more common types are : rock-basin lakes in cirques and upon the treads of glacial steps; lakes back of a terminal and lateral moraines; lakes formed by the damming of tributary valleys by a valley train in a main valley; and those in the basins of knob and basin topography. Lakes are ephemeral and forms and their sites soon become *lacustrine plains* by filling with in wash or lowering of their outlets by erosion. The small lacustrine plains which are often found in glaciated mountain valleys are often striking because their flatness contrasts sharply with their surroundings.

1.3.4 EVOLUTION OF LANDFORMS BY THE AEOLIAN PROCESS :

Δ **Topographic effects of wind erosion** : Wind erosion manifests itself in three forms : abrasion or corrosion, the natural sandblast action of wind-blown sand; deflation, the lifting and removal of loose material from the earth's

surface; and attrition, the mutual wear of particles carried along by the wind. It was formerly common practice to attribute great erosional power to wind and to explain most desert land forms in terms of wind erosion, but it has become increasingly apparent that few major topographic features of arid regions are so formed.

Wind abrasion may aid in the shaping of some of the details of major forms but is itself hardly capable of producing features of great areal extent. Possibly the only topographic form that can be attributed unequivocally to wind abrasion is the *yardang*.

Wind abrasion can operate only near the ground because of the inability of wind to lift sand more than a few feet. Bagnold (1941) stated that sandblast action on posts is seldom appreciable above a height of 18 inches and concluded that sand is seldom lifted more than 2 metres above the ground. According to Blackwelder (1928) wind abrasion manifests itself through (a) polishing and pitting, (b) grooving, and (c) shaping and faceting. He believed that effects of the first are rather widespread throughout arid regions, but that they are of little geologic significance. It is important, however, that we recognize that polishing is a natural effect of abrasion or sandblast action, for many etched boulders and surfaces have been attributed to wind abrasion that lack polishing and are more likely the result of differential weathering than of eolian erosion. Wind abrasion may produce *ventifacts* which exhibit one or more polished and faceted surfaces, but these are relatively rare because strong winds, abundance of sand and absence of vegetation are essential to their best development.

Certain topographic forms commonly called pedestal rocks, which consist of residual masses of weak rock capped with harder rock have been explained by wind abrasion. It is doubtful whether this process has contributed significantly to their shaping.

That deflation is responsible for the formation of many depressions called *blowouts* is rather generally agreed. These are common in areas of sand accumulation where they form small basins on or within dunes and other types of sand accumulation. Blowouts may also develop in areas where nonindurated or poorly indurated materials lie beneath the land surface.

Δ **Lag deposits** : In the process of removal of sand and smaller-sized particles by deflation there is a sorting of materials according to size with the coarser materials being left behind. These concentrations of pebbles and boulders have been designated by the general name of *lag deposits*. Locally they are striking features of deserts. *Desert pavement* and *desert armor* are terms often applied to them. Such surfaces characterize the *regs* or stony desert portions of the Sahara. The areas covered with large-sized rocks are called *hamadas*; those areas covered with small rock and gravel are known as *serirs*. They are in striking contrast to the areas of sand accumulation known as *ergs* (Gautier, 1935).

Δ **Eolian deposits** : Materials undergoing transport by the wind are subjected to a winnowing which gives rise to two distinct types of wind deposits, accumulations of sand and deposits of silt and clay known as loess.

Δ **Environments in which wind deposits are found** : Wind deposits are by no means restricted either in origin or location to deserts. Not only is loess found in humid or subhumid regions but eolian sand deposits are common in several nondesert environments. The four nondesert environments in which sand deposits may be found are (a) along shore lines, (b) along stream courses in semiarid regions, (c) in areas where loosely cemented

sandstones have disintegrated to supply sand, and (d) in areas of glacial outwash. Along most shorelines, either ocean or lake, unless the shore be rocky and rugged, dunes sands are likely to be found.

Δ **Types of sand deposits** : Several classifications of sand deposits have been proposed, but none is completely satisfactory. Bagnold's classification is probably most applicable to areas of freely moving sand and was based upon extended observations in the deserts of Africa and Asia as well as laboratory experiments in wind tunnels designed to establish the mechanics of wind transportation and deposition. It seems, however, to be somewhat inadequate for areas where vegetation has played an important role in the shaping of sand deposits. He recognized the following types of eolian sand deposits :

- A) Small-scale forms : small ripples and ridges on sand surfaces which have little geomorphic significance.
- B) Large-scale forms
 - 1) Sand shadows and sand drifts
 - 2) True dunes
 - a) The barchan or crescentic dune
 - b) The seif or longitudinal dune
 - 3) Whalebacks or sand levees
 - 4) Undulations
 - 5) Sand sheets

A *sand shadow* is an accumulation of sand to the lee of and in the shelter of an obstruction, such as a boulder, bush or cliff, which interferes with stream-line air flow and checks the wind's velocity. Sand accumulates to the lee of an obstruction until its front slope has attained the limiting angle of repose, which is about 34 degrees. Sand will then slide down the front of the advancing deposit along what is called its *slip face*. The height of a sand shadow is limited by the size of the ground plan of the wind shadow in which it forms, for any sand which gets outside this shadow will be swept into the main sand stream by the stronger wind outside it. Sand shadows may also form where wind sweeps sand over a cliff or escarpment. Such deposits have also been called *sandfalls*.

Sand drifts are found to the lee of a gap between two obstructions. This gap acts as funnel through which sand trails out to the leeward. This usage by Bagnold of the term sand drift may be confused with the more common application of it to thin deposits of sand that are called by him sand sheets.

Bagnold (1933) defined a *dune* as a "mobile heap of sand whose existence is independent of either ground form or fixed wind obstruction". Groups of dunes are often referred to as *dune complexes*, *dune colonies* or *dune chains*. Unlike sand shadows and drifts, their existence is not dependent upon an obstruction or topographic break. They usually attain their maximum development upon relatively flat terrain.

Bagnold considered the barchan and the seif as the only true dune forms; they, according to him, rarely exist together. Dunes are most typically found in deserts, for a long coasts and rivers moisture and vegetation usually interfere with their growth and prevent their orderly development as well as limit their size. A *barchan* is a crescentic-

shaped dune with tips extending to the leeward, making this side concave in plan and the windward side convex. The slip face, if present, is transverse to the wind. Barchans tend to arrange themselves in chains extending in the direction of the most effective winds. Changes of wind directions may modify their form and even temporarily cause a reversal in positions of the slip faces, or destroy them with resultant development of conical-shaped sand masses. In areas of abundant sand supply, barchans are large and closely packed, whereas in areas of lesser sand supply they are smaller and more dispersed.

A seif or longitudinal dune instead of being transverse to the prevailing wind is parallel to it. Many seifs in the Egyptian Sand Sea attain heights of 100 metres, and some in Iran are as much as 210 metres high. According to Bagnold, the width of a seif is roughly 6 times its height. Seif chains have been described which are 300 kilometres long. Their crests form knife-like ridges with many peaks and sags. One side of the crest may be rounded and the other shows a collapsing front (slip face) at right angles to the prevalent wind direction. There may be gaps or corridors between adjacent seifs in which bare desert floor is exposed. Contrary to the more commonly held idea that longitudinal dunes or seifs are built mainly by winds blowing sand rather constantly from one direction, Bagnold maintained that they grow in height and width largely through the action of cross winds and grow in length during periods when the prevailing wind is parallel to the trend of the seif chain. He believed that the seif is a modification of the fundamental barchan form produced by strong cross winds transverse to the prevailing wind direction and that their knife-like crests represent barchan forms superposed upon a ridge of sand extending parallel to the prevailing wind.

Whalebacks or *sand levees* are flat-topped sand ridges which extend parallel to the prevailing winds but lack the collapsing fronts which mark seifs. They also have much larger dimensions. A whaleback may be 100 miles long, 2 miles wide and 150 feet high. Locally there may be seif forms upon their tops. They seem to be confined largely to the Egyptian Sand Sea. Bagnold suggested that they may constitute the residue from the march of a dune chain or series of dune chains downwind. *Undulations* are sand deposits somewhat similar in nature to whalebacks but shorter in length and lacking the definite form of the whaleback. They form billowy surfaces between whalebacks or seifs. Both whalebacks and undulations appear to be areas of "dead" or nonmobile sand in contrast to the mobile barchans and seifs.

Bagnold applied the name *sand sheet* (more commonly called sand drift) to a sand area marked by an extremely flat surface and absence of any topographic relief other than small ripples.

Hack (1941) proposed an empirical classification of dunes which he thought fitted the Navajo country of Arizona. He recognized three types of dunes, transverse, parabolic, and longitudinal. *Transverse dunes*, including the common barchan, are nearly always free of vegetation; their form seemed to him not to be influenced by vegetation. The tips of a transverse dune extend to the leeward. *Parabolic dunes* were defined as "long, scoop-shaped hollows, or parabolas of sand with points tapering to windward" and with a windward slope much more gentle than the leeward in contrast to the transverse dune. He thought that these dunes were formed where the wind removed sand from windward hollows and deposited it on leeward slopes. This type of dune seemed to be associated always with a cover of vegetation. It would appear that they are not original dune forms but rather forms produced by sand blowouts and subsequent redeposition. Hack described them as "parabolic dunes of deflation".

Δ **Loess** : This name is applied to wind-blown silts which commonly are buff-colored, nonindurated, calcareous, permeable, particularly in the vertical direction and consist of angular to subangular particles of quartz, feldspar, calcite, dolomite and other minerals held together with a montmorillonite binder. Unweathered loess is usually gray in color, but because of its permeability exposures of unweathered loess are uncommon.

Not all loess, however, was derived from glacial outwash. In some parts of the world loess deposits exist which appear to have been derived from arid regions by means of deflation and accompanying exportation. This type of loess has been called *desert loess*. The extensive loess deposits of northern China have been attributed to this origin. Some of early descriptions of the loess of China ascribed thicknesses of 1000 to 1200 feet to the loess here, but more recent studies indicate that the maximum thickness is more likely of the order of 200 feet (Obruchev, 1964). Recent work here also suggests that not all of the material which has been called loess is eolian origin. In view of the fact that few of the world's great deserts are bordered by thick loess deposits, some skepticism may exist as to whether the Gobi Desert was the main source of the China loess. Possibly the floodplains of such rivers as the Yangtze and Hwang may have served as source areas for some of the material just as today the floodplain of the Yukon River in Alaska is proving to be an important source of modern loess.

□ 1.3.5 EVOLUTION OF LANDFORMS BY THE COASTAL PROCESS :

Δ **Movements of water in oceans and lakes** : Movement of water goes on at nearly all times both at the surface and at depth. We are not concerned here with large-scale circulations whereby there is interchange of water between the surface and depths or slow latitudinal movement of oceanic waters. The three types of movement that carry on gradational work are waves, currents and tides and only rarely are tides of geomorphic significance.

Wave height is determined not only by wind velocity but also by the extent of water over which the wind can blow. This is called the *fetch*. Extremely large waves can develop only where the fetch is great, which explains why such waves cannot develop on lakes or enclosed arms of the sea.

Another factor, besides wind velocity and fetch that affects wave height is the duration of the wind.

As waves move into shallow water there is a decrease in velocity with a consequent crowding together of the waves and steepening of their fronts. Where the orbital velocity in the crest of the wave exceeds the rate at which the wave is moving forward there develops a curl in the top of the wave front which, lacking sufficient water to fill the cavity beneath it, lunges forward or breaks to form *surf*. In breaking, waves of oscillation change to waves of translation, which dash forward onto the shore as *swash*. After the swash reaches its maximum forward position the water runs back down the seaward slope as *backwash*.

Tsunamis are waves that are produced by submarine earthquakes, volcanic eruptions, or land sliding and slumping. They are often called tidal waves but are in no sense related to tides. They have wave lengths of the order of 100 miles and travel at velocities of 400 miles or more per hour. In open sea, their heights may be only a foot or two but when they come onto shore they may develop into waves having extraordinary heights.

The highest waves ever formed are those known as landslide surges. These develop in more or less restricted bodies of water when a mass of rock breaks off an adjacent slope and plunges precipitously into the water. The largest of these on record occurred on July 9, 1958 in Lituya Bay, Alaska, (Miller, 1961) when a mass of rock

estimated at 40,000,000 cubic yards plunged from an altitude of 3,000 feet into Lituya Bay.

Currents differ from waves in that there is continued and progressive forward movement of water. Currents may be created in several ways, but we shall discuss only those that have geomorphic significance.

Winds are responsible directly or indirectly for the major surface ocean currents; hence they have a definite relation in direction to planetary wind belts, modified, however, by the deflective effect of the earth's rotation and the configuration of the continents. Where currents impinge upon land they are forced to deviate in direction and hence the directions of some currents are determined to a large degree by the outlines of the continents. This is one cause of *longshore* or *littoral currents*, a type of current which may have important geomorphic effects. Longshore currents are also generated when waves break obliquely against a shore. Wind drift currents seldom affect water to depths in excess of 300 feet. Their velocities are rarely great enough to produce significant erosive action, but they can move material along the sea bottom and thereby contribute directly or indirectly to the development of depositional features along shorelines. *Longshore drifting* as effected by currents should not be confused with a different type of longshore transportation called by Johnson (1919) *beach drifting*.

Tidal currents locally may attain velocities sufficient to transport material, but generally they are of minor geomorphic significance.

It generally is assumed that the water which is carried landward by waves returns seaward mainly by a bottom flow called *undertow*. Some (Shepard, Emery and LaFond, 1941) have claimed that trip currents are more important in returning water seaward than undertow. *Rip currents* consist of localized lanes or streaks of water moving seaward in contrast to the no localized flow of undertow. Rip tide has also been applied to them, but it is a misnomer, for they are in no way related to tides. The presence of rip currents along a coast is indicated by such features as (a) brown streaks of sediment-laden water, (b) streaks of green water resulting from the greater depth of water where a rip current forms, (c) streaks of agitated water, (d) gaps in lines of breakers, (e) foam belts, especially where the bottom is rocky, and (f) seaward movement of floating objects. Shallow channels on the sea floor often attest to their presence. It is likely that rip currents represent a more significant method of carrying seaward suspended sediments than is generally realized.

Δ **Marine erosion** : Waves, particularly storm and tsunamis, are the most important agents of marine erosion. Smaller waves, such as those associated with surf, may carry on attrition of material and minor amounts of abrasion, but, just as a stream during a single flood may do more geologic work than it will for months or years at low-water stage, so storm waves during a short period may effect more change than ordinary waves will in months. With minor local exceptions, as in constricted passageways, currents are probably relatively unimportant erosional agents; but they are an important means of transportation.

Δ **Conditioning factors** : Aside from variations in wave strength, there are several factors that influence the rate at which marine erosion proceeds. Among the more important controlling factors are :

- a) The kind and durability of the rocks along shore;
- b) Structural features of the rocks, particularly its attitude and degree of jointing and fracturing.
- c) Tidal range.

- d) Openness of shore to wave attack.
- e) Depth of water offshore.
- f) Configuration of coastline.
- g) Presence or absence of a protective beach.
- h) Abundance and size of abrasive tools.
- i) Stability of sea level.

Δ **Processes of marine erosion** : At least five processes contribute significantly to marine erosion : corrosion, corrosion or abrasion, attrition, hydraulic action and the shock pressure of breaking waves. Marine erosion is most effective where abundant tools are available but wave impact itself, particularly upon highly jointed and fractured rocks, may exert great pressure.

The most effective process in marine erosion is the corrosive or abrasive action of sand, gravel and pebbles moved by waves against the shore. This includes what has been called the "artillery action" of tools hurled against solid rock and the more common but less spectacular corrosive action as rock particles are moved back and forth over bedrock. In the absence of tools, marine abrasion may be ineffective, as is shown by the preservation of glacial striae along some coasts. Attrition contributes indirectly to marine abrasion by reducing rock particles to sizes that can be carried seaward by undertow and rip currents.

Δ **Topographic features resulting from marine erosion** : In many places the seaward limit of a coast is marked by a nick or scarp resulting from wave erosion and commonly called a *sea cliff*. Usually extending seaward from the base of the sea cliff is a *wave-cut bench*. This bench may be bare rock or it may be covered with impermanent deposits of sand, gravel and pebbles called a *beach*. Elevated wave-cut benches marking former high sea levels are usually called *marine terraces*. Sea cliffs and wave-cut benches are the two most common features created by marine erosion. Details of sea cliffs vary greatly, depending upon the kind structure and attitude of the rocks in which they are being cut. The configuration of a sea cliff cut in rocks dipping seaward will be different from one in rocks that dip landward, and still different from one cut in horizontally-bedded rocks; cliffs cut in granite differ in appearance from those cut in basalts; and cliffs in easily erosible material such as glacial till or nonindurated Tertiary deposits will be bold and marked by much slumping and landsliding.

Above the base of a sea cliff there may be found a bench or high-water platform which owes its existence largely to the spray or splash of storm waves, aided by a subaerial weathering and rainwash. Whether a high-water platform is present seems to depend upon whether the processes of high-water leveling keep ahead of the more general process of marine terracing in the zone of breakers. Gotton (1945) has pointed out that optimum conditions for high-water platform development are attained in enclosed bays where normal wave action is not too effective, but these platforms are also found along open coasts, and so it seems that they are normal features of shore lines. Care should be exercised not to interpret high-water platforms as marine terraces exposed by recent shoreline emergence.

A shoreline which is being extended landward by wave attack is said to be *retrograding*. During

retrogradation, recession of the sea cliff will proceed at different rates along a coast, depending upon the durability of the rocks and the openness of the shore to wave attack. Thus the lesser indentations of a coast may result from differential marine erosion. These are known as *coves*, *bays* or *bights*. It is doubtful, however, if wave attack is ever responsible for large bays and estuaries. These have resulted from submergence of topography produced in other ways. *Head-lands* will be left in areas of more resistant rocks. Headlands subjected to wave attack from two sides may have *sea arches* or *caves* cut in them. Portions of headland detached from the shoreline are known by such names as *stacks*, *chimneys*, *skerries* or *islands*.

Retrogradation of a shoreline may go on so rapidly that small streams are unable to keep pace in down cutting with the rate of sea-cliff recession. As a result, these streams enter the sea from hanging valleys. These are common along the chalk cliffs on either side of the Strait of Dover.

A shoreline that is retrograding across a veneer of sedimentary materials over a under mass of older rocks may exhume or resurrect part of the buried topography. The shoreline is then said to be *contraposed* (Clapp, 1913). Its characteristics may undergo notable change when waves encounter the resistant undermass. Contraposition of a shoreline is analogous to superposition of valleys.

Δ **Topographic features resulting from marine deposition – Beaches :** The term beach should be restricted to the temporary veneer of rock debris which accumulates along and on a wave-cut bench.

Beaches may extend for hundreds of miles along a shore, as along the southeast coast of the United States, or they may be patchy. Along rugged shorelines, beaches are mainly limited to strips at the heads of bays or coves. Their greater number in these locations is the result of the fact that waves converge upon headlands and diverge in bays, causing an intensification of wave erosion on headlands and a lessening of its in bays. Longshore currents and beach drifting also tend to move sand into the bays to either side of a headland.

Movement of beach material goes on at all times, but reduction in beach thickness is most significant during periods of extreme storm waves, whereas beach accretion takes place during periods of relatively quiet water.

The materials of a beach come both from land and sea.

Δ **Bars :** The term *bar* may be used in a generic sense to include the various types of submerged or emergent embankments of sand and gravel built on the sea floor by waves and currents; they may be emergent at low tide. Specific names are given to individual forms depending upon their form and position. One of the most common types of bar is that known as a *spit*. It was defined by Evans (1942) as "a ridge or embankment of sediment attached to the land at one end and terminating in open water at the other." Most commonly the axis of a spit will extend in a straight line parallel to the coast, but, where currents are deflected landward or unusually strong tides exist, growth of a spit may be deflected landward, with the resulting creation of a *recurved spit* or *hook*. Several stages of hook development may produce a *compound recurved spit* or *compound hook*. Spits may be attached at both ends to headlands to form what is called a *winged headland*. Convergence of two spits offshore or recurving of a simple or compound spit until it becomes attached to the shore at both ends produces a *cusped spit*. Growth of successive cusped spits farther and farther seaward will cause prograding of the shoreline, and the seaward extension thus formed is known as a *cusped foreland*. Cape Henry, Virginia and Cape Kennedy (Canaveral).

Florida are examples of such features.

Δ **Barrier beaches and barrier chains :** Following Price (1951) and Shepard (1952) what was long called an offshore bar is here termed a *barrier beach* or *barrier island*. The term bar seemed inappropriate to them because this name was used originally to designate slightly submerged sand ridges. Barrier beaches consist of elongate sand ridges, generally parallel to the shore, which rise slightly above high tide. They are separated from the coast by a lagoon. A series of these features extending for a considerable distance is known as a *barrier chain*.

1.3.6 SEA LEVEL CHANGES AND RAISED BEACH FEATURES :

Sea level change is one of the main factors in stimulating coastal change and in the long term it has been controlled by changes in the volume of landlocked ice sheets. If sea level rises, then coasts must be able to change dynamically to keep up, or down. On the scale of a human lifetime we may consider sea level to be unchanging, but even into the twenty-first century it now appears that sea level has been rising due to human activity and may continue to rise well into the future.

Δ **Introduction**

The level of the sea is not constant, it is always rising and falling, either through the passing of waves (including tides), meteorological influences or gravitational effects in the form of the earth's *geoid*. Long-term and significant sea level changes, however, reflect changing levels of both land and sea; *eustasy* refers to absolute changes in global sea level, and *isostasy* refers to the vertical movement of land due to local geological factors. It is the balance between these two processes at a given coastline that produces observed changes in sea level, referred to as *relative sea level change*, because an absolute (eustatic) rise in sea level may not be required to allow the sea to rise relative to the land. Explores this balance in a simplified way, and a number of cases can be put forward to illustrate the relationship :

- Relative sea level rise will occur if (a) eustatic sea level rises whilst the land is isostatically subsiding, static or uplifting at a slower rate than eustatic rise; (b) eustatic sea level is static whilst the land is isostatically subsiding; and (c) eustatic sea level falls whilst the land is isostatically subsiding at a faster rate than eustatic fall.
- Relative sea level fall will occur if (a) eustatic sea level falls whilst the land is uplifting, static or subsiding at a slower rate than eustatic fall; (b) eustatic sea level is static whilst the land is isostatically uplifting; and (c) eustatic sea level rises whilst the land is isostatically uplifting at a faster rate than eustatic rise.
- A sea level stillstand will occur (a) eustatic sea level fall and isostatic subsidence of the land occur at the same rate; (b) eustatic sea level and the isostatic land level are both static; and (c) eustatic sea level rise and isostatic uplift of the land occur at the same rate.

The main reason for eustatic and isostatic changes in the level of the sea and land is the growth and decay of continental ice sheets during the Quaternary. Glacierization requires ice, which is introduced into glacial systems as snow, mainly derived from oceanic evaporation. Therefore, during Quaternary glacial stages water is extracted from the oceans and transferred to ice sheets where it is locked up for the duration of the glacial stage. This

glacial control on eustasy is known as *glacio-eustasy*. Similarly, these same continental ice sheets growing on land exert great overburden pressure, causing the land to isostatically subside, a process known as *glacioisostasy*. At the end of a glacial stage, melting ice returns water to the ocean causing sea levels to rise, and at the same time, the release of ice overburden results in the isostatic rebound or uplift of the land. Therefore, previously glaciated areas often possess very complex relative sea level histories, reflecting the balance between glacio-eustatic and glacio-isostatic changes (e.g. Gehrels *et al.* 1996; Dawson and Smith 1998). Furthermore, knowledge of these processes has led to the development of models that attempt to predict evolving coastlines since the last glacial through the construction of shoreline maps for different time periods (e.g. Lambeck 1995).

The Quaternary Period spans from 1.81 million years ago to the present day and comprises the *Pleistocene Epoch* (1.81 million years to 10,000 years ago) and the *Holocene Epoch* (10,000 years ago to the present). During the Quaternary sea levels fluctuated widely, with sea level high stands during the interglacials and lowstands during the glacial stages. Many interglacial sea level highstands attained or exceeded the present sea level altitude, and many fossil shorelines may be encountered above present sea level and appear to be 'raised', hence the term '*raised*' beach. The word 'raised' can be a misnomer, because many of these fossil Pleistocene shorelines are in situ.

In the British Isles, mapping isostatic readjustments following the last glacial at the end of the Pleistocene shows that northern Britain is currently uplifting by up to 2 mm y^{-1} , whilst southern Britain is subsiding by up to 2 mm y^{-1} . The two regions are separated by a fulcrum line that extends in a stage zone from the Tees Estuary in northern England, through the Dee Estuary and on to the Lley Peninsula in northwest Wales. This post-glacial isostatic movement is reflected in the distribution of *submerged forest* sites in Great Britain. The majority of sites are located in southern Britain, whilst only a few are found north of the fulcrum line. This indicates that post-glacial eustatic rise coupled with isostatic subsidence in southern Britain has led to relative sea level rise and the drowning of previously wooded landscapes, whilst in northern Britain isostatic rebound is outpacing eustatic rise, so that relative sea levels are actually falling there are resulting in coasts emergence rather than submergence.

Only in isostatically stable areas of the world can records of eustatic or absolute sea level change be possible found i.e. areas that have never been glaciated or suffer earthquakes, volcanic or other tectonic activity (earth movements) that may vertically displace the land. Such areas are rare, but the island of Bermuda is considered to be a stable mid-ocean carbonate platform from which a eustatic sea level record has been constructed (Lowe and Walker 1997). Eustatic sea level curves from such areas indicate that the last glacial sea level 18,000 years ago was approximately 130 m below the present sea level. A period of very rapid rise ensued until approximately 6,000 years ago, but since then sea level has been comparatively stable with a less than 10 m rise to the present day. Sea-level from 18,000 years ago to the present and how the coastline of Great Britain has developed.

Δ Future sea level rise and its impacts

The results of a number of model experiments reported by Warrick *et al.* (1996) indicate that a sea level rise of 49 cm by 2100 is possible, but with a range of uncertainty of 20-86 cm and represents a rate of sea level rise that is significantly faster than that experienced over the majority of the twentieth century. These projections of future sea level rise are presented under a number of scenarios that use different levels of greenhouse gas

emissions. The scenarios are labeled IS92a-f, and IS92a is equivalent to the 'business-as-usual' scenario of Houghton et. al. (1990), which is a projection using the present greenhouse gas emission levels. The other scenarios are for a number of lower or higher emission cases. For each scenario, a high, low and best estimate is calculated, each associated with a global atmospheric temperature of 4.5°C, 1.5°C and 2.5°C respectively.

The various contributors to sea level change in the period 1990-2100 have been assessed individually. A thermal expansion estimate of 28 cm has been calculated using the computer model of Wigley and Raper (1993). A 16 cm contribution from mountain glaciers and small ice caps was calculated in relation to changes in glacier volume in response to global warming; the warmer the climate, the greater degree of melting, leading to a decrease in glacier volume. The Greenland ice sheet is estimated to contribute 6 cm, whilst the Antarctic ice sheet is likely to contribute negatively by -1 cm. In summary, under the IS92a scenario, the best estimate for the 2100 is that global sea level would have risen by 49 cm, although because of the range of uncertainty it could be as little as 20 cm or as much as 86 cm.

The projections also show that throughout the twenty-first century, even under the scenarios of lower greenhouse gas emissions, sea level continues to rise. This suggests that even if greenhouse gas emissions were cut completely now, sea level would continue to rise. This is mainly because of a lag effect caused by the ability of the oceans to store heat, in what is called its thermal inertia. This means that society is committed to sea level rise. However, global sea level rise will have different impacts on different coastlines, mainly due to isostatic movements, but there is considerable concern for the majority of the world's coasts. Concern is not centred on the day-to-day increase in tidal range, which can in most cases be contained, but in the increased ability of extreme events, such as storm surges and very high tides, to flood coastal lowlands and cause extensive coastal erosion. Deep storm depressions are predicted to become more frequent under global warming, so that increased sea levels, coupled with surge conditions and high precipitation in coastal lowland river catchments, are inevitably going to mean higher frequency and higher magnitude flood events.

Δ Socio-economic impacts of future sea level rise

Bijlsma (1996) on behalf of the Intergovernmental Panel on Climate Change outlines the socio-economic implications of predicted sea level rise due to an enhanced Greenhouse Effect. These include:

- General negative impacts on various coastal amenities and activities, such as tourism, freshwater supply and quality, fisheries and aquaculture, coastal agriculture, human settlements, financial services and human health.
- Protecting low-lying islands, such as the Maldives and Marshall Islands and countries with extensive low-lying deltaic plains, such as Bangladesh, Nigeria, Egypt and China, will probably be extremely costly.
- The numbers of people globally who currently suffer from coastal flooding (40 million) could double or triple by 2100.
- Attempting to adapt to sea level rise may mean that some societies will face difficult decisions regarding which coasts to protect and which to abandon, and trade-offs between environmental, economic, social and cultural values will have to be made.

Furthermore, Houghton (1994) indicates that for the United States alone, the financial cost of coastal protection

and land loss will amount to US\$7 thousand million, which doesn't include the cost from seawater contamination of water supplies, agricultural losses and health problems that may be caused by sea level rise. This figure is approximately 1 per cent of the US Gross Domestic Product (GDP), and a similar percentage is predicted for other developed countries, whilst for some developing countries costs could be as much as 6 per cent of GDP.

1.3.7 SELF ASSESSMENT QUESTIONS :

- a) Discuss the evolution of land forms resulted by fluvial erosion.
- b) Discuss the forms, processes and evolution of deltaic coast.
- c) Write short note on : Interruptions of the fluvial cycle; sea level changes; Raised Beach features.

1.3.8 REFERENCES :

- W. D. Thornbury (1984). *Principles of Geomorphology*. John Wiley and Sons, Inc pp. 594
- S. W. Wooldridge and R. S. Morgan (1959). *An outline of Geomorphology*. Longman, Green and Co. Ltd. London. pp. 409
- Simon K. Haslett (2000). *Coastal systems*. Routledge, London pp. 218

Part – I
PAPER-I :: Module – I :: Unit-04

Geography and Environment Management.

Module Structure

- 1.4.0 Lithology (Rocks) and Landforms
- 1.4.1 Processes of Weathering, Mass wasting and erosion and resultant Landform.
 - 1.4.1.1 Introduction and Definition
 - 1.4.1.2 Importance of weathering in Geomorphology
 - 1.4.2.3 Factors of weathering
 - 1.4.1.4 Mechanical weathering-stresses and processes
 - 1.4.1.4.1 Different stresses for mechanical weathering
 - Geological stress
 - Geomorphological stress
 - Thermal stress
 - Hydrological stress
 - Biological stress
- 1.4.1.2 Processes of Mechanical Weathering
 - Insolation Weathering
 - Insolation cracks
 - Block disintegrations
 - Granular disintegration
 - Boulder cleaving
 - Unloading Weathering
 - Sheeting
 - Spalling
 - Crystal Weathering
 - Ice Crystal Weathering
 - Salt Crystal Weathering

■ Moisture Weathering

■ 1.4.1.5 Chemical weathering

■ 1.4.1.5.1 Factors:

- Ph
- Eh (Redox or Oxidation Potential)
- Ionic Potential

■ 1.4.1.5.2 Processes of Chemical weathering

- Solution
- Carbonation
- Hydrolysis
- Hydration
- Oxidation
- Reduction
- Chelation

■ 1.4.1.5.3 Weathering Landforms

A. Constant Volume Weathering and landform

- Regolith and saprolite
- Desert varnish
- Patina
- Rounded landforms
- Color banding
- Karst landforms

B. Weathering with expansion and related Landforms

- Unloading dome
- Unloading sheet
- Scree

C. Landforms due to differential weathering

- Weathering pit
- Tor
- Inselberg
- Etchplain
- Weathering residues

■ 1.4.3 Mass Wasting and Landforms

A. Factors responsible for decrease in Shear Strength

B. Factors responsible for increase in Shear Stress

A. Factors responsible for decrease in Shear Strength

- Weathering
- Seepage Erosion
- Liquifaction
- Remoulding
- Reduction in angle of internal friction
- Progressive failure

B. Factors responsible for increase in Shear Stress

- Increase in the length and height of slope
- Unloading
- Loading
- Shocks and vibration
- Removal of underlying support

■ 1.4.3.1 CLASSIFICATION OF MASS WASTING

■ 1.4.3.2 Landforms developed by Mass Movement

Dislodgement of debris

- Free face
- Land slide scar
 - Circular scar
 - Planner scar
 - Wedge like scar
 - Toppling scar

Transportation of dislodged materials on slope

Deposition of transported materials at foothill

- Colluvial cone
- Scree or talus
- Solifluction lobe

■ 1.4.4 Slope Development and Slope facets

Components of a slope

Process of slope formation

Theories of slope development

View of Penck

Theory of Parallel retreat

View of Wood (1942)

View of Davis – Slope decline theory

View of L.C.King

view of Savigear

■ 1.4.5 Concept of Grade, Profile of Equilibrium and Base Level

■ 1.4.5.1 Grade;

■ 1.4.5.2 Profile of Equilibrium

- Concept of Equilibrium
 - Decay Equilibrium
 - Steady State Equilibrium
 - Dynamic Equilibrium
 - Dynamic Metastable Equilibrium
- Equilibrium Profile

■ 1.4.5.3 Base level of erosion:

- Types of base level;
 - Local Base level
 - Temporary Base level
 - Grand Base level

■ 1.4.6 Self Assessment Questions

■ 1.4.7 References

■ 1.4.0. Lithology(Rocks) and Land form

Geomorphologists are equivocal about the role of rock composition and rock structure on the development and evolution of landform. Both the historical and functional approaches in land form study accept this lithological control on landform evolution. The proponent of historical approach, W.M. Davis proposed that "landform is a function of structure, process and stage". Here 'structure' refers not only to the rock orientation but also the composition and character of rocks. The functional approach, on the other hand, advocates the study of landform as a function of number of factors and the change in landform occurs as any of these factors change. The lithology thus accepted as the most important factor in landform development and evolution. The pioneer of this functional approach, J.T. Hack, in his Dynamic Equilibrium Concept proposed that the evolution of landform and the difference in landform from one part of the globe to other is not due to the effect of "stage" or time but due to the effect of lithological variation. He cited one example in this connection. A land if is constructed of alternate layers of sandstone, shale/mudstone and again sandstone, the river will develop a narrow and steep valley on sandstone at first. After the erosion of sandstone, the river will develop a wide valley on softer shale and again due to denudation of the shale, if the river is imposed on sandstone, the narrow and steep valley is again developed. So the shape of valley does not depend on the 'stage' but on the rock or lithology.

In one of the fundamental concept in geomorphology (concept no-2, following W.D Thornbury, 1985) it is clear that "Geologic structure is a dominant control factor in the evolution of landforms and is reflected upon them". Here 'geologic structure', in broad sense, refers not only to the rock orientation but also the composition and character of rocks.

The character of rocks how affect the landform evolution is discussed below with examples.

The **resistance** is an important determinant of effectiveness of erosional agents. The soft rocks like shale, limestone (in humid area) are easily eroded to form lowlands, basins, valleys etc. where as the hard rocks resist erosion or greater intensity of energy is required for effective erosion of hard rocks. The resistant rocks, by resisting erosion, develop highlands like mounds, ridges, inselberge, bornhardt, tor etc.

The **chemical composition of the rocks** is another factor to determine the character of rocks as well as development of unique landform. The acidic lava being viscous, readily solidifies and does not flow farther from the zone of eruption and thus develops volcanic cone. The granite, a rock of this kind, develops resistant batholith and other plutons. The cumuldome, tholoids, spine etc. are developed from acidic volcanic rocks. The exfoliation as a result of horizontal fracture in granite develops round shaped landform.

The basic lava, on the other hand, can flow to a long distance due to more liquidity forms shield like landform after solidification.

The **permeability** can exert immense influence on landform development. The selective permeability through fractures can help in the development of sink holes at the junctions of the fractures and thus karst topography develops. The mass permeability of the limestone and other soluble rocks does not help in the development of Karst. The conglomerate and sandstone are considered as resistant as most of the runoff infiltrates down through the pores and thus availability of surface runoff becomes less and so intensity of erosion decreases. The shale, on the other hand, is eroded intensively due to less infiltration and more runoff and so most of the badlands with intensive development of rills and gullies are developed on shale or mudstone.

Presence of Bedding plane and its surface exposure helps in intensive erosion and step like topography is developed as a result of this. The erosion on sedimentary rocks are more than those on igneous and metamorphic rocks.

The **solubility** of rock beds determines the intensity of erosion as solution is the most important process of denudation. The limestone, dolomite and other soluble rocks are easily eroded and depressions are formed on these rocks.

	Rocks	Characters of rocks	Developed Landforms
	Granite	Resistant, hard, massive, acidic	Batholith and other plutons, volcanic neck, steep sloped volcanic cone, exfoliation dome and round topped landforms
	Basalt	Resistant, hard, massive, basic	Shield, plateau, Decan trap like landforms in the region of more than one phase of eruption, dissected plateau, mesa, butte, flat topped landform.
	Sand stone	Resistant, permeable and porous	Highland, cliff
	Shale	Non permeable and soft	Lowland, basin, valley, mudflow scar, gentle slopped landform
	Limestone	Soluble and fractured	Karst landform
	Schiest	Foliated and soft	Slip along foliation surface and planner scar due to slope failure.

■ 1.4.1. Processes of Weathering, Mass wasting and erosion and resultant Landforms

■ 1.4.1.1 Introduction and Definition

'Weathering the breakdown and alteration of materials near the earth's surface to products that are more in equilibrium with newlyimposed physico-chemical condition' - Ollier,C.(1974).

According to the functional approach, weathering is the response of the rocks or minerals to the change in the interactive combination of the weather elements i.e physical environment to achieve temporary equilibrium condition.

The definition by Richie (1950), advocates this concept. 'Weathering is the response of materials which are in equilibrium within the lithosphere to conditions at or near its contact with the atmosphere, the hydrosphere, and perhaps still importantly, the biosphere.'

■ 1.4.1.2 Importance of weathering in Geomorphology

- The process of denudation depends on the rate and types of weathering and the evolution of land forms largely depend on the denudation.
- The geometric change of landform depends on the rate and types of weathering
- The drainage and rate of infiltration depends on the weathering products viz. regolith and saprolite.
- The rate of removal of the weathered product is the primary factor of landform evolution.

■ 1.4.1.3 Factors of weathering

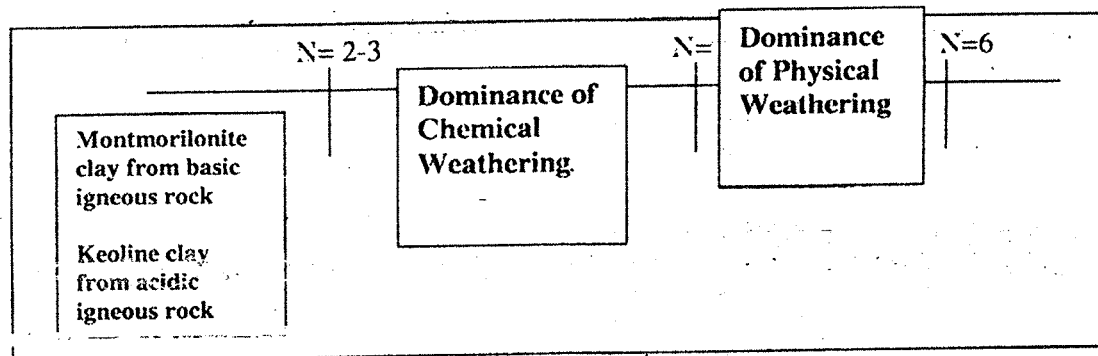
Jenny(1941) was of the opinion that soil is the combined effect of weathering and is the function of climate, topography, parent materials, organism and time.

- **Climate** mainly the availability of moisture and temperature are the determinants of both types and rate of weathering. The availability of moisture depends on effective precipitation and it determines the amount of Hydrozen ion and biomass. The rate of oxidation and reduction, hydrolysis, presence of clay minerals etc. depend on the availability of moisture.

Weinert in 1961,64 & 65 proposed that

Potential Evaporation- Precipitation ratio (N) = $12 E_j / P_a$.

where E_j is the precipitation of January, P_a is the annual Precipitation.



Dominance of different weathering according to Weinert, 1965.

Keller in 1957 divided the weathering zones into broad three groups viz. a) sub-aerial zone b) sub-surface zone c) sub-aqueous zone according to the availability of water. The sub-surface zone receives water from atmosphere in the form of rainfall and it takes part in both physical and chemical weathering. The subsurface zone is composed of the zones of percolation, capillary action and fluctuating ground water table and is affected by alternate freeze and thaw, through flow, illuviation, eluviation, oxidation and reduction. The sub-aqueous zone is characterized by anaerobic condition and is the zone of solution and precipitation of soluble matter.

- **Topography** viz. the landform is characterized by height and slope and these two regulate the temperature and rainfall and this causes the variation in the type and rate of weathering and thus the property of soil varies according to the distance from water divide on both sides of it and is termed by Ollier as **Toposequence**.
- **Parent material** is an important determinant of the type and rate of weathering due to their variation in character like mineral composition, texture, porosity, structure and bulk property. The presence of feldspar, calcium plagioclase, calcium carbonate and homogeneous composition etc. do not favour mechanical weathering but invites chemical weathering. Fine grained granite, limestone, quartzite, gneiss are resistant to mechanical weathering. Calcareous sedimentary rocks, limestone, basic igneous rocks, marble, dolomite etc. are more prone to chemical weathering.
- **The biotic factors** mainly the presence of burrowing animals and plants affects both bio-mechanical and bio-chemical weathering by either transferring or mixing the matters and bringing new matters at the contact of atmosphere and hydrosphere for more and more weathering.

■ 1.4.1.4 Mechanical weathering-stresses and processes

The disintegration of rocks by the hydrological, thermal or crystal stresses into more durable products is called mechanical weathering.

▲ Different stresses for mechanical weathering ▲

■ Geological stress

The superincumbent pressure(for sedimentary rocks), stress from crystal formation(granite or marble), compression and tension from tectonic force are responsible for mechanical weathering by fracturing and disintegration.

■ Geomorphological stress

Gilbert (1904) suggests that the unloading by denudation causes the release of pressure and which is responsible for the upward expansion of rocks causing the formation of exfoliation or unloading sheets. This unloading sheets, according to Chorley et.al. 1985, may be realized upto a dept of 100m and may be of few cm to 30m thickness. In the granite rock, due to the greater elasticity, such type of unloading sheets are common and surface domes are developed.

■ Thermal stress

The thermal gradient in the rocks causes the stress and develops the fractures to be disintegrated into pieces. In the monanzite quartz at Mohave desert the thermal gradient is $0.5^{\circ}\text{C}/\text{cm}$. The research shows that the thermal gradient only can not bring any remarkable change in the rock but the presence of moisture may cause considerable change in this respect.

■ Hydrological stress

Alternate wetting and drying causes expansion and contraction of rocks and thus fractures are generated in claystone and montmorillonite clay. The formation of ice crystals and salt crystals associated to the hydrological and thermal regimes develops the stresses.

■ Biological stress

The small creatures like earthworms can make the holes of 1.5-2m and can transfer the soil upto the diameter of 1mm. On an average, earth worms can transfer 43 tonnes of materials per acre per year.

The plants can initiate and enlarge the fractures in the rocks by the taproots and the fractures may be as long as 3m.

■ 1.4.1.2 Processes of Mechanical Weathering

Mechanical weathering is a combination of the following processes.

- Insolation Weathering
- Unloading Weathering
- Crystal Weathering
- Moisture Weathering

■ Insolation Weathering

○ Insolation cracks ○

The alternate expansion and contraction of rocks due to the change in temperature helps in the development of cracks and subsequently the rocks become fragmented.

○ Block disintegrations ○

The temperature gradient developed in the rocks due to differential heating causes stress in the rocks and it lead to the development of sets of horizontal and vertical fractures and thus the blocks are separated.

○ Granular disintegration ○

The different specific heat of different minerals, which the rock is composed of, helps in the differential expansion of minerals. Sometimes the anisotropic expansion of minerals causes granular disintegration of rocks.

○ Boulder cleaving ○

The rocks or boulders being exposed on slopes as an elevated exposures are prone to break in to halves to be leveled with the slope and this is called boulder cleaving.

■ Unloading Weathering

○ Sheeting ○

Gilbert, 1904 proposed that due to removal of vertical pressure and resultant expansion, sheets are separated by the development of horizontal fractures. Bradley, 1963 was of the opinion that in massive sandstone such sheeting is also possible.

○ Spalling ○

Sheets are generally flat and thin but in the cave or tunnel the outer layer is separated in the form of small fragments and is called as spalling. The tension cracks and shear cracks (due to compression) are responsible for spalling.

■ Crystal Weathering

o Ice crystal weathering o

The increase in the volume after freezing may increase in the pressure up to 30000 lb/in² at -22° C temperature and it causes the pressure helping in fracturing. Formation of ice wedge thus helps in downward increase the fracture. The alternate freeze and thaw also helps in the fracturing and mechanical weathering.

o Salt crystal weathering o

In porous rocks, the crystallization of the salts within the pores at the surface of the rock may lead to granular disintegration or possibly to some sort of exfoliation (Ollier, 1975).

■ Moisture Weathering

According to Nepper-Chirstensen (1965) in flint and basalt the alternate wetting and drying causes alternate expansion and contraction and is responsible for fracturing. Nishoka and Harada in 1958 opined that sandstone, shale, limestone, granite and other 22 rocks expand and contract by alternate wetting and drying. The change in the rocks and minerals by alternate wetting and drying is termed as slaking.

■ Chemical weathering

■ Factors:

Chemical weathering depends on Ph, Eh(Redox or Oxidation Potential) and Ionic Potential of the concerned solution and soil.

• Ph

The negative logarithm of the Hydrogen ion concentration is the value of Ph and this amount of Hydrogen ion affects the rate of chemical weathering. The solubility of the materials largely depends on Ph e.g. the solubility of iron increases 10000 times if the Ph of the solution increases from 6 to 8.5.

- Eh (Redox or Oxidation Potential)

The potentiality of a solution to be oxidized or reduced is called Redox or Oxidation Potential or Eh. In other words, the energy required for electron transfer in oxidation or reduction is the indicator of Redox or Oxidation Potential or Eh. The stability of a solution in its existing oxidation state depends on Eh.

- Ionic Potential

Hydration depends on the attraction of water by the ion. This attraction power of the ion is a ratio of charge of ion and radius of the particle. The attitude of the particle towards water is called ionic potential.

■ 1.4.1.5.2 Processes of Chemical weathering

- Solution

Most of the weathering, denudation and erosion process are constituted by solution and it depends on temperature, flow of water, availability of moisture, Ph of the solution and Eh. The following table indicates the importance of solution.

Sl. No	Continent	Total Soluble Matter (10^{14} g)	Rate of solution (ton/km ²)
1	N. America	7	33
2	S. America	5.5	28
3	Asia	14.9	32
4	Africa	7.1	24
5	Europe	4.6	42
6	Australia	0.2	2

- Carbonation

In this process the carbonate and bicarbonate ion is added to the rock and the soluble matter is developed and is easily replaced.

- Hydrolysis

In this process the water breaks in to hydrogen and hydroxyl ion and this hydrogen ion is more mobile. This hydrogen ion is smaller and have more charge which helps to enter easily into he lattice of the minerals and thus the break of the minerals become easier.

- Hydration

The combination of water with mineral and mineral combination is the hydration. It may bring changes in minerals in different ways.

- b) New mineral may develop after combination with water with associated volume change (Anhydrite Calcium Sulphate may develop gypsum being added to water with the increase in volume.)
- c) Water may enter into the wide lattice of the clay minerals and it helps in the swelling of it. The sodium rich Mancos Shale in Colorado increases by 60% of its volume in addition with water.
- d) The salt in the porous rocks may hydrated readily and frequently and it creates stress and matter breaks into pieces.
- e) The hydration may bring preparation for the further degradation by hydrolysis or solution.

- Oxidation

The change involving the discharge of electron i.e. increase of positive valence and decrease of negative valence is called oxidation. Iron, manganese, sulphur etc. are easily oxidized. For ready oxidation, immediate drainage, exposure to weather, easy decomposition of biotic matter are necessary. By oxidation electron is discharged from the lattice of a mineral and thus the lattice is easily broken and other ion thus can easily enter into the lattice and cause further weathering.

- Reduction

It is the change involving increase in the negative valence. This type of change occurs below gound water level, at the bottom of water bodies where anaerobic condition prevails.

- Chelation

The bond structure of biotic matter and metal ion may be replaced together inviting the condition of excess humus and acidic change. Podsol soil is developed by this change.

■ 1.4.1.5.3 Weathering Landforms

A. Constant Volume Weathering and landforms

- Regolith and saprolite

The decomposed rocks in situ is called saprolite and the combination of in situ decomposed rocks and partially transported materials are called regolith.

- Desert varnish

In desert region quartz particles, boulders etc are characterized with a lusturous coating. This coating is made up of iron oxides, manganese dioxides and some times of silica.

- Patina

The exposed particles and boulders develop oxidized layer and it helps in distinction of those from the others in colour, porosity and other characters and thus called as patina.

- Rounded landforms

According to Ollier, due to chemical translocation and decomposition of matters within the rocks, the outer layer of it is exfoliated and rounded boulders are thus developed.

- Color banding

Due to deposition and dislocation of iron oxides, different bands with separate color is developed in rocks or soil by alternate oxidation.

- Karst landforms

Solution landforms like sink hole, polji, uvala, caverns etc are characteristic feature of karst region.

B. Weathering with expansion and related Landforms

- Unloading dome

The unloading by exfoliation may cause the development of a round topped dome of ten to few hundred feet height, covered with fragmented unloading sheets and generally devoid of vegetation.

- Unloading sheet

The sheet of rock, the outer layer, detached from the interior by unloading of pressure by denudation may be of considerable thickness and generally found on either side of a v-shaped valley.

- Scree

The sharp edged fragments due to insolation effects and development of ice wedge are generally deposited at the foot of steep slope and is called as scree.

C. Landforms due to differential weathering

- Weathering pit

The concentrated weathering causes the development of weathering pit. Such pits on steep slope is called as weathering pan.

- Tor

The differential weathering help in exposing the core rock to the surface by removal of the surrounding softer rocks. The exposed core stone then is fragmented by deep seated weathering and ultimately is covered by boulders and called as tor.

- Inselberg

The differential weathering and stripping of the weathered materials are the basic cause for the development of inselberg. Ackermann, 1962 was of the opinion that the development of penitent rocks and subsequent removal of saprolite may cause the development of inselberg.

- Etchplain

The weathering and subsequent stripping of the saprolite may develop different levels of weathered surface characterizing the weathering of different ages and is termed as etchplain by Wayland in 1933.

- Weathering residues

The clay minerals, residual iron oxides, bauxite residues etc. are developed by chemical weathering and subsequent leaching.

■ 1.4.3 Mass Wasting and Landforms

The general term "mass movement" is applied to those processes that involve a transfer of slope forming materials from higher to lower ground, under the influence of gravity, with out primary assistance of a fluid transporting agent. They merge imperceptibly with processes in which transporting media such as air water or ice are involved and which are generally called 'mass transport process' (D. Brudsen, 1979). Mass wasting, on the other hand, is the down slope movement of materials primarily with the effect of gravitational pull being assisted by other geomorphic processes.

Causes of Mass Movement

Mass movement is caused as a result of interaction of the processes responsible for detachment (shear stress) and displacement of materials and the processes responsible for the resistance or shear strength of materials.

$$\text{Safety Factor} = \frac{\text{Shear Stress}}{\text{Shear Strength}} \leq 1$$

Mass wasting may be developed as a result of two basic causes

A. Factors responsible for decrease in Shear Strength & Factors responsible for increase in Shear Stress

A. Factors responsible for decrease in Shear Strength

- **Weathering**

The disintegration and decomposition help in the reduction in the cohesion and angle of internal friction and thus help in the down slope transfer of materials easily. The solution process help in the down slope transfer of material.

- **Seepage Erosion**

Water after entering into the soil reduces cohesion, reduces surface tension of drained soil, remove the cementing material thus makes the material unstable.

- **Liquifaction**

The saturated soil loses the shear strength and behaves like fluid and moves down slopes. The weight of saturated soil increases and so down slope movements becomes easier. If the pore water pressure increases, the dislocation of soil particle is further increases.

- **Remoulding**

As a result of earthquake, explosion or landslide the structure of soil breaks and becomes softer to initiate landslide.

- **Reduction in angle of internal friction**

Angle of internal friction depends on the texture of the material and moisture condition. The low angle favours down slope movements in saturated condition.

- **Progressive failure**

According to Terzaghi, 1960 – 62 the stability of matter of fractured slope depends on friction and cohesion. If the jointed and detached parts are removed one by one, the stretch is increasingly concentrated on the remaining intact pieces and the process of down slope movement accelerate as a chain reaction, Brudsen, 1979.

D. Factors responsible for increase in Shear Stress

- Increase in the length and height of slope

The down cutting by river and glacier increases in the length and height of slope and as these two components of slope increase the shear stress on the slope forming materials. The critical height for the initiation of slide depends on weight of unit materials involved, slope gradient, angle of internal friction and cohesion (Carson, 1971).

- Unloading

The denudation causes the reduction of downward pressure and this may cause the lateral spread, extension of fractures, increase of porosity, softening of rocks in long run etc. and thus invites down slope movement.

- Loading

The increased deposition on slope by either natural (deposition of scree, ice, pressure from large trees) or anthropogenic (construction of building, water storage through construction of dam etc.) processes increases downward pressure and this down ward pressure helps in downward movement.

- Shocks and vibration

The earthquake, explosion and vibration from vehicles breaks the bonding of the soil particles and reduces the cohesion and help in downslope movement.

- Removal of underlying support

The down cutting and basal erosion by rivers and wave removes the materials from the base of the slope and thus removes the underlying support and down slope movement takes place.

If water is added to the shalestone lying below other hard rocks, it removes the basal support and the underlying support. The mining from slope exposure will loose the underlying support for the upper slope segment. The landslide of 1968 and long continuing subsidence of Tindharia Cricket Colony is due to the removal of coal from the slope exposure at the base of this village (Maiti R.K, 2004,2006).

■ 1.4.3.1 CLASSIFICATION OF MASS WASTING

A Vertical Movement

- Falls
 - Rock falls
 - Earth falls
 - Debris falls

- Topple
- Subsidence
 - Collapse
 - settlement

B Lateral Movement

- Block glide
- Spreading
- Cambering
- Sackung

C Diagonal Movement

- Creep
 - Soil creep
 - Rock creep
 - Talus creep
 - Rock glacier creep
- Slide
 - Rock slide
 - Debris slide
 - Soil slip
 - Slump
- Flow
 - Soil flow
 - Debris flow

■ 1.4.3.2 Landforms developed by Mass Movement

The displacement of slope forming materials from upslope is mainly responsible for change in landform. Mass movement is thus most effective geomorphic agent and all the other geomorphic agents change the landform with direct help of it. Mass wasting affects landform directly in three ways

- Dislodgement of debris
- Transportation of dislodged materials on slope
- Deposition of transported materials at foothill

▲ Dislodgement of debris ▲

The landform developed as a result of the displacement of regolith and saprolite depends on the quality of the materials and level of consolidation, amount of clay, rate and the type of mass wasting.

- Free face- the entire free face (near vertical slope) is the result of dislodgement of loose debris.

- Land slide scar – On steep slope the displacement of materials develop the scar on the slope face. Landslide scars have their shape according to the nature of weathered matter, orientation of fractures, structure of the rocks etc.
 - Circular scar- in deep soil and structureless rocks, circular scars are developed.
 - Planner scar- in the rocks with horizontal structure like slate and schist such plane like scar is developed.
 - Wedge like scar- wedge like scar is developed as two fractures intersect at acute angle on deep soil.
 - Toppling scar- in the rocks with columner structure, if the gradient of the fracture is steeper than the slope gradient, the dislodged block will fall with forward rotation.

▲ Transportation of dislodged materials on slope ▲

During transport of dislodged matter on slope with huge energy, a considerable change along the path is made. The longitudinal scar developed depending on nature of the transported materials, elasticity, steepness of the slope, length of the flow, speed of the flow etc.

▲ Deposition of transported materials at foothill ▲

After the deposition, the materials develop the following landforms

- Colluvial cone- the deposition of dislodged materials at the foothill develops the conical landform, the colluvial cone, having concave longitudinal and convex cross profile.
- Scree or talus- the sharp edged rock fragments being deposited at the foot hill or at the foot of the slope is the scree or talus.
- Solifluction lobe is the tongue like landform developed at the foot of the slope by solifluction. is the solifluction lobe.

■ 1.4.4 Slope Development and Slope facets

Slope is the prime component of any landform, varies from 0° to 90° or above (i.e. overhanging slope). The evolution of slope is highly debated problem in geomorphology. Some important concepts of slope evolution are being discussed below.

▲ Components of a slope ▲

An ideal slope is composed of convex crest (waxing slope), steep free face (escarpment), constant slope (rectilinear) and concave slope (waning slope). The waxing slope is developed in an environment of increasing erosion. The free face, being the steeper part is eroded faster giving rise the development of a gentle slope at the base where available energy is only equal to the required energy for the transport of already eroded materials.

▲ Process of slope formation ▲

Slopes are mainly formed by the endogenetic processes and are modified by the exogenetic processes.

▲ Endogenetic processes ▲

- Movement
 - Epirogenic
 - Endogenic
- Earthquake
- Volcanism

▲ Exogenetic processes ▲

- Mass Movement
- Fluvial process
- Glacial process
- Wind action
- Wave action etc.

▲ Theories of slope development ▲

● View of Penck

The fundamental analysis of slope problem is done by W. Penck in his 'Die Morphologische Analysis'. Penck stated that shape of landform of slope is dependent on relative rate of upliftment and that of erosion.

1. If the rate of upliftment is more than the rate of erosion, the height of the landform will be increased and the energy of erosional agent will also be increased subsequently as a result (according to Penck) the Waxing slope or Convex slope will be developed.
2. If the rate of upliftment becomes less than the rate of erosion, the land will be decreased in height and the energy of erosional agent will also be decreased and more deposition will occur. Thus the Waning slope or Concave slope will be formed.
3. If the rate of upliftment and rate of erosion become equal, the height of the landform will remain constant and a Constant slope will be formed.

Theory of Parallel retreat

Penck's concept of slope formation is mainly based on the condition of parallel retreat. According to him if the rate of upliftment and rate of erosion become equal, the slope will retreat parallelly. In case of a non eroding river, if the slope is made of homogeneous composition, the slope will retreat parallelly due to removal of equal amount of weathered material. The removed material is deposited at the base of the slope and formed an another gentle basal slope termed as Haldenhang after Penck. This basal slope also retreat parallelly but due to the gentle gradient the weathered material will take longtime to remove. In case of active eroding river, if the slope is made of homogeneous composition, the relation between the rate of denudation and rate of down cutting is

$t_a = t_o / \cos \alpha$, where t_a = rate of down cutting, t_o = rate of denudation and α = slope angle.

● View of Wood (1942)

Allen Wood assumed a face in the beginning of slope formation as the result of tectonic action or erosional process. As weathering proceeds a talus slope is formed in front of the free face due to the deposition of weathered material from the free face. Due to exposure to weathering the free face become convex in shape and the talus slope is enlarged and become tangent to the convex rock slope. This slope is termed as 'Constant slope' by A.Wood. Due to further erosion, the talus slope is subjected to erosion and shaped as concave which become tangent to the local base level. According to A.Wood, slope has four components – free face, constant slope, waning slope, waxing slope.

● View of Davis – Slope decline theory

Davis described the slope form in respect to time. He discussed the development of slope in different stages.

Youth In the youth stage, there is a initial slope with little steepness. As the consequent river starts it's active down cutting, the valley sides slopes are developed with steeper gradient and are being increased in length by the time. The shape of the slope in this stage become convex.

Mature The length of the slope becomes maximum in this stage. The water divides are reduced to sharp crest. The angle of slope is decreasing during this stage. The Graded slope is formed in maturity phase. The graded slope may be defined as the slope with such an angle / dip where the enegy of the erosional agent is finished to transport the already weathered of eroded materials.

Old Old stage is characterized by almost flat or Concave shaped slope. The slope angle becomes minimum, during this stage. The area is turned to peneplain with some moriadnocks(residual hills).

● View of L.C.King

L.C.King accepts the four components of slope i.e. free face, constant slope , waning slope, waxing slope as proposed by A.Wood. King stated that free face retreat parrallely and leave a gently sloping concave slope termed as 'Pediment' at its foot. Ultimately the whole area will be covered by the pediment with some residual hills termed as inselberg. This landform is called as 'Pediplain'.

King recognized two types of water flow – Laminar flow and Turbulent flow. There is no significant erosion in laminar flow whereas turbulent flow makes a significant erosion.

● View of Savigear

According to Savigear the shape of slope depends on the rate of erosion and rate of removal of materials from the base. He cited an example of Carmathan bay of South Wales; where the eastern coast is steeper due to active wind action and western coast is gentle due to inactive wind and lower rate of removal of materials.

■ 1.4.5 Concept of Grade, Profile of Equilibrium and Base Level

■ 1.4.5.1 Grade;

Grade is a condition where the available energy is so less to encourage further erosion and there is no deficit of energy to initiate deposition. That means the available energy is equal to the energy required for the transport of eroded materials. Thus it is an equilibrium or constant condition where no change in slope is possible. As the slope attains such a gradient in adjustment with the other conditions so that the energy becomes equally distributed on the surface indicating the maximization of entropy and low availability of energy only equal to that required for transportation of debris down slope. The gentle slope on peneplain, pediment, panplain and other planation surfaces are of such type on which regulated and restricted energy is released to fit it to the equilibrium condition. The graded profile, according to Davis, of a river both along long and cross profile proposed to be developed at the mature or old stage of a cycle is also a smooth curve established at the mouth of a stream first and gradually proceeds head ward slowly. Any change in the factors of erosion may bring the modification and break in the equilibrium slope inviting a new slope below the existing slope.

■ 1.4.5.2 Profile of Equilibrium

- **Concept of Equilibrium**

Equilibrium is a state of adjustment where rate of energy and mass transfer becomes minimum being characterized by entropy maximization and low availability of energy. This low energy condition is represented by less erosion or no erosion and thus a stage is achieved in ultimate condition and every system tries to achieve this situation.

There are few types of equilibrium

- **Decay Equilibrium-** It is the condition when rate of change of form declines or decay through time. This time is mainly the cyclic time which spreads for millions of years and through this span a landform passes through sequential and well determined phases of landform development guiding towards an equilibrium at ultimate stage in the form of peneplain where the entropy maximization is achieved through equal distribution of energy.

- **Steady State Equilibrium**-it is condition when form oscillate around a stable average value due to interacting feedback loop in the system. Every system in short time oscillate around such a average state some times with higher and some times with lower value and thus following an average pattern.
- **Dynamic Equilibrium**- It is the condition of Equilibrium about a mean value of form which is changing through time. Any change in the input of a system may bring a manifold change in the output and this system will try to achieve equilibrium in this newly achieved condition.
- **Dynamic Metastable Equilibrium** - It is the condition of Equilibrium about a mean value of form which is trending through time and at the same time is subjected to step like discontinuities as a threshold effect promote a sudden change of form. A sudden change in the system after achieving a threshold condition by positive feedback will cause a drastic reduction of height and slope by negative feedback.
 - **Equilibrium Profile**- It is a smooth profile extending from the mouth of a stream towards source having gentle slope which can release limited energy only equal to that required for the transport of mass of sediments down slope. It is a kind of equilibrium profile on which no erosion and deposition is occurred for short duration.

■ 1.4.5.3 Base level of erosion:

It is the ultimate level of erosion beyond which no erosion is possible. The energy of a geomorphic agent (processes) comes from the acquired potential energy.

PE (potential energy) = $m \times g \times h$;

where 'm' means the mass of the agent, 'g' means the gravitational acceleration and the most important "h" means the height above base level.

This means more will be the height above base level more will be the energy of a geomorphic process. The change involving the increase in 'h' means more energy available for erosion and this can invite the situation of rejuvenation. As Geomorphology deals with the change of landform, and as any change in the shape of landform requires the movement of materials from one part to another and it requires energy. The most important regulator of energy is thus height above base level. So the concept of base level is an important prerequisite for geomorphic understanding.

Not always the sea level act as the base level. It can only act directly on the river and glacier systems which emerge into the sea at the estuaries.

- **Types of base level;**

- **Local Base level**

- If a tributary joins a main stream or a lake, the level of that main stream or lake act as the base level for that tributary because no erosion is possible at the tributary basin area below that level of mouth.

- **Temporary Base level**

- The hard rock, situating along the long profile of the stream can act as the base level for the upper reach as no erosion is possible for this portion to become denuded below this level. This hard rock may be eroded at a slower rate and ultimately be obliterated and thus is termed as Temporary Base level.

- **Grand Base level**

The sea level act as an ultimate level of erosion and is called as Grand Base level.

■ 1.4.6 Self Assessment Questions :

1. How does the rock structure affect the development of landforms?
2. Explain why the porous rocks are considered as the resistant rock.
3. Define permeability. How does the permeability affect the development of landform?
4. Describe the effect of climate as a factor of weathering.
5. Briefly describe the factors of weathering with special reference to parent materials.
6. Explain the stresses responsible for mechanical weathering.
7. Describe the effects of Ph and Eh on the progress of Chemical Weathering.
8. Describe the processes of Chemical Weathering.
9. Describe the landforms developed by Weathering.
10. Describe the causes of Mass Movement with special reference to Safety Factor.
11. Describe the landforms developed by Mass wasting.
12. Describe the components of slope. Critically discuss the theories of slope development.
13. Describe the concept of grade. How is the concept of equilibrium related to the concept of entropy maximization?
14. Define base level. Describe different types of base level. What is the importance of base level in the study of Geomorphology?

■ 1.4.7 References :

- Chorley, R.J, Schumm, S.A & Sugden, D.E. (1984)- pp 12-14, 201-229, 230-255..
- Bloom A.L., (1992) *Geomorphology, A systematic analysis of late Cenozoic landform* pp1-12, 175-209.
- Singh, S. (1998), *Geomorphology*, pp. 12-23, 24-56, 257-266.
- Bland, W. and Rolls, D. (1998) *Weathering*, Oxford University Press, 240p.
- Ollier, C. (1975), *Weathering*, Longman Group Limited, 295p.
- Wooldridge, S.W. and Morgan R.S. (1988) *An outline of Geomorphology*, pp 173-193

PART-I
Paper - I, Module - I
GEOTECTONICS AND GEOMORPHOLOGY
Unit – 05

MODULE STRUCTURE

- 1.5.0 Evolution of Landform
 - Introduction
 - Objective
- 1.5.1 Cyclic Evolution Theories
 - Introduction to cyclic concept
 - Normal Cycle of Erosion by W.M. Davis
 - Assumptions
 - The Model
 - The Theory
 - Merits
 - Demerits
- 1.5.2 Penck's concept of Primal Evolution
 - Assumptions
 - The Model
 - The Theory
 - Criticism
- 1.5.3 L.C. King's concept of Pediplanation
 - Introduction
 - Assumptions
 - The Theory
 - Similarity with Davisian Concept
 - Dissimilarity with Davisian Concept
 - Criticism
- 1.5.4 Non- Cyclic Evolution Theories
 - J. T. Hack's Concept of Dynamic Equilibrium Theory

- Introduction
- Assumptions
- The concept of Dynamic Equilibrium
- Criticism

☐ 1.5.5 Dynamic Metastable equilibrium Theory of Chorley & Schumm

- Introduction
- The concept of Dynamic Metastable Equilibrium
- Complex Response
- Geomorphic Threshold

☐ 1.5.6 Applied Geography

- Introduction
- Objective

☐ 1.5.7 Application to Hydrology

- Hydological problems in karst region
- Ground water in glaciated areas:-

☐ 1.5.8 Application to Economic geology :

- Surface expression of ore bodies:-
- Weathering residues:-
- Study of Unconformity
- Placer deposits:-

☐ 1.5.9 Application to Engineering

- Construction of Road
- Sand and Gravel Quarry:
- Construction of Dam

☐ 1.5.10 Application for the mitigation of environmental hazard

- Management of Flood
- Management of Landslide and Erosion

☐ 1.5.11 Self Assessment Questions

☐ 1.5.12 References

PPART-I

Paper - I, Module - I

GEOTECTONICS AND GEOMORPHOLOGY

Unit – 05

|| 1.5.0 EVOLUTION OF LANDFORM

Introduction

Geomorphology deals with the origin and evolution of landforms and so the study of landform evolution with respect to long cyclic time, short- graded time or instantaneous time is quite relevant even today. The so-called time-independent evolution theories are associated to either of these times. The change in landform as an systematic interaction of all the active factors are drawing the increasing attention of the recent geomorphologists. The evolution of landform can be viewed through either historical or functional approach though the later one is more popular at present.

Objective

The basic objective of this part of the module is to make the learner familiar to the processes how the approaches for the study of landform evolution is changed from time to time and how do the evolution of landform is viewed at present.

The concept of W.M. Davis is the pioneer in this concept but this concept is also inherited from 'cyclic nature of earths history' of James Hutton (1785). The process of evolution should follow an inevitable path in the framework of time and it is the unidirectional process guiding towards a situation of entropy maximization is the prime concept of Davis but the recent theories states that this process of evolution is by-directional and is not continuous but divided into numbers of geomorphic episodes including positive and negative feedback in the process of change. The basic message of this part of the module is that all the evolution theories are connected with the other and is the result of improvement upon the former one.

|| 1.5.1 CYCLIC EVOLUTION THEORIES

o Introduction to cyclic concept

The change of landform in long periods of cyclic time can be well represented through the models of Devis, Penck and King. The historical approaches and functional approaches of landform evolution found rational synthesis in systematic description of landform systems in

relation to spatial variation of structure and processes viewed in temporal sequences. The models of Davis, Penck and King answer the mechanism of evolution of land form in respect to time following a sequence leading towards an equilibrium.

The (cyclic) concept of evolution implied an inevitable, continuous and broadly irreversible process of change producing an orderly sequence of landform transformation – Chorley et al 1985.

In general sense, cycle means an circular movement, where initial and terminal position remains same. In the concept of W.M. Davis (1989) the evolution of landform is guided rather forced to a unidirectional path following some sequences towards a final stage of equilibrium (Entropy Maximisation).

Some geomorphologists believe that this cyclic concept of land form evolution was inherited from the concept of 'cyclic nature of earths history' of James Hutton in 1785 and the other concept of 'uniformitarianism' according to Chorley, Schumm and Sugden. Cyclic concept of W.M. Davis followed the 'graded concept' of Gilbert. In the cycle the graded condition along both long and cross profile of a stream has to be achieved in late mature and old stage to attain equilibrium.

Recently time-independent functional and systematic approach of land form evolution became popular among the geomorphologists. The evolution of land from a state of high concentrated energy to low distributed energy giving rise an equilibrium is believed to be important.

There are two sets of concepts for land form evolution:-

A	Time dependent historical approaches	Cyclic/sequential stage wise evolution	A1. Stand Still Theory of normal cycle of erosion by W.M. Davis. A2. Dynamic Theory of Primarumph evolution by W. Penck A3. Cyclic concept of Pendiplanation by L.C. King
B	Time dependent functional theory	Equilibrium concept	B1. Dynamic Equilibrium concept of J. T. Hack B2. Metastable Equilibrium concept of Chorley & Schumm.

o Normal Cycle of Erosion by W.M. Davis

During 1884 – 1899 American Geomorphologist W.M. Davis developed his concept of cycle of erosion following the evolution concept of Charles Darwin and developed by T.H.

Huxley and Herbert Spencer. In his cyclic concept the evolution of land form was fixed following the life cycle of an organism. 'Davis should have developed a model of sequential land form changes through time, so that, as he put it, land forms like organic forms shall be studied in view of their evolution' – Chorley -1984.

Each stage in Davisian cycle was associated with declining potential energy, as the relief was worn down, and each stage was characterized by an assemblage of landforms— an irreversible equalization of energy levels through out the land form assemblage, leading ultimately to a spatially uniform terrain (the peneplain)-Chorley et. Al 1984.

The cyclic model of Davis is composed of two phases (a) Tectonic phase and (b) Erosional phase. Tectonic phase is of short duration during which rapid upliftment takes place and erosion phase, on the other hand, is a long period of continued erosion during which land form passes through an orderly sequential stages of youth, maturity, and old. During tectonic phase land form starts upliftment from sea level and after the ultimate stage of land form evolution land is reduced to sea level again and thus a cycle is complete with the same initial and terminal point.

Basic Assumptions of Davis.

- Its is assumed that the lithology of the entire region is same.
- Land is uplifted rapidly from sea level and during this, there is no significant erosion. Potential energy is realised due to this initial upliftment.
- Erosion starts after the upliftment is over.
- The region is situated in a humid topical climate to facilitate the fluvial actions.
- The region should be stable for long to allow the cycle to be complete.

Davis proposed his cyclic concept of land form evolution in 1988 in which a well defined and well determined path for the change in landform was put forward. "There are sequential changes in land form through time (passing through youth, maturity and old stages) and these sequential changes are directed towards well defined end products development of land form" singh 2000.

The Model

1. **Central concept** - The geographical model is a period of time during which an uplifted landmass undergoes its transformation by the process of land sculpture, ending into low featureless plain or peneplain (Singh, 2000).
2. **The inevitable relation between time and geometric shape of land form** -By this model, time became not a temporal framework within which events could occur but a

processes itself leading to an inevitable progression of change. Thus the passage of time, according to Davis, imprinted on land forms' progressive changes in their geometry, from which the passage of time could be inferred as surely as from the changing geometrical relation of the hand of clock (Chorley et al, 1984).

3. 2nd Law of Thermodynamics and Entropy Maximisation

2nd Law of thermodynamics states that in a closed system entropy gradually increases and this change is unidirectional and irreversible. entropy (Degree of disorderness) means En – inside; Trope = Transformation. In a system how much energy is exhausted and how much energy remained as excess to perform the work is indicated by entropy. According to Davis as cycle proceeds, the concentrated high energy is gradually reduced and distributed till entropy is maximised by equal distribution of available energy with lowest possible intensity in a unit space. Davis termed this condition as peneplain.

4. The input of energy in the system comes from potential energy generated by initial upliftment.

The source of energy in this system comes from initial upliftment of primary landmass and this energy gradually reduces as the land is lowered by erosion and ultimate reduction to sea level.

The theory :

Davis proposed that the evolution of land form is irreversible and follows an inevitable path through the stage of youth, mature and old. He put forward a remarkable equation for landform evolution "Landform is a function of structure, process and stage". The stage got more emphasis in his cyclic model.

After the upliftment of a land from sea level is over, the erosion starts in a sequential way which passes from youth through maturity and ultimately to the old stage where the land is reduced to the sea level. Again it can uplifts to its initial position.

After the upliftment of land, river valleys are developed by the accumulation of surface water along the channels. Valleys are developed being separated by wide and extensive interfluvium. Rivers start active down cutting and so the depth of the valleys increase at a faster rate. The upper parts of the valley side get eroded which causes the widening of the valleys and aerial reduction of the interfluvium. At the end of the youth stage the valley attains the maximum depth and the interfluviums are reduced to sharp crest. At this point the relative relief becomes maximum. The down cutting stops at this moment and lateral erosion starts. the flood plains on both side of the river become extensive, the height and area of the interfluvium are reduced to the residual

hills (Monadnocks). At the ultimate stage, the extensive flood plains situated near the sea levels are called as peneplain.

W.M.Davis elaborated his concept in a graph by the two lines the upper one indicating the height of water divide & the lower one indicating the height of valley floors. When AB denotes the initial relief and X,Y denotes the maximum relief at the end of youth. It become least at the ultimate stage by the development of almost plane like erosional surface (Peneplain).

Merits:

- i) Cycle model of Davis is easy to understand and has a wide range applicability.
- ii) The proposed landforms have direct relation to the actual field.
- iii) Davis' Model brings a revolutionary change in the study of landform development.
- iv) This Model can explain the historical interpretation of landform i.e. the history of land from development.

Demerits:

1. The nature of upliftment as proposed by Davis is criticised by W. Penck. According to plate-tectonic theory the upliftment is slow and is a continued process.
2. According to Davis no erosion will take place during upliftment but erosion is a natural process it doesn't wait for the completion of upliftment.
3. For the completion of Devisian cycle, the crustal stability is necessary but always is not possible.
4. According to Davis, land form is a function of structure, process and time. Many other Geomorphologist. e.g. J.T. Hack Strahler and Chorley did not consider the time factor in land form development.

□ **1.5.2 PENCK'S CONCEPT OF PRIMARUMPH EVOLUTION**

Penck proposed his concept of "morphological system of land from evolution' as a criticism to that of W.M. Davis. In his concept he tried to make a relationship between endogenetic and exogenetic proceses. According to him the development of different shapes landforms is time independent but depends on the relative rate of upliftment and rate of erosion.

Assumptions:

Penck's concept depends on the following asumptions

1. Unlike Davis Penck supposed that the erosion and upliftment perform together.
2. The shape of landform (or slope) depends on the relative rate of upliftment and

that of erosion.

1. If the upliftment is more than erosion, the land gets increasing height and energy of erosional agent gradually increases and waxing (convex) slope developments.
2. If the rate of upliftment is less than that of erosion, the land gradually lowers with the resultant decrease in the erosional energy. The waning (concave) slope develops by increasing deposition.
3. If the rate of upliftment is same to that of erosion the slope is unchanged.
3. Penck supposed that the evolution of a land starts from the upliftment of Primarumpf (feature-less primary peneplain)
4. The evolution of landform does not depend on time but on the relationship between upliftment and erosion.

Discussion of the Theory:

Penck proposed that the evolution of landform starts from the upliftment of a featureless primarumpf. The evolution of landform and its shape is divided into few stages (Entwicklung).

He proposed three such entwicklung

1. Aufsteigende entwicklung.
2. Gleichformige entwicklung.
3. Absteigende entwicklung.

1. Aufsteigende entwicklung

State-I

If the rate of upliftment is greater than that of erosion, the height of the landform increases and as a result, the erosive power will increase. Thus the amount of down cutting will be greater and so, the relative relief increases.

Stage-II

In this stage the upliftment remains greater than the erosion and so the absolute relief increases but the rate of down cutting become equal to the rate of Water divide erosion and so relative relief remain constant.

2. Gleichformige entwicklung

Stage-III

In this stage, the rate of upliftment and erosion become same and the rate of down cutting

and denudation of water divide also become same. Thus the relative relief become constant.

3. Absteigende entwicklung

Stage-IV

The rate of upliftment become less than erosion and as a result the absolute relief decreases but due to the equal rate of down cutting and denudation of watershed the relative relief become constant.

Stage-V

The rate of upliftment become less than erosion and the rate of downwasting of water divide is more than the down cutting and so, both absolute and relative relief decrease and the land is reduced to the base level and endrumph is developed.

If the rate of upliftment becomes equal to that of erosion, slope starts retreat parallelly with the reduction of its length, if the slope is eroded everywhere at the same rate because of homogeneous composition. As the slopes retreat parallelly there will form a concave pedepain and the residual part of the water divide made up of hard rocks remains as inselberg.

Criticism.

Penck's view of landscape evolution is not so popular because of a) his untimely death b) explanation in tough German language. c) ill-defined terminology etc.

His concept was criticised in U.S.A. mainly for his view of parallel retreat.

□ 1.5.3 L.C. KING'S CONCEPT OF PEDIPLANATION

L.C. king proposed the series of cycle model viz. landscape cycle, pediplanation cycle, hill slope cycle etc. on the basis of his extensive study in Africa. So his concept is associated to the arid and semi-arid landform. Through his concepts is evolved as a criticism to that of Davis, yet the geomorphologists find a similarity between the concepts of both of them.

Assumptions

Kings concept is based on few assumptions

1. King assumed that there is uniform development of landform in varying environmental conditions.
2. Major landscapes in all continents have been evolved by rhythmic global tectonic events.
3. All the landforms are composed of few basic elements.
 - a) Summits b) Scarps c) Debris slopes and d) Pediments.

4. The evolution of landform is mainly due to the parallel retreat of scarp and extension of the pediment.

5. Every cycle starts with the rapid rate of upliftment followed by long and continued crustal stability.

The Concept

The parallel retreat of slope and pediplanation are related to three stages i.e. youth, maturity and old.

Youth- After the rapid upliftment of earlier pediplain the cycle starts. The rivers start rapid down cutting and the long continuity of the graded long profile of the river break with the location of numbers of knick points etc.. The gorges and canyon start developing.

Mature- The active down cutting is ceased and lateral erosion starts. The scarps on the valley side start retreating backward with the remarkable development of pediment at the valley floor. The broad and extensive valleys are developed with remarkable formation of pediments on the either sides of it. By the retreat of scarps the interfluvies become reduced and narrowed down and ultimately those are reduced to residual hills with concave side which are termed as inselberg. At the late maturity all the individual pediments are jointed together to form extensive pediplain.

Old stage: The inselbergs are turned rounded due to continuous weathering and erosion as bornhardts. The extended weathering is responsible for the dissection of bornhardts and ultimately these are reduced to the heaps of boulders and are termed as castle-koppies or tors. Ultimately the total region is reduced to multi-concave surface, where inselbergs or castle-koppies mark the position of breaks in concavity.

Similarity with Davisian Concept:

- 1) As Davis, King proposed his concept in different stages i.e. youth, mature and old.
- 2) Similar to Davis, King also assumed rapid rate of upliftment and long crustal stability.
- 3) The end product of the erosional cycle which is termed as peneplain with monadnocks by Davis is termed as pediplain with inselberg by King.

Dissimilarity:

1. Davis gave emphasis on the down cutting and resultant valley deepening where king emphasised on lateral erosion and retreat of scarp.
2. Division peneplain has no further extension but according to king pediplain may be extended by lateral erosion or parallel retreat of scarp slope.

Criticism:

1. King's concept is based on his study on Africa and related to arid or semiarid landform. So it has no universal application.
2. The uniform development in different environmental condition is not possible as king assumed.
3. In the time when he proposed his concept, most of the geomorphologists are interested to the quantification and laboratory analysis and not to the evolution theory
4. He proposed his concept as the criticism of Division theory but in many respect it becomes similar to him.

o Non- Cyclic Evolution Theories

II 1.5.4 NON CYCLIC CONCEPT OF J.T. HACK

Introduction : As the criticism of division cyclic model Penck postulated his time independent concept of morphological system of landform evolution. In 1960, 65 J.T. Hack proposed his time independent concept of dynamic equilibrium as a further criticism of Davision concept. Hack's concept was based on his intensive study of Shenandoah valley of Appalachian (U.S.A.). According to him there will be no appreciable change in landform until the controlling factors of landform development and evolutionary processes and energy are changed.

Assumption:- Hack's concept is based on the following assumption.

1. There is a balance between rock resistance and rate of erosion.
2. There is uniform rate of down wasting on all components of landscapes.
3. The difference in landform are due to geologic variation.
4. The process which operate today are responsible for the land form development.

■ Concept of dynamic equilibrium:-

1. According to Hack the rate of erosion depends on rock types and characters i.e. hardness or softness, presence of joint or fracture, permeability and other lithological characteristics, angle of dip, climate, vegetation, rate of upliftment etc but not on time. He stated that land form does not change in respect of time.
2. Hack suggested that there is an equilibrium between the rate of erosion and rock resistance. When the rock resistance is changed, a new equilibrium is developed. Thus his concept is termed as dynamic equilibrium concept.

He cited two examples to elaborate or explain his time independent concept.

Example – I

He took two rocks with different resistance e.g. Quartz and shale. Quartz requires more intensity to be eroded and transported than shale. So to perform the same rate of erosion quartz requires greater slope and relief. Here two different sets of equilibrium are developed on different rock (shale & quartz).

Example – II

Hack assumed a multilayered surface of alternate hard (sandstone) and soft (clay stone) rock. In hard rock due to more rock resistance narrow 'V' shaped valley is developed.

After the removal of hard layer the denudation process acts on the softer layer (clay stone) and develops wide valley. But this change in landform for short duration. Again the denudation process acts on the hard layer and forms the previous narrow 'V' shaped valley. According to him the shape of landform does not change with respect to time but with rock resistance. This he proved again that land form is not dependent on time factor.

Hack considered that there is an equal rate of erosion on summit, scarp, valley bottom etc. So the shape of land form is not changed only the height is reduced. So one ridge and valley topography can retain itself always and he considered it as the ultimate shape of landform.

Criticism:-

1. Hack did not consider the influence of base level in landform development. The shape of landform will be changed as the height of the land is reduced to base level, because the energy of the erosive power will be least or finished and instead more deposition will start.
2. Hack assumed an equilibrium condition in case of every land form. But it is noticed that the state of equilibrium is rare. All landforms are changing to reach the equilibrium state.
3. He ignored the influence of the processes which were operating in the past to give rise to the present land form.
4. It has been proved that time has an important impact on landform development.
5. He explained only the rock resistance factor in landform development and other factors are neglected.

□ 1.5.5 EPISODIC EROSION MODEL BY SCHUMM AND CHORLEY (DYNAMIC META-STABLE EQUILIBRIUM)

The dynamic equilibrium concept of J.T.Hack is the combination of Progressive decaying

concept of Davis and Steady State Equilibrium Concept of Gilbert. If the concept of Threshold condition which brings catastrophic change in land's character is associated to the concept of Hack, we get the idea of Dynamic Metastable Equilibrium proposed by Schumm and Chorley.

If the height of a landform which is experiencing a positive or negative feed back guiding towards an equilibrium suddenly increases, and the progress towards equilibrium is suddenly disturbed, those episodic upliftments are called as 'Geomorphic Thresholds'. Based on this concept, Schumm and Chorley proposed the idea of Dynamic Metastable Equilibrium concept in 1975.

According to Schumm most of the Landform Evolution Theories are too much generalised and the complex nature of landform evolution can not be described by these. If a landform evolution is viewed in the light of long cyclic time, the episodic nature of detail evolution of land form is obscured but if the time duration is reduced to graded time the nature of positive and negative change of landform is well illustrated. The short duration period of graded time and steady/ instantaneous time is the basic pre-requisite for Dynamic Metastable Equilibrium concept.

According to schumm the evolution of landform is composed of few episodic events and not like the continuous evolution towards an equilibrium proposed by Davis.

The concept of Dynamic Metastable Equilibrium is composed of the idea of **Complex Response and Geomorphic Threshold**.

Complex Response

If a river basin experiences rejuvenation, the entire basin do not respond to it instantly in the same way but actually down cutting starts at the lower end and deposition may occur at the upper any portion of it. The river basin responds to this rejuvenation by erosion at one part and deposition at elsewhere. This type of mixed response to a particular event or change on the factors of landform evolution is called 'complex response'.

Geomorphic Threshold

A drastic change in the factors of landform evolution may bring the manifold change in the rate or intensity of erosion and such changes are known as **Geomorphic Threshold**. One example may help in understanding the idea. If a deposit lies within the channel, the slope in front of the deposit will increase till it is broken or destroyed by the increased energy on the steep slope developed at its front. This geomorphic threshold help in change from positive to negative feedback or vice versa.

According to Dynamic Metastable Equilibrium concept if once rejuvenation starts in a

drainage basin, the entire basin may not start down cutting but responds in a complex way. The down cutting starts at the mouth at first and then this condition is gradually creep towards the source. If there is down cutting at the mouth, there may be deposition at other part. Gradually the renewed erosion will be effective at the upper source region, and then deposition starts at the lower reach. Thus the whole system does not respond to a particular change in a same way but actually in a complex way and the evolution phase is composed of alternate positive and negative feedback and thus the process is ultimately guided towards a new equilibrium.

□ 1.5.6 APPLIED GEOMORPHOLOGY

o Introduction

The application of geomorphological knowledge for the solution of societal problem and also to other discipline is termed as the applied geomorphology. During the time of second world war geography began to loose its existence due to misuse of geographical knowledge. Since then it took a new path and devotes itself in the favour of social welfare and thus the concept of applied geography become popular.

o Objective

The basic aim of this part of the module to make the learners acquainted with the probable sections of society, environment, economy etc where the knowledge of geomorphology can perform a positive role in decision making thus making the subject socially relevant. The application of theoretical knowledge for the solution of social problems like water scarcity, inaccessibility etc. environmental problems like flood, landslide soil erosion etc. is the need of time and by going through this chapter the learners can take part in the development processes.

□ 1.5.7 APPLICATION TO HYDROLOGY

1. **Hydological problems in karst region** are best understood by the understanding of geomorphological condition of the region . To know the available water resource one has to analyse that in which condition or stage of cycle the region belongs to. In youth and old stage, the surface water is available . But in mature or late youth stage there will be no surface water . So to achieve the water resource ,intensive study is necessary and at this point, the knowledge of geomorphology is inevitable.

In Indiana, a dam was built on limestone having many solution channels and ultimately all the water of the dam site reservoir has percolated down and afterwards, leaving that site, the dam and reservoir were reconstructed on the non- permeable rock.

A. Ground water in glaciated areas:-

In glaciated areas ground water is one of the important source of water resource. So to

identify the probable and potential ground water reserves the knowledge of geomorphology is necessary. The glacial deposits e.g. outwash plain, till deposits, buried glacial valleys, periglacial valley etc are the import source of ground water resource. To identify the glacial valley study of landform, geology, soil etc. with the knowledge geomorphology and a background of Remote sensing are necessary. The water holding capacity is also another important factor regarding hydrology. Sand has more water holding capacity for more pore space whereas the clay has less water holding capacity. The proper analysis of glacial deposits on the surface with probable subsurface expansion can help in identification of water bearing layers underground.

□ 1.5.8 APPLICATION TO ECONOMIC GEOLOGY :-

- o **Surface expression of ore bodies:-** Some minerals can be identified by their surface expression as high land, low land, sheared, breccia etc. location of lead-zinc may be identified by highland e.g. Brokenhill of Australia. Location of calcite may be identified by low land e.g. at Qatman (Arisona). Shinkage of ore bodies during oxidation gives rise to mireralisation slump e.g. clay minerals and limestone etc in Sierra Majoda, in Mexico. Iron are may be identified by highlands e.g. near Superior Lake in U.S.A.
- o **Weathering residues:-** Some minerals are found in weathering residues. To search such minerals, geomorphology can play an important role e.g. bauxite, calcite etc, are found as weathering residues.
- o **Study of unconformity:-** Nickel, thorium and manganese etc. have direct relation to the unconformity. Mill and Eyrich (196) studied the relation between plane of unconformity and occurrence of minerals. Out of 68 mineral districts of U.S.A. and Canada –
 - ii. 11 districts show the concentration of minerals at upper part of unconformity along pores and fractures.
 - iii. 15 districts show the location of minerals with igneous intrusion dissecting the unconformity.
 - iv. 12 districts show the location of minerals along non permeable rocks above the plane of unconformity.

Placer deposits:- Placer is the deposit at depression due to greater specific gravity. To identify the location of such placer deposit geomorphological knowledge is inevitable. Placer deposits are of different types such as alluvial, glacial, Aeolian, colluvial placer etc. The alluvial placer is most important among them. These placers are associated with minerals e.g. alluvial placer may associated with platinum as found

in Russia and Columbia. Colluvial placer may be associated with gold as found at the foothill zones of New Zealand. Aeolian placer may be enriched of gold as found in bajada landform of Australia. Gold may also be found at beach placer as in Alaska, California.

II 1.5.9 APPLICATION TO ENGINEERING

Consideration of topography and its geology is obviously one of important input for any kind of engineering work.

- **Construction of roads:** Topography plays an important role in the constructions of roads. Different type of landform or terrain impose varying problem in engineering work. In karst region, the cut and fill process should be followed to construct a road. In glacial Terrain and fluvial deposition, the coarse deposits should be preferred for the road construction because clay leads to plastic flow under pressure. In sloppy region the road follows the contour lines.
- **Sand and Gravel Quarry:** Sand and gravels have commercial and industrial uses. Sands and gravel may be found at flood plain, river terrace, alluvial fan and cone talus, glacial deposits of various types etc. Every location has a particular quality and quantity of sands and gravels and geomorphological knowledge is required to understand these characters.
- **Constructions of Dam:** Dam site selection needs geomorphological knowledge. Bryan (1929) pointed out five requirements for a good reservoir site
 - i. Constricted valley to minimize the construction cost.
 - ii. Broad upper reach to ensure sufficient storage
 - iii. Spillway construction to protect the dam and maintain the flow of the river
- i. Availability of building materials
- ii. Minimum amount of siltation in the reservoir

II 1.5.10 APPLICATION FOR THE MITIGATION OF ENVIRONMENTAL HAZARD

- o **Flood:** As the input of water of a river basin exceeds the amount of out put, the flood will occur.

Frequency of flood may be reduced or stopped with the help of geomorphological knowledge. Because geomorphological knowldges help us to understand the causes of flood,

helps to study the factors controlling the runoff, the lag time etc. which are the function of slope and undulation on the surface. Flood management processes requires the geomorphological knowledge:

- ii. Plantation of trees and bushes is necessary at the upper catchment area to restrict the soil erosion.
- iii. The dams should be constructed on the upper course of tributaries instead on the main river. The size of the dam and reservoir should be reduced.
- iv. Hydal project plan can be introduced to restrict the deforestation by arranging for the supply on alternative energy.
- v. Terrace cultivation has to be introduced to reduce the amount of soil erosion.
- vi. Climatic stations should be established at the farthest part of the basins to know the probable weather condition.
- vii. Check dams should be constructed at the upper catchment to reduce the amount of siltation of the main river.
- viii. The left channels could be connected to main river by flood channels to distribute the excess water during flood.
- ix. Jacketing of river should be stopped. Restricted flood should be allowed to encourage the deposition on flood plain and raising of land at lower reach.

Landslide and soil erosion. Landslide denotes to dislodgement of soil, rock etc. from the upper slope. land slide and subsequent soil erosion is a serious problem specially in the mountainous regions. The causes of land slide can be understood in the light of geomorphology. So to restrict land slide geomorphological knowledge is necessary.

Landslide management processes:

- i. Plantation of trees with extended root and grasses is the utmost requirement to increase the shear strength.
- ii. Metal conduits can be used to release the pore water pressure from the back of retaining wall.
- iii. Construction of retaining wall helps to restrict the easy movement of dislodged materials.
- iv. Terrace cultivation has to be followed to reduce the potentiality of soil erosion.

- v. According To Thomson (1970) geotextile method may be used to stabilise the slope.
- vi. Cementing of the surface can be practiced to reduce the infiltration.

1.5.11 SELF ASSESSMENT QUESTIONS

1. How do the concepts of landform evolution experience change with time?
2. Discuss the basic premises of Davisian Cycle. What are the basic concepts on which the cyclic theory was oriented?
3. Critically discuss the cyclic concept of W. M. Davis. How do you relate this theory with the concept of equilibrium?
4. Define Entwicklung? Describe the processes of land evolution with the relative rate of upliftment and erosion through different Entwicklungs.
5. Describe the king's concept of pediplanation. How do it related to that of Davis?
6. Describe the concept of Dynamic Equilibrium. Describe how do the equilibrium is changed with the change in rock resistance following Hack.
7. Describe the concept of Geomorphic Threshold and Complex Response System. How do these two concepts are associated to Dynamic Metastable Equilibrium concept?
8. How do the knowledge of geomorphology take part in the water supply system at Karst and glacial areas?
9. How can geomorphology find its place in the field of Engineering Geology?
10. What are the avenues in Economic Geology open for Geomorphology?
11. how do a Geomorphologist can participate in the management of Environmental Hazards with his academic skill?

1.5.12 REFERENCES

- Bloom A.L., (1992) *Geomorphology: A systematic analysis of late Cenozoic landform* pp1-12, 99-101, 297-324,.
- Chorley, R.J., Schumm, S.A & Sugden, D.E, (1984)- pp 3-14, 571-582.

H.J.De Blij, Peter O Muller Physical Geography of the Global Environment (4th Edition)599p

Small,R.J., (1992) The Study of Landforms. Pp, 154-182.

Singh,S.(1998), Geomorphology, pp.12-23 , 24-56, 57-88, 554-589.

Thornbury,W.D. (1985)-Principles of Geomorphology

Wooldridge, S.W. and Morgan R.S. (1988) An outline of Geomorphology, pp 173-193

Strahler,A. and Strahler,A. (4thEdition) Modern Physical Geography 638p.

Strahler,A. and Strahler,A. Introducing to Physical Geography 638p

PART-I
Paper - I, Module - II
OCEANOGRAPHY AND HYDROLOGY
Unit – 01

MODULE STRUCTURE

- ☐ 2.1.0 : Introduction
- ☐ 2.1.1 : Objective
- ☐ 2.1.2 : Study of th Continental and Oceanic Crusts
- ☐ 2.1.3 : Theories of the Origin of Ocean basins
- ☐ 2.1.4 : Ocean Waters— Salinity and Temperature and Chemical Compositions
- ☐ 2.1.5 : Distribution of Salinity— Horizontal Distribution
- ☐ 2.1.6 : Self Assessment Questions
- ☐ 2.1.7 : References

Paper - I, Module - II
OCEANOGRAPHY AND HYDROLOGY
Unit – 01

□ **2.1.0 INTRODUCTION :**

Oceanography and Hydrology are two major aspects of physical environment. The present unit of the Module-II includes physical setup and geological history of the ocean basins. The physical and chemical compositions of ocean waters are also included in the unit to know their role as environmental factors.

□ **2.1.1 OBJECTIVES :**

- i) To study the diverse composition of continental and oceanic crusts.
- ii) To identify the possible origin and permanency of the ocean basins.
- iii) To know how ocean waters are diverse in physical and chemical compositions.

□ **2.1.2 STUDY OF THE CONTINENTAL AND OCEANIC CRUSTS :**

Δ **Drift and arrangement of land and sea :** Wegener postulated to main directions of movements : towards the West and towards the Equator. "The forces invoked by Wegener to produce the drift were differential gravitational forces acting upon the protruding blocks of sial". The relation of the centre of gravity and the centre of buoyancy of the moving mass is the cause for the equatorward drift of continental block of sial. As moon was then rotating faster, the tidal forces of sun and moon were responsible for the westward drift of the crust.

According to him, the carboniferous period is supposed to be the beginning of the major drift process on the earth, before which the Pangaea and Panthalassa formed two major land and sea units. Though the central and the western portion of 'Pangaea' was practically covered by the shallow seas, yet the rest of the area was free from encroachment of water. In late Paleozoic era or early Mesozoic era the whole mass broke and drift started. The Jurassic period witnessed the sundering of the southern and western parts of this great land mass. At the same time, the rifting and drifting apart of the Americas towards the west and away from each other caused the birth of Atlantic Ocean and further, the part of Gondwanaland moved northward thus creating the Indian Ocean. By Pliocene time, the distribution of land and ocean had come into being with the only exception that the pole was not at its present position. The North Pole in the Silurian period was in the lat. 14°N and long. 124°W; in Carboniferous period in lat. 16°N and long. 147°W; and in Tertiary period, in lat. 51°N and long.

150°W. The creation of oceans and separation of continents resulted in a tension which, in the opinion of Wegner, was the cause for the formation of great rift valleys of Africa, Asia and also Rhine river valley in Europe.

The evolution of the Pacific and other oceans has been dealt more elaborately by Alex. L. Du Toit who, like Wegener, also conceived the Pacific as twice its dimension in the beginning. Its history is marked by three-sided movements of the earth crust, of varying degrees and magnitude, the result of which was the shrinking in the size of the ocean. During the westward drift of the Americas which is supposed to be due to lesser opposition from the foreland of the Pacific, a marked regular advance of North America is observed. Hence the resulting Eastern boundary of the Pacific is comparatively straight and less complicated. The movement of Asian and Australian blocks towards west resulted into the lagging behind of the continental masses. Thus the western part of the Pacific region became island strewn. The general asymmetry of the two Pacific coasts is thus accounted for by difference in degree of mobility, volcanism and the metallization agencies. In the south the Antarctica, failed and lagged behind the advances of South America and Australia. The result of this is the general stability of the Southern Pacific Ocean.

The general outline of the Atlantic basin was formed mainly due to the westward drift of Americas.

The Atlantic Ocean, equivalent to half of the Pacific Ocean, makes S-shaped outline, with the westward bulges of the coast of African Sahara, and the recession of the coast of South America in the form of Caribbean embayment. The North American block is believed to have advanced slightly more than South America. Thus, while Cape Sao Roque projects outward, the Gulf of Guinea recedes in the same direction.

As regards the mid-Atlantic Ridge (viz. Dolphin-Challenger Rise), the subsurface rise in the centre of the Atlantic, extending from North to South, which supports the volcanic islands of Iceland, Azores, Ascension, Tristan da Cunha and Bouvet, various theories have been advanced. According to Wegner the ridge is regarded as the former land surface left behind when the Americas moved in the westward direction off the coast of Europe and Africa. Further, the creation of the Arctic Ocean and the North Polar Sea is supposed to be due to the equator-ward drift of the land masses. The actual place of rifting between Arctic and Atlantic is marked by the existence of 'Poseidon' depression as called by Schuchert. Unlike the Antarctic basin, the Arctic basin at no time remained a stable region. During every epoch it was crossed over by synclinal depressions and marked by ejections of basalt, especially during the Cretaceous-Tertiary period.

The Indian Ocean, slightly lesser in extent than the Atlantic Ocean, also consists of a medial rise, extending from Gauss Berg in the Antarctica through St. Paul, New Amsterdam, Chagos, Maldives

and Laccadive islands. The ridge and basin formation of the Indian Ocean owe their origin to the sub-parallel slices of Gondwanaland that had rifted and drifted apart, but without complete severance. As a result, India moved away from Africa. The southern region of the Indian Ocean in the proximity of the Antarctica is formed due to the westward and equatorward drift of the land. This area is comparatively stable in many respects.

The principle of Isostasy as such does not clarify in particular, the ocean and land arrangement, but it accounts for the sustenance of the depressions and the elevations, once formed on the earth's surface. It also takes into account the mechanical stability or balance between the upstanding areas and low basins on the rotating earth, due to adjustments brought about by the materials of which they are formed and the level of compensation beneath it. In the upper part of the crust the density and the gravity of matter varies according to the rock formation, beneath which, after a certain given depth, a level of compensation exists, where density is uniform throughout. The density, as such, results in inverse proportion to the height of the area. In other words, the higher areas or mountains or land surface in general formed of rocks of lesser density, and hence are areas of higher elevation, as compared to depressions, lower plateaus area or oceanic basins, in general, are constituted of denser material. The continental elevations are supposed to be of lighter density sial rocks ($Si=Silica+Al=Aluminium$) in contrast to heavy density sima rocks ($Si=Silica+ma=Magnesium$) of the ocean bottom. Each area exerts an equal amount of downward pressure at the level of compensation. As soon as the structure above changes its form, the downward pressure is consequently adjusted. The idea of isostasy, as applied to Drift Hypothesis, helps to interpret original breaking and drifting of Pangaea, formed of Sial rocks on the heavier crustal Sima rocks forming the ocean floor. The hypsographic curve of Krummel shows the relation between the deeps, deep sea plain, continental slope and shelf in complete harmony with the isostic principles.

Very recently Umbgrove (1947) and Vening Meinesz (1945), have put forth the idea that just below the continents, the concentration sub-crust takes place due to the convection currents, which originate beneath sima layer of the ocean basins and flow in the form of sub-crustal currents, transferring magma from one end and piling up beneath the continents. This accentuates the difference between the relief features of the earth crust, but still keeping the isostatic balance between the two.

Δ ORIGIN AND PERMANENCY OF THE OCEAN BASINS :

The omnipresence of changes in nature has made the problem concerning the origin of the ocean basins and land and water arrangements rather difficult. Due to great antiquity of the earth (2 billion years), its origin and the particular arrangement of ocean bodies and continents render doubts regarding the authenticity of any single idea which attempts to prove the origin of present distribution

of land and sea. The arrangement of these oceanic basins and continental nuclei bear a definite geometrical arrangement due to which various theories have been put forth to explain their origin. The broad features of the earth, i.e., land and sea, are termed as 'Relief features of the first Order, by R. D. Salisbury, on which the relief features of second order (plains, plateau and mountains) and the third order (arid, glacial and river, topography) are formed.

Δ Distributional Pattern of Land and Sea

Before taking up the problem of origin of land and oceans, the following general pattern of the major structural units on the face of the earth, has to be studied :

- 1) a) The continents of the Laurasia group.
b) The intervening North Atlantic and Arctic oceans.
- 2) a) The continents of the Gondwanaland group.
b) The intervening South Atlantic and Indian oceans.
- 3) The oceanic basin of the Pacific covering the above groups like a huge girdle.

□ 2.1.3 THEORIES OF THE ORIGIN OF OCEAN BASINS :

Δ Nebular Hypothesis : Nebular Hypothesis was expounded by the German Philosopher Immanuel Kant in 1755, who believed in the earth and our solar system as condensed from a rotating diffuse cloud of primordial matter of dust and gas. Laplace, while avoiding Kant's idea regarding the primordial matter of Nebula, assumed the sun as the rotating disk of hot gaseous matter of radius greater than the distance from sun to the outermost planet. The constant cooling, contraction, increase in speed of rotation and consequent increased centrifugal force caused the separation of rings from the equatorial circumference, which ultimately aggregated to form the planets.

The same process resulted in the origin of moon as a satellite to our earth. In course of time moon and earth cooled down, resulting into rocky crust. According to the assumption, the superficial layer, however, cooled rapidly, but the internal mass of earth cooled and contracted slowly, leaving the upper crust unsupported, which later sunk down and fitted on the lower layer, thus resulting into crumples, folds and deformations. The higher areas thus became continents and lower areas gave birth to oceanic depressions in which the water vapour previously held in the atmosphere rained and formed the primitive oceans.

This theory was later on discarded on the basis of the fact that the increased angular momentum of all planets caused by increased rotation is not sufficient enough to break up the nebula.

Δ Planetesimal Hypothesis : The concept of formation of planets by the piecemeal accretion of

smaller particles, called Planestesimals, scattered through the space was proposed by T. C. Chamberlian (1904) in the form of Planetesimal Hypothesis. Due to interaction of sun (eruptive) and another star (projectile of suitable mass in the form of tidal disruption) ejection of small jets of solar matter became nuclei for different planets which later on grew by constant aggregation of planetesimals.

The oceans began to form in depressions created by volcanic action, and the condensed vapour in atmosphere rained to fill up the oceans. Further, evolution of the relief features is accounted for by the weathering process. The arrangement of oceans and continents is explained by a series of meridional and oblique yield tracts, which divide the earth into quadrilateral areas. These yield tracts in the growing earth are supposed to be the result of a mechanism which allows the earth body to change its shape according to the change in the rotation. The stresses are greater near the centre and least at points near the poles. The dislocation, fracture of the surface, and the consequent rearrangement of the material of the earth crust causes the stressed to be relieved.

According to the yield tract idea, the oceans are arranged in Pairs: North and South Atlantic, North and South Pacific and the Indian Ocean, corresponding to the depression, occupied by the Mediterranean, the Baltic, the Black, the Caspian, the Red Sea and the Persian Gulf in the north. The symmetry between the two is that the northerly ocean areas lie obliquely to the west with respect to southerly oceanic depressions. The areas delimited by yield tracts are the loci of continental plateaus.

Δ Tetrahedral Hypothesis : A tetrahedron is a solid body which has four equal plane surfaces, each of which is an equilateral triangle. It has four solid angle where three edges or di-hedral angles meet. The edges of the tetrahedron represent those parts of the surface which are farthest from the centre.

The concept of a tetrahedron was taken into consideration while keeping in view the contraction of a hot rotating spheroidal matter, that is, the shrinking of the earth. The deformation of the sphere during evolutionary cooling and contracting, resulting in an irregular, collapse of crust of the earth on a shrinking interior, forms a tetrahedron body. In spite of the obvious changes in the volume of the sphere, the surface area in a tetrahedron remains almost the same.

The concept of a tetrahedron is the close conformity with the shrinking sphere.

Lowthian Green (1875) initiated the Tetrahedral hypothesis to attempt an explanation of the following facts : the marked concentration of land in the northern hemisphere and water in the southern hemisphere combined with the reversal of this contrast in the polar regions; the antipodal arrangement between land and sea; and the southerly extension and tapering off of the three continental blocks of South America, Africa and Australia. The three solid angles of tetrahedron correspond with the three ancient land areas of the globe, which are the Laurasian or Canadian Shield in the Hudson

by area, the Baltic shield in the Baltic Sea covering tracts in Finland, Sweden and North West Russia and Angara shield in East Central Siberia. The pivot of the tetrahedron covers the Antarctic continent, approximately of same geological antiquity as the above shields. Further, the antipodal arrangement of land and sea is explained by the fact that the four flat faces of a tetrahedron are occupied by the oceanic basins, opposite which are the edges, forming the continental areas, converging and tapering towards the south. Arctic Sea is formed on the northern flat face around which the continental margins are clear.

The tetrahedral theory, though consonant with the views of cooling and shrinking globe, is criticized on the ground that the final figure of a tetrahedron does not correspond to the figure of equilibrium. The present land and sea arrangement shows the maximum land area at 66° north latitude (80.1% of the land is above sea level here), and 43.6% at 30° north latitude. Contrary to this, the tetrahedron theory requires the maximum land area at 30° north latitude. Similarly, the south tapering land of the continents should fall regularly at an interval of 120°. But in reality, the intervals between continents from east to west are 115°, 115° and 90° (continents taper at 20°E, 135°E and 70°W longitudes respectively).

Δ Other concepts : Lap worth assumed the formation of the continents and oceans due to folding on the earth crust on a very large scale. The anticlines of composite nature formed the continents and the synclines formed the oceans. The continent of America, according to him, is a broad arch with a median sag is place of central plains. Similarly, the Atlantic Ocean is a vast down fold with a central upfold of Challenger – Dolphin Ridge. Love (1907) modified the above idea with the help of a scientific basis, that is, the centre of gravity is not the geometric centre of a body. Jeans and Sollas advocated the idea of earth's colloing and deformation in the shape of a 'pear'. The birth of the moon and the formation of the resulting Pacific depression are considered to be accidental. The stock of the pear is assumed to represent the small islands, antipodal to stock of the pear is assumed to represent the small islands, antipodal to the land hemisphere and the oceans are represented by the areas around the girdle of the pear. According to Jean and Sollas, African continent is the centre of the land hemisphere situated antipodal to the water hemisphere and the Pacific Ocean. The Atlantic Ocean is supposed to be a later stage break in the continuity of land hemisphere.

The above theories in general bring to light one fact in common, that is, the permanency of the ocean basins. Sir Charles Lyell's concept of 'Unformitarianism' and James D. Dana's concept of 'permanency of ocean basins', also point towards the relationship between the origin of the earth, its consequent cooling, and the final formation into basic land and sea distribution which changed very little in subsequent geological times. Later on, due to evidences in the field of geology, palaeoclimatology, geophysics and palaeontology, the concept of permanency of ocean basins gave birth to

the idea of horizontal displacement or drift of blocks of continents on the sub-crustal floor of Sima. In other words, the outlines of the oceans are the result of constant evolution and displacement of the crust during geological era.

Δ **Drift Hypothesis** : As early as 1620, Francis Bacon presented the concept of drifting of earth crust which resulted in the arrangement of oceans and land distribution. Afterwards, Buffon (1780), Young (1810), R. Owen (1857), A. Snider (1859), E. Wettstein (1880), O. Fisher (1882), C. B. Warring (1887), W. H. Pickering (1907), were the main promoters of the concept. But the real advocates of continental drift theory were, F. B. Taylor in 18908, H. B. Baker in 1911 and A. Wegener in 1912, who convincingly put forth, geological, physical and biological evidences in support of the drift of the surface layer, and the formation of oceans in between. This idea of drift is also responsible for an interpretation of various other problems, such as tectonic forces, ice age, animal and plant migration, etc. on the surface of the earth.

Δ **Taylor's theory** : The idea of continental drift as developed by F. B. Taylor in 1910 was originally put forth to explain the formation of greater mountain ranges of tertiary era. As the relief features of this age were of greater magnitude, he imagined the source of energy for drift to be external. In late Cretaceous times due to the creation of moon, and its tidal force, the rotation of the earth increased, resulting in drifting of land towards equator from north and south poles, and towards west.

The formation of Alpine-Caucasus-Himalayan mountain belt supports the idea of equator-ward continental drift. This caused stretching of the earth crust in north and south, resulting in bays and seas, such as Baffin Bay, Labrador Sea, Davis Strait in the north and Great Australian Bight and Ross Sea around Antarctica in the Southern hemisphere. Arctic Sea separating Greenland and Siberia is the result of equator-ward stretching of the crust. South America, Taylor suggested, was separated from Africa in the Permian time. This gave birth to the Atlantic Ocean. The evidence to this fact is brought by the remarkable similarity of West Coast of Africa and Europe, to the east coast of the North and South Americas.

There are severe objections to this theory. The first and the foremost is against the assumption that the drift was due to the external force of moon's tidal force and the continuous increase in the rotation of the earth. This fact seems quite vague and contradictory as the moon's force, if sufficient to drift the continents, can also put brake on the rotation due to friction. Secondly, this concept is also silent in explaining the early orogenic cycles prior to the Cretaceous period in which the moon took birth. But in the opinion of Arthur Holmes, "Taylor must be given credit for making an independent and slightly earlier start in this precarious field. His immediate object was to account for the distribution of mountain ranges."

Δ **Wegener's theory** : Alfred Wegener, a Professor of Meteorology and Geophysics in the University of Graz, in 1912, after rejecting the idea of permanency of oceanic basins and creation of earth's relief features by sinking of the earth crust, put forward the concept of 'Displacement Hypothesis'. The three successive editions of his book, in 1915, 1920, 1922, and the English version of the last edition of 1922, viz., *The Origin of Continents and Oceans* by J. G. A. Skerl in 1924, put forward the thesis of Continental Drift.

The theory of drift was proposed by Wegener in order to explain the land and sea distribution as well as the larger climatic aberrations of the past. For this he gathered evidences from the works of geologists, palaeoclimatologists, palaeontologists and geophysicists. The present arrangement of continents and oceanic basins is regarded as a result of the fragmentation by rifting and drift apart of the original continental mass known as '*Pangaea*' surrounded by a universal ocean '*Panthalassa*'. Wegener, like Suess, supposed that the continental mass called '*Pangaea*' in the Carboniferous time (due to lack of evidence he does not take precarboniferous conditions) was a continuous layer of 'Sial' spread over the layer of Sima which formed the surface of the primaeval Pacific Ocean (Panthalassa).

□ **2.1.4 OCEAN WATERS – SALINITY AND TEMPERATURE AND CHEMICAL COMPOSITIONS :**

The study of the temperature of the oceans is the subject-matter of Marine Meteorology which deals with that portion of the atmosphere and its interaction with the sea which overspreads the great water masses.

Δ **Source of heat and isolation on oceans** : Sun is the original source of energy and heat to the earth, in the absence of which the oceans would freeze and the temperature everywhere on the globe would be that of the cosmic space (a degree or two above the absolute zero). It is a flaming gaseous mass whose surface temperature is estimated to be 10,300°F and the central part the temperature rises to 50 million degrees F. The diameter of the sun is 100 times greater than that of the earth and its volume is about 10,00,000 times that of the earth.

From such a sun a perennial radiation of its energy is found towards the earth. According to Trewartha, "The radiant energy received from the sun, transmitted in a form analogous to short waves (1/250 to 1/6700 mm in length) and traveling at the rate of 1,86,000 miles a second is called solar radiation or isolation." Byers define it as "the rate at which the radiant energy is received on a horizontal surface."

Other minor sources of energy which heat the ocean are the inner heat and the increase of energy by pressure factor. But these sources make minor additions to the total amount of energy and heat

received by the oceans from the sun.

Process of Heating : There are various processes of heating the ocean water, i.e., by absorption of radiation from the sun, convection of heat through the ocean bottom from the interior of the earth, transformation of kinetic energy into heat, heating due to chemical processes, convection of sensible heat from the atmosphere and condensation of water vapour.

Δ Absorption of radiation from the sun : The incoming solar radiation waves are of three types : (A) *Short wave radiation*. These are high temperature radiation waves, which are not visible to human eyes, and pass through the lower part of the atmosphere without being absorbed. The length of the waves is less than 0.4μ [1 micron (μ) = .001 mm] and it comprises 6% of the total solar energy. (B) *Medium waves*. These are visible light waves between 0.4μ to 0.8μ length and comprise 52% of the solar radiation. (C) *Long waves*. These are low temperature radiation waves and are vastly absorbed by the surface. They comprise 42% of the total solar radiation. Johnson mentions 6%, 49% and 42% values for the above three waves respectively.

The atmosphere acts as the green house to the solar radiation, i.e., it is transparent to solar radiation but non-transparent to earth's radiation. The fraction of the solar energy that is received on the surface after being depleted (by absorption and scattering) in the atmosphere, is again reflected back into the space. This over-all reflectivity is known as the *albedo* of the earth. It is 42% in general, meaning thereby that out of 100 solar radiation only 58% is left to be absorbed or scattered by the land and the sea. According to Moller the reflected energy is only 35% (2% reflected from the earth's surface, 6% from atmospheric particles, and 27% from cloud tops).

Besides latitude, the seasonal changes also affect the angle of the rays at a certain place. Throughout the year the angle of the ray at a fixed latitude does not remain the same, but changes according to the seasonal shifts in the sun's position. Hence in the month of June, the northern hemisphere receives greater isolation as compared to the southern hemisphere and each latitude receives greater isolation in the month of June than in January. The reverse is the case in the southern hemisphere where January is the warmest month.

The second factor which affects the amount of isolation is the duration of day-light, which differs due to the inclination of the earth's axis. Hence equal hours of day-light are not found at each latitude. Higher we go in the latitudes longer will be the day. At equator it is 12 hours day-light, whereas at 49° latitude it is 16 hours and at $66\frac{1}{2}^\circ$ it is 24 hours day-light. The normal deduction in this case would be that greater insolation will be received in the higher latitudes. No doubt the outer atmosphere at poles receives greatest insolation in July and January, but the total insolation to reach the earth is less than that of the equator. The main causes for this are the slanting rays bringing less energy in

bigger area, longer summer days balanced by shorter winter days, and lastly, slanting rays passing through greater passage of atmosphere where the energy is depleted.

The depletion of the solar radiation in the atmosphere is due to three different processes : direct reflection from the surface into the space; diffuse reflection of the radiation waves (comes back to earth again) by particles larger than the wave length (mainly dust and cloud droplets); selective scattering of short waves by particles smaller than the wave length (e.g., scattering of blue colour waves which have very small wave length by very fine dust particles); and the absorption of selected wave lengths, mainly the larger ones, by water vapour, carbon dioxide and ozone.

Another fact that determines the reflection is the cloud cover which in relation to land is greater on water surfaces. Hence, out of the same amount of energy that is received on the land and the sea, reflection is greater on the sea surfaces and the turbidity factor is greater on sea than on land. On an overcast day, when all the incoming radiation, the amount of reflection from water surfaces would be 17% according to Schmidt, and 8% according to Powell and Clarke. The above values are mostly true when the sea surface is smooth. If strong waves and currents are found then the percentage of reflection specially at low altitude of sun would be slightly greater. This feature is particularly marked in regions of high latitudes.

Δ **Convection** : According to Geikie, the temperature of the earth increase 1°F for every 64 feet of descent. This results in the increase of temperature at the bottom of the ocean due to the nearness to the centre of the earth. The convectional currents thus originating due to internal heat of the earth account for heating of lower layers of water which, as we come upward, becomes insignificant. It has been estimated that the heat flow by convectional currents amounts to between 50° and 80 g.cal/cm²/year, that is, 0.95×10^{-3} to 0.152×10^{-3} g.cal/cm²/min. This flow is insignificant in comparison with the solar radiant heat received and absorbed by the sea surface. Further, the movement of oceanic water also reduces the effect of bottom heat and makes it negligible due to the quick transfer from one region to another. It amounts to one thousandth part of the solar radiation received at the surface. Its role in the heat budget of the ocean is difficult to estimate due to the absence of very accurate thermometers and difficulty in deducting the effect of pressure of upper layers of oceanic waters in the generation and increase of temperature at the bottom.

Δ **Kinetic energy** : Due to the surface wind and the tidal currents, the energy generated in the form of stress on the sea water is capable of raising the temperature of the sea water locally. The contribution of heat amounts to one-ten thousandth part of the solar radiation and is very insignificant in the total amount of the heat of oceans. But if the tidal currents are strong enough in a small shallow enclosed water body, then the heat thus generated can be of some local importance. For example, in

the Irish Channel, the dissipation amounts to 0.002 g. cal/cm²/min. or 1050 g. cal/cm²/year, which results in an increase of 0.2°C temperature on an average (the depth of 50 m remaining the same throughout the year).

Δ **Process of cooling** : Three types of heat transfers helping cooling of the surface are recognized by physicists and meteorologists, viz., conduction, convection (horizontal as well as vertical) and radiation. Both conduction and convection depend on the existence of a material medium either solid, fluid or gaseous, whereas radiation does not require material medium for transfer of heat to the space.

Δ **Isothermal lines and Horizontal temperature distribution** : Each isotherm is a boundary, delimiting an area where the temperature has a certain mean value. For showing the summer and winter temperature conditions of the sea the months of August and February are taken into consideration.

Represents the sea surface isotherms in February. The most conspicuous factor which is noted here is crowding together of the isotherms in the south of Newfoundland, on the western coast of European continent and North Sea, and widening out to make northward bulge on the coast of Norway. This fact is due to the cold water of the Labrador current which flows southward along the American coast and makes the temperature lower than the average temperature for those latitudes. The warm water of Gulf Stream moves towards west coast of Europe and brings higher than the average temperature for those latitudes. Hence within a short distance from the American coast a change of sea surface temperature from 20°C to 0°C is met, whereas a wide extent of ocean water on the Eastern Atlantic has to be crossed to meet such a great change. In the South Atlantic the isotherms though do not show such clustering, yet while extending east and west they make an angle with the latitudes. In the western part, isotherms bulge towards southwest (due to warm Brazil current), but in the eastern part they bulge towards north-west (due to cold Benguala current). Hence on the same latitude higher temperatures are found on South American coastal waters than on African coast. Southward, due to constant prevailing west wind drift, the isothermal lines observe a parallelism with latitudes. A comparative view of asymmetrical distribution of temperature and isotherms of the South and the North Atlantic brings out the fact that 5°C isotherm reaches 70° latitude in the north but it does not go beyond 50th parallel in the south due to the obvious fact that the gulf stream is more powerful and extensive than the Brazil current. Further, from east to west the Atlantic shows a considerable difference in temperature, e.g., 0°C temperature is found off Labrador between 50°N and 60°N latitude, whereas 9° to 13°C temperature is found on the European side between the same latitudes. In Central Atlantic the isotherms are much disorted due to unstable weather and sea conditions.

The temperature conditions in the marginal seas vary according to their latitude and location. The seas in the lower latitudes, e.g., Mediterranean, show higher temperature than the neighbouring

areas of the Atlantic and the seas in the northern latitudes like epicontinental seas of the Baltic and the Hudson Bay are colder than the adjacent Atlantic.

In the North Pacific Ocean there is a perfect parallelism between isotherms and latitudes, though on the coast of North America the isothermal lines make minor loops towards north due to the warm Kuroshio current and along Japan coast the crowding of isotherms is slightly intensified due to cold Oyashio current coming from north. The western side of Equatorial Pacific shows high temperature due to powerful south equatorial current whereas in the eastern part a region of lower temperature crops up due to the effect of cold counter current and the upwelling along west of south America. This temperature characteristic separates east and west Pacific. Further, in the South Pacific, the isotherms have a tendency to run parallel with slight bulges along the west coast of South America due to the cold Humboldt current flowing northwards. The zero degree isotherms line runs parallel to the rim of the Antarctica.

In the Indian Ocean the isothermal lines of 25° , 27° and 28°C cover the central portion of the ocean. In the South, the pattern of isotherms is similar to that of the Pacific, that is, almost parallel to the latitudes, with a loop of isotherms created at the Cape of Good Hope due to the cold Agulhas current. On the coast of North Africa the isotherms make a bulge southward because of a cold current flowing south westward from Cape Guardaui, the north-east corner of Africa, along the coast which keeps down the sea surface temperature. The same isotherm bends northward in the Arabian Sea, and slightly south-ward in the Bay of Bengal due to the effect of monsoon drifts. The enclosed sea shows somewhat different conditions. Red Sea has high temperature in the south due to mixture of open ocean water but in the north the cold effect of the land creates low temperature. The temperature in Persian Gulf is lower than in Indian Ocean due to contact with the cold land area.

Summer isotherm conditions in August bring out a marked change in certain parts from that of February. The arctic Sea surface in this season is partially free from the ice. Hence all the isotherm lines make an upward loop in the Davis Straits where the temperature is 5°C . Consequently, the great upward loops in isotherms caused by the Gulf Streams on the Norwegian coast become insignificant. In general the isothermal lines in the North Atlantic show a marked northward shift in August. The isotherms of South Atlantic do not differ so much in February and August due to preponderance of water and greater mixing.

In the Pacific, particularly in the southern part, there is a closer approximation to parallelism of the isotherms with the lines of latitudes than in the Atlantic regions. In the west Pacific, there is a large area of sea in the Austral-Asiatic region where temperature rises over 28°C . This is due to the fact that the westerly flowing equatorial current transfers warm water towards the west Pacific and

piles up against the Austral-Asiatic Islands.

As regards the Indian Ocean the highest sea surface temperature of 28°C is recorded in the Arabian Sea and Bay of Bengal area. In the western part of the ocean isotherms become almost parallel to the lines of longitudes with a decreasing temperature towards the African coast. In the southern Indian Ocean the isotherms are parallel to the latitudes. The enclosed seas, e.g., Red Sea and Persian Gulf, attain a maximum temperature of 30°C and 33°C respectively due to the contact with the warm land areas. Thus the enclosed seas show contrasting temperature conditions in the year, that is, lesser temperature than the open sea in the month of February and higher temperature in the month of August.

From the above analysis of the isothermal maps we can safely conclude that a relation exists between the isotherms and the currents of the ocean. Under the influence of very strong permanent winds which can bring the surface water in the form of a current, from the regions of higher intensity of solar radiation to the contrasting areas situated in the higher latitudes, heat is carried in different sections of the ocean. Caribbean Sea has slightly higher temperatures than the equatorial region due to the piling up of warm water under the stress of trade winds.

Δ Vertical temperature distribution : The sea temperature decreases with the depth, as the sun is the original source of energy. Attempts have been made to estimate the water temperature at various depths ranging from 100 to 3000 and 4000 m. It was found that up to the depth of 100 m the temperature corresponds to the surface temperature, but further lower there is a general lowering of the temperature. It is in this region that the influence of oceanic current is noticeable in creating warm layers in cold water. Normally 90% of the heat that penetrates the surface of the sea is absorbed in the upper 60 feet of water. The sun's rays have no effect below 600 feet, therefore, in spite of movement of water a vast bulk of sea is relatively cold. Hence in general, the stratification in sea may appear as lighter water of high salinity on the top of the denser cold water. Probably 80% of water in the oceans has a temperature permanently below 40°F .

The following facts are marked in connection with the vertical distribution of temperature of the sea.

First, inspite of a general decrease of temperature towards the bottom, the rate of fall is not equal at all the depths.

Secondly, the vertical rate of fall of the temperature from equator to poles is not same as the surface temperature decreases with the increasing latitudes, whereas the bottom temperature remains almost the same.

Thirdly, the surface temperature and its decrease towards bottom may be found to be influenced by the factor of upwelling of bottom water also.

Fourthly, sometimes the sinking of the dense surface water takes place due to convergence with dense bottom layer or with an intermediate water mass.

Fifthly, in equatorial regions an almost reversal of the average decrease of temperature conditions is found.

Sixthly, there are, however, certain exceptional areas with higher temperatures even at greater depth, e.g., the Saragasso Sea, the Red Sea, the Sulu Sea and the Mediterranean Sea, etc.

In the enclosed seas of higher latitudes quite different observations are made. At the surface the temperature is low due to abundant fresh water, while denser water of slightly higher temperature and high salinity moves in as a sub-surface current from outer ocean and remains at the bottom. Hence a versal of temperature conditions is found.

Δ Salinity in Oceans : Among various physical properties of the ocean, salinity is perhaps the most important variable besides temperature and pressure. Commonly defined, it is the ratio between the weight of the dissolved material and the weight of the sample of sea water.

The variations of salinity between 34 and 37½ per mille determines other physical properties of sea water such as compressibility (depending on salinity and temperature which produces density differences), thermal expansion, refractive index and the temperature of maximum density.

Again, salinity introduces altogether new properties like osmotic pressure (pressure depending on the temperature and the chlorinity of water) and electric conductivity in the ocean waters.

Δ Composition of sea water : Sea water is almost a universal solvent and, therefore, it is a complex solution of many mineral substances found in dilute form. It is assumed that the sea water was faintly saline in the beginning and its saltiness has been increasing ever since.

Various methods have been used to determine the amount of salt in the sea. The titration method is based on the constancy of composition of the dissolved matter; the hydrometer method determines the density of the sample of sea water by which salinity is calculated; and the last method is by measuring the electric conductivity or refractive index of the sea water at a known temperature.

Chemical analysis by Dittmar during his 'Challenger' expedition in 1884 reveals the existence of 47 different sea salts. Among all the salts in the sea, seven are the most important.

Different Salt Percentage in Sea Water		
Salts	Amount %	Percentage
Sodium Chloride	27.213	77.8
Magnesium Chloride	3.807	10.9
Magnesium Sulphate	1.658	4.7
Calcium Sulphate	1.260	3.6
Potassium Sulphate	0.863	2.5
Calcium Carbonate	0.123	0.3
Magnesium Bromide	0.076	0.2
	35.000	100.0

According to both Forchhammer and Dittmar, the proportion of different salts remains unaltered in the ocean everywhere. Only the total percentage varies from 33% to 37%.

In addition to the above salts, a few other elements and salts are found in sea water – iodine, fluorine, arsenic, boron, lithium, caesium, barium, strontium, manganese, nickel, cobalt, zinc, lead, silver, gold and radium. But the percentage concentration of these is too small for precise measurement, though the total quantity present is great. It is worth noting that there are thousands of tons of gold and 1400 tons of radium in solution in the sea water.

Δ **Origin of salts in the sea :** The river brings about 5.4×10^8 of the total solids of the ocean.

The quantity of salts brought by the rivers is too small and estimated to be two millionth part of the ocean.

Salts are also added either due to erosion of the earth crust or by some agencies which add minerals from the obscure sources buried deep within the earth, e.g., volcanic ash brings some material – especially chlorine, boron, sulphur and iodine.

Δ **Controls of salinity :** The amount of salt in the oceans is not same everywhere due to varying factors which control the amount and distribution of salinity.

Δ **Evaporation :** Okada, the Japanese investigator has carefully observed the influence of rate of evaporation on salinity distribution. By the process of evaporation in the sea water, the salts are left behind which increases the salinity. On the other hand, salinity is comparatively low in equatorial regions due to evaporation and high relative humidities.

Δ **Precipitation :** It is another factor which leads to diversity in the amount of salinity. In equatorial

regions, inspite of high temperature, the salinity is less because these are the zones of heavy rainfall which reduces the salinity. North and south of this, there is an intra-tropical zone of high temperature and strong north-west and south-east trade winds. Hence there is an increase in evaporation resulting in an increase in salinity.

Δ River water : The amount of salinity in sea water also varies due to the influx of fresh water by rivers. Hence near the mouths of rivers Amazon, Congo, Niger, Ganges and St. Lawrence, comparatively lower salinity is found. This effect is also marked along the coasts in the enclosed seas.

Δ Atmospheric pressure and wind direction : Salinity changes slightly due to winds resulting from difference in atmospheric pressure. There is a prevailing low pressure over Baltic Sea as compared to the North Sea. To compensate it, a great amount of high salinity water from the latter comes into the former whose salinity is thus slightly increased.

Δ Movement of sea water : It leads to a very thorough mixing of the surface water and hence determines salinity. Among these movements are, first, the tidal currents which everywhere surge backwards and forwards in the neighbourhood of the land; secondly, currents and drifts which actually carry warm and saline waters from the equator to the pole, from the western coasts to the eastern coasts, from open seas to the dependent seas; and thirdly, the vertical movement which lifts up water from the bottom to the surface in some regions and causes surface water to sink to the bottom elsewhere. This mixing, when it is permanent and long continued, would lead to an absolute uniformity of composition at the surface.

Δ Periodic variations of salinity : The periodic variations in salinity depend on the difference between precipitation and evaporation as related to the months of the year. Bohncke has computed the periodic variations in salinity in the North Atlantic between $18^{\circ} - 42^{\circ}\text{N}$ latitude to be 36.70% in March, (highest) and 36.59% in November. Sometimes, higher salinity is found at the end of June and minimum in December.

□ 2.1.5 DISTRIBUTION OF SALINITY - HORIZONTAL DISTRIBUTION :

- i) **Latitudinal :** Salinity distribution on seas is primarily controlled by latitude, that is, it decreases from the equator towards the poles. It is, however, not maximum at the equator due to an excess of rain over evaporation, but the region between $20^{\circ} - 40^{\circ}\text{N}$ latitude records the highest average salinity of 36% due to higher evaporation. In the southern hemisphere between $10 - 30^{\circ}\text{S}$, 36% salinity is found.

The coastal waters all along the continents record lower salinity as compared to the central part of the sea due to maximum of fresh water.

- ii) **Regional** : Salinity distribution on the surface of the sea is represented by Isohalines. These are lines joining places of equal salinity. The world distribution of surface salinity in oceans presents a detailed picture of horizontal salinity.

Atlantic Ocean : the average salinity in this ocean is about 35.67%. Between 0° – 5°N it is 34.98% which again increased to 36% near 15°N . In the South Atlantic is similar latitudes comparatively higher salinity is found ranging up to 37.77%. There are two regions of higher salinity north and south of the zone of equatorial lower salinity. The northern region has its centre roughly about 25°N and 40°W while the southern centre is about 15°S and 25°W . Around these circular areas of higher values the salinity decreases to 35% which again is not uniform in north and south.

In the South Atlantic, isohalines are mostly found parallel to the latitudes due to an excess of water over land. Here the 40^{th} parallel is approximately the southern boundary of water having a salinity of 35% or more. In the South Atlantic between 10° – 30°S latitude, saline water is found mostly in the western part of the ocean while cold water upwells in the eastern part. At the river mouths also low salinity is observed : St. Lawrence 31%, Amazon 15%, Congo 34%, Niger 20%, Senegal 34% and Rhine 32%.

In partially enclosed seas of the Atlantic, salinity is dependent on factors other than those prevalent in the open sea.

In the Mediterranean, evaporation is comparatively greater and mixture with open sea water is less. At the Gibraltar Strait the salinity is 36.5%.

Indian Ocean : The observations regarding salinity in this ocean are comparatively less. In the northern part the isohalines run almost across the latitudes because of the duration of the S. W. monsoons. 35% isohalines practically divide the ocean into eastern and western section. The west, however, has higher salinity than the east.

In the partially enclosed seas of the Indian Ocean at head of the Persian Gulf the salinity is 37.0% whereas in the interior it amounts to 40.0%. The Red Sea records perhaps the highest salinity among enclosed seas due to high evaporation and lesser precipitation and mixture. At the head near Bab el-Mandeb, the salinity is 36.5% and near the Gulf of Suez it is 41.0%.

Pacific Ocean : The Pacific Ocean is slightly different from the other oceans due to its unique shape. At the equator, the salinity is 34.85%, whereas north and south of this upto 15°N and S latitude it varies between 34% and 35%.

Inland seas and Lakes : The degree of salinity increases in a lake or in the inland sea due to an absence of any outlet, more evaporation and little addition of river water. This salinity percentage increases with the age of the lake or inland sea.

Δ Vertical distribution : Salinity in the ocean decreases or increases towards the bottom according to the nature of the water mass. Though mainly the atmospheric reactions influence the salinity which in general decreases at the bottom. But an occurrence of cold or warm water mass at depth may result in drastic changes. At the southern boundary of the Atlantic, surface salinity is 33.0 which at 200 fathoms increases to 34.50, still deeper at 600 fathoms it reaches upto 34.75%. At 20°S the surface salinity of 37% decreases upto 35% at the bottom. At the Equator region of Indian Ocean the surface salinity of 34% increases with the depth to 35% due to greater mixture of fresh water at the surface. A map of Gibraltar Strait would illustrate that the rate of decrease on two sides of the strait in the Atlantic and in the Mediterranean are not similar. At the entrance of Mediterranean, 6.50% salinity is found which continues upto 600 fathoms.

□ 2.1.6 SELF ASSESSMENT QUESTIONS :

- a) Explain the nature and composition of continental and oceanic crusts.
- b) Discuss the origin and permanency of the ocean basins with supported theories.
- c) Write short note on : Temperature-salinity diagram;
 Chemical compositions of ocean waters;
 Ocean basins; Oceanic crusts.

□ 2.1.7 REFERENCES :

- R. C. Sharma and M. Vatal (1986). *Oceanography for Geographers*. Chaitanya Publishing Hous. Allahabad. pp. 304.
- J. A. Steers (1937). *The Unstable Earth*. Kalyani Publishers. Ludhiana. pp. 345.

PART – I
Paper-I : Module- II
Unit - 02

Module Structure

- 2.2.1 Air- Sea Interaction
- 2.2.2 Waves:
 - 2.2.2.1 Wave energy :
 - 2.2.2.2 Wave Refraction
 - 2.2.2.3 Wave Breaking
 - ▲ 2.2.2.3.1.Types of Breaker
 - ▲ 2.2.2.3.1.1 Surging;
 - ▲ 2.2.2.3.1.2 Collapsing
 - ▲ 2.2.2.3.1.3.Plunging
 - ▲ 2.2.2.3.1.4.Spilling
 - ▲ 2.2.2.3.2.Quantitative Analysis of breaker
- 2.2.3 Current
 - 2.2.3.1 Shore normal currents
 - 2.2.3.2 Long-shore currents :
 - 2.2.3.3 Tidal Currents
- 2.2.4 Cell circulation :
- 2.2.5 The tides :
 - 2.2.5.1 The energy source :
 - 2.2.5.2 Types of Tide :
 - ▲ 2.2.5.2.1 Progressive wave tides :
 - ▲ 2.2.5.2.2 Rotating tides :
- 2.2.6 Marine Ecosystems :
- 2.2.7 Deep-sea Deposits :
- 2.2.8 Self Assessment Questions
- 2.2.9 References

■ 2.2.1 Air- Sea Interaction

The interaction between air and water is crucial to sustain biological and geochemical activity. There are a numbers of significant interfacial transfer, incorporating energy and mass. If the air water boundary is disturbed by differential motion of the adjoining masses, stress is applied along the interface. Energy is transferred from faster to slower moving masses- usually air to water. Large quantities of energy (10¹⁵J) are generally passed in this way from ocean to land every year. Some of this energy is stored in the wave and some as heat. However, an energy balance must be maintained through commensurate loss (Carter,1988). The height of the wave increases as it comes closure to the land showing high potential energy and is balanced by lowering of velocity i.e. kinetic energy. According to Carter,1988, the increasing height of the wave near the coast is balanced by the decrease in wave length.

The result of interaction between air and ocean gives rise to chemical accumulation of solutes particularly the salts, particularly those of sodium, magnesium, calcium and potassium. The temperature differentials near the water surface helps in evaporation and accumulation of salts in the water. In deep ocean the amount of salt is nearly 35 ppt. where as that near coast is lower due to supply of more fresh water.

This salt, in turn, return the land in the form of wave spray as the wave breaks by wind and then subsequently deposited on the soil and leaves and may cause corrosion of the structures (Carter,1988,1995).

The gas exchange between air and ocean is fundamental to the biosphere because huge amount of gases mainly Co₂ is absorbed by the ocean by photosynthesis. The photic zone or the shallow water is thus very crucial in maintaining life on the earth. The respiration of biotic organisms are also important in such exchange. The gas is also incorporated into the sea near photic zone by wave breaking and become saturated with dissolved Oxygen. The concentration pattern of carbon dioxide is also complex. Excess carbonate (Magnesium and Calcium) is incorporated into skeletal organisms and may be released after death and come back to the surface in the form of limestone. The ocean and oceanic deposits constitute an enormous reservoir of carbon dioxide

which provides an effective and increasingly important buffer against the atmospheric fluctuation of the gas.

■ 2.2.2 Waves :

Wave forms when the water surface is disturbed, for example, by wind, earth quake or planetary gravitational forces. During such disturbances energy and momentum are transferred to the water mass and transmitted in the direction of the impelling force. This energy powers many coastal processes, hence a knowledge of wave hydrodynamics is vital in understanding many coastal phenomena, from sediment transport to ecological adaptations and even in siting of ports and harbour. (Carter, 1988)

Waves can be treated as an energy-source, the transformation of solar energy into the mechanical energy of the wind is subsequently converted into wave potential energy by the deformation of the still-water sea surface. The resultant 'piles of water' possess potential energy in exactly the same way that river water does at the head of its course.

Each wave possesses a characteristic length (L), defined as the distance between successive wave crests, and wave period (T), defined as the time between successive crests to pass through a fixed point. The velocity at which the wave travels is given by the ratio of these two.

Another extremely important attribute of waves is their height (H) defined as the distance between trough and crest. Unfortunately wave height is not related to the other wave characteristics and must consequently be found by direct measurement.

The inter-relationships between these four wave properties are the subject matter of several wave theories. There are three main theories each of which contributes to the overall understanding of wave dynamics – none of which provides a complete explanation. These theories are :

1. Airy wave theory. Named after Airy (1845). Applies to waves of very small height in deep water.

Stock's wave theory. Named after Stokes (1847). Applies to waves water depths.

3. Solitary wave theory, named after the distinctive detached nature of waves in shallow water.
(Pethik, 2000)

Sl No	Wave Parameters	Deep Water $d/L > 0.05$	Transitional Water $0.05 > d/L > 0.05$	Sallow water $d/L < 0.05$
1	Wave Length	$gT/2\pi$	$gT/2\pi \tanh 2\pi d/L$	$(gd)^{0.5}$
2	Wave Phase Velocity	$gT^2/2\pi$	$gT^2/2\pi \tanh 2\pi d/L$	$T(gd)^{0.5}$

■ 2.2.2.1 Wave energy :

The energy of a wave a) potential, due to the deformation of the wave above still-water level, and b) kinetic, due to the orbital movement of the water particles within the wave form. Airy wave theory predicts that these two forms of energy are equal and the total of these is directly related to the square of the wave height :

$$E = \frac{1}{8} \rho g H^2$$

Where : E = Wave energy
 ρ = water density
 g = acceleration due to gravity
 H = Wave height

This expression is for energy per unit wave crest. The rate at which this energy is transported across the surface of the sea is related to the velocity of the wave form – C. However there is a complication here, for the energy in a group of waves does not move at the same rate as a single wave within the group. In fact individual waves move twice as quickly as the group in deep water so that a single wave can be seen progressing through the group and disappearing at the front to be replaced by a new wave at the rear.

This group velocity is included in the expression for the rate of energy transfer (P) :

$$P = EC_n$$

Where P = Wave energy flux or wave power

n = group velocity

(n = $\frac{1}{2}$ in deep water). (Pethick, 2000)

Wave refraction in shallow-water : like light wave, the marine waves change their direction near the coast due to differential depth and resultant difference in velocity. It is an important wave behavior in shallow water. Wave refraction is the process by which the wave crests are bent until become parallel to the submarine contours .

■ 2.2.2.2 Wave Refraction

Waverefraction is the gradual reorientation of the waves propagating at an angle to the bottom slope or against a current (Carter, 1995). A series of wave (wave train) which is approaching a coastline with its crests at an oblique angle to both shore and bed contours will, at any given instance, have the shoreward portion of its crests in shallow water and the seaward portion in deeper water. Since the wave phase velocity is directly related to water depth this means that the shoreward part of a wave crest will be moving more slowly than the same crest further out to sea. The result is that the seaward portion swings forwards and the wave crest becomes curved. This process continues until all parts of a crest are in the same water depth, at which stage equal wave phase velocity is restored and the wave crest is parallel to bed contours and shoreline.

This process applies to waves approaching head and bay complex where the submarine contours follow the shoreline configuration. In such a case wave refraction takes place at different rates according to the position of the crest and the waves are refracted until they parallel the shore once more. This process results in a spreading out, or divergence, of the wave rays (lines drawn at right-angles to wave crests) in the bays and their convergence at the headlands. This convergence/divergence also results in a concentration of energy at the headlands and a dispersion of energy in the bays so that wave height increases at the headland and decrease in the bay.

The amount of wave refraction which a given wave will undergo depends on the proportion of its wave crests which are moving in shallow water and therefore are affected by the shoaling transformation of C (Pethick, 2000).

$$C = (gd)^{0.5}$$

According to Airy theory the velocity is proportional to the depth of water and so the portion of the wave lying at shallow water experience slower velocity than the other part and the portion with speedy movement at deeper portion bends to be parallel to the bathymetric contours.

■ 2.2.2.3 Wave Breaking

Wave when approaching towards the coast its velocity decreases and, on the contrary, height increases in order to maintain the energy balance by attaining more potential energy to substantiate the decrease of kinetic energy. Another important change in the wave character is the reduction in the wave length at shallow water. The increase in wave height and decrease in wave length is termed as wave attenuation and is responsible for the gradual increase in wave steepness. The steep wave after attaining threshold condition, breaks along a breaker line or breaker zone to dissipate its energy brought from deeper sea.

▲ 2.2.2.3.1 Types of Breaker

Although breakers exhibit a continuous series of forms, Galvin classified these into four basic types :

1. Surging
2. Collapsing
3. Plunging
4. Spilling

▲ 2.2.2.3.1.1 Surging;

Bases of the wave rushes up the slope of the beach with only minor production of foam or bubble. It is developed on steep beach slope. On the steep slope of laterite chunks at Old Digha, such breakers are observed during high tide.

▲ 2.2.2.3.1.2 Collapsing

Sometimes on steeper slope the wave front collapses with huge energy.

▲ 2.2.2.3.1.3.Plunging

Here the front face of the wave steepens until being vertical and the crest curls and falls into the base of the wave often rebounding. At Old Digha, the breakers are plunging in type.

▲ 2.2.2.3.1.4.Spilling

In this type the front face of the wave gradually breaks forming the foam bubbles and turbulent water. Wave breaks within a wide zone and thus the dissipated energy is distributed. On gentle gradient beach such breaking is experienced. At New Digha, on gentle slope, the breakers are spilling in type.

▲ 2.2.2.3.2 Quantitative Analysis of breaker

The energy brought to the coast by the wave and the percentage of the energy dissipated by breaking depend on the type of breaking. The type of breaking again depends on the wave length, beach slope and wave height during breaking.

Galvin (1968, 1972) proposed a breaker coefficient incorporated both these controlling factors. Using deep-water wave characteristics his breaker coefficient B_o is

$$B_o = \frac{H_o}{L_o s^2}$$

or, using the wave height actually attained at the break-point (much easier to measure in the field):

$$B_b = \frac{H_b}{g s T^2}$$

Where : B_b = breaker coefficient (Shallow Water)

H_b = wave height during Braking

L = wave length

g = acceleration due to gravity

s = beach slope

T = wave period

Galvin showed that the breaker coefficient not only defined the position of the break-point but also the form of the breaking wave.

Their surf scaling factor was proposed by Guza and Bowen in 1975 and is very similar to that of Galvin :

$$e = \frac{a.2\pi}{gT \tan^2 \beta}$$

Where a = wave amplitude (height)

g = acceleration due to gravity

T = wave period

β = beach slope.

Galvin showed that the breaker coefficient not only defined the position of the break-point but also the form of the breaking wave.

Table: Comparison of breaking coefficients

Author	Theory	Expression	Breaker type transition	
Galvin 1968	Breaker Coefficient	$B_b = \frac{H_b}{g_s T^2}$	Surging to plunging	Surging to plunging
			0.003	0.068
Guza and Bowen 1975	Surf Scaling Factor	$e = \frac{a.2\pi}{gT \tan^2 \beta}$	2.5	33
Kemp 1975	Phase Difference	$P = \frac{t}{T}$	0.5	1.0

Breaking coefficient is thus an important indicator of availability of excess energy which bounces back into the sea taking some materials off into the sea. The coastal dunes and other protective structures generally increases the beach gradient and thus the breaker type is changed. If a beach

presently face surging breaker in place of earlier spilling, it turns to be prone to more erosion instead of protection by the protective structures.

■ 2.2.3 Current

Currents are the horizontal movement of water towards or away from the coast. It is of most important process to study because the orientation, angle of approach and intensity of this current determines the efficiency of the coastal processes to displace and transport the sediment and thus regulates the change in coastal configuration. Ocean currents are drifts of water responding to prevailing wind patterns and to variations in water temperature and salinity. True ocean currents have little significance for coastal process and landforms, except where they influence the distribution of coastal ice, coral reefs, or mangrove swamps.

Coastal Currents are of two types a) Shore normal currents and b) Long-shore currents :

■ 2.2.3.1 Shore normal currents

Most of the currents which develop at right angles to the shoreline are produced by the onshore-offshore movements of water particles in association to the swash and backwash. The potential energy of the waves is constantly being transformed into kinetic energy and back into potential as the wave moves forwards. In deep water this kinetic energy takes the form of circular orbits of water particles, but the transformation in shallow water causes these to be modified and results in a set of shore-normal currents which vary in magnitude and duration as the water depth decreases. Wave breaking at the breaker line as a result of attenuation in shallow water is responsible to transfer the circular orbital movement to be changed in to a horizontal one and is then termed as transcurrent wave where there is no vertical component of particle movement.

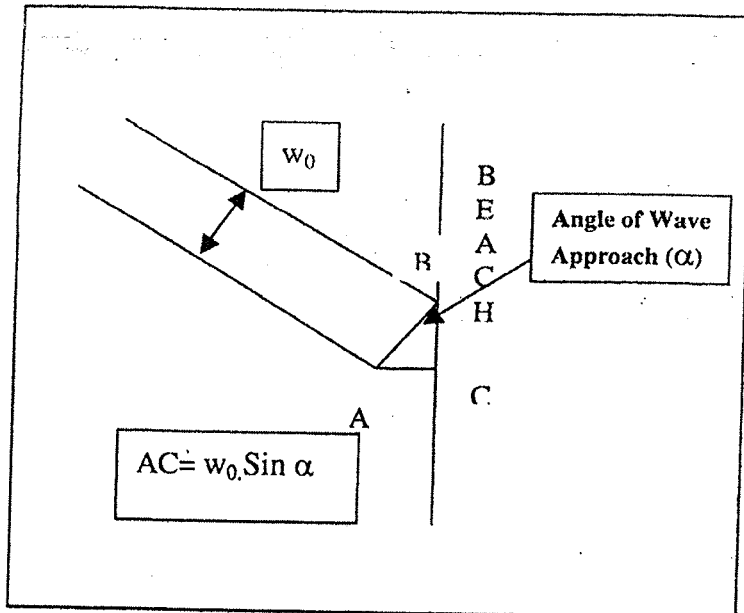
■ 2.2.3.2 Long-shore currents :

One approach to the partition of oblique wave approach into shore-normal components has been to consider the mass transport current and long-shore components and apply to it the principle of continuity.

The unit width of wave crest approaches the beach at an angle α and the water movement associated with it is envisaged as dividing into two components : one shore-normal, the other long-shore. The long-shore component can be determined by the simple geometric

procedure of calculating the width of wave crest that will travel in the direction. As figure shows this is simply :

$L(\text{Width of long-shore component: BC}) = w_0 \sin \alpha$ (Pethick , 2000).



Inman and Bagnold (1963) determine the width of the shoreline affected by this portion of the long-shore component. Figure shows that this will be =

$$L = \frac{w_0}{\cos \alpha} = \frac{1}{\cos \alpha}$$

■ 2.2.3.3 Tidal Currents

Tidal currents, produced by the ebb and flow of tides, repeatedly and regularly reverse in direction. In the open ocean their velocities rarely exceed 1 m/s, but where the flow is channeled through narrow straits, or entrances to estuaries and lagoons, they may exceed 4 m/s locally. The velocity of currents in these confined channels is related not only to the capacity of the channel but also to the tidal range, being greatest in areas of large range (sometimes called macrotidal areas, as opposed to mesotidal and microtidal areas with medium and low tidal ranges).

A long coasts of certain areas like East Anglia, local tidal currents may reverse along the shore during the few hours preceding and after high tide, but such local effects are usually minor

compared with the literal drift induced by wind-driven waves. Wind-driven masses of surface water can increase or decrease the tidal range, locally and temporarily, as they are piled up by onshore winds or driven away from land by offshore winds.

In the near shore zone water carried onshore by waves is often carried away again as *rip currents*. Rip currents may be regularly or irregularly spaced. On some beaches rips develop in zones where breaker height is greatest and re fed by water from the highest breakers, but their origin and mechanism is still not well understood.

The power of currents to erode the coast is very limited and is usually confined to the transport of sediment which has already been placed in suspension by wave action or by a river. The sea floor, however, can be eroded and stilts and sands set in motion. Many hollows, shallow troughs, and large dunes in shallow seas have been attributed to current action, but few convincing data are available.

■ 2.2.4 Cell circulation :

sometimes on the beach, a cellular circulation of water is noticed comprising of shore normal component and shore parallel (long shore) component. The water brought into the beach by swash part of shore normal current and then flows towards each other to be joined together to flow back to sea in the form of backwash. This cellular movement of water is called as cell circulation. The separation of shore-normal and long-shore currents is an obvious enough device for the simplification of a complex problem. It is clear that the two currents will be superimposed on one another to have an important contribution to sediment transport. Yet the near-shore zone on many beaches is characterized by circulatory current system in which both shore-normal and long-shore currents are both present but occupy quite distinct pathways.

This circulatory system was first recognized by Shepard, Emery and LaFond (1941) whose work was followed by that of Shepard and Inman (1950). In this cell circulation the slow, onshore, mass transport is transformed into long-shore currents in the zone shore wards of the breakers. These long-shore currents feed narrow zones of offshore water movement – the rip currents. These rip currents balance the amount of water brought into the near-shore by the mass transport; they move this water to seawards across the breakers and then broaden into a rip head

before disappearing(Pethick,2000). The development of rip cell is a debated question. Some groups of Geomorphologists believe that this shore parallel component of the circulation is developed on the irregular coast where water moves from higher to lower elevation causing shore parallel movement. But the others believe that on regular coast the water gradient is generated by the interaction of edge wave and incident wave and thus favouring the development of shore parallel movement of water and then cell circulation becomes possible.

■ 2.2.5 The tides :

Tidal waves are not very high in the open ocean – perhaps no more than 50 cm. They do however increase in height as they reach the shore just as wind waves do and may reach heights of 5 m on some coasts. Velocity of 80 Km per hour, means that enormous volumes of water shifts position at the coast each day and enormous amounts of energy are changing the structure of the coastline and the coastline is adjusting constantly to these tidal waves. Land forms such as mudflats, marshes, estuaries and extensive beach profiles are all due to tidal energy and thus the tide can be considered as one of the important coastal process.

■ 2.2.5.1 The energy source :

The gravitational pull of the moon and to a lesser extent, the sun, provide the energy input which drives the tidal waves. Despite the fact that the sun is much bigger than the moon, it is much further away so that the sun's tidal attractive force is just less than half that of the moon, in fact 0.46 times as great.

These forces are felt as a vertical and horizontal lift by all the particles which make up the earth – but it is the water of the ocean which responds most obviously to the attraction. As the water passes under the moon so it is drawn upwards and along by this force – forming a bulge. This bulge under the moon is matched however by a similar bulge on the side of the earth facing away from the moon. We must consider why this is so.

If the moon and earth were stationary their gravitational attraction would make them fly together. But they are not stationary – they rotate about a common axis. This produces a so-called centrifugal force which attempts to pull each planet outwards. The gravitational and centrifugal

forces are exactly equal and opposite so that planets remain at a constant distance apart. Each particle on earth experiences both the centrifugal and the moon's attractive force.

This balance of forces is true only as an average for each planet. In fact each particle on earth experiences the same centrifugal force – but the moon's attraction is greater for those particles closest to the moon.

This means that a parcel of water directly under the moon experiences a slightly greater attraction towards the moon than the centrifugal force exerts in the opposite direction. Its net movement is consequently towards the moon-hence the 'bulge'.

On the opposite side of the earth (facing away from the moon), the moon's force is slightly smaller than the centrifugal force so a water parcel there is attracted away from the moon by the larger centrifugal force. Thus two bulges are produced – one under the moon, one on the other side of the earth (Pethick,2000).

■ 2.2.5.2 Types of Tide :

▲ 2.2.5.2.1 Progressive wave tides :

When a wave reaches shallow water it suffers several changes :

1. Its speed is checked.
2. Its wavelength decreases.
3. Consequently its energy is packed into a smaller area and its height increases.

Tidal waves at the coast change in exactly the same way : that is, they slow down and increase in height. The open-coast tides may in fact be as much as 2m in height, compared to the 50 cm maximum of the ocean tide. The increase in tidal wave height is responsible for increase in the tidal range and this has profound effects on the landforms of the coast as well as on plant, animal and even human life. Another effect is that current speeds which increases as more water has to change places to accommodate the increase in range.

▲ 2.2.5.2 Rotating tides :

The rotating crest of the tidal wave around an amphidromic point obeys the same rule as a progressive wave in that it slows down and increases in height in shallow water.

The rotating tide of the mid-Atlantic shows the shallow water transforming points which have high tides at the same time. The lines can be imagined as the crest of the tidal wave 'photographed' each hour on the same film.

The range of the tide increases as the wave slows down as predicted from elementary wave theory. Thus the tidal range in the open ocean is less than 50 cm but the range increases to 1.25 m on the Venezuelan Coast as the wave speed begins to decrease.

The range of the rotating tide however is dependent on another factor – the distance from the amphidromic point. Tidal range at the amphidromic point itself is always zero and the height of the wave increases as the distance increases away from the centre. Looking again at the tidal map of the mid-Atlantic it can be seen that the tidal range in the Caribbean Islands is almost zero – since they lie around the amphidromic point. The Venezuelan coast lies further away and experiences a 1.00 m range, while the same rotating tide on the West African coast has a range of 2.5 m.

The importance of this variation in tidal range is obvious enough. As well as enhancing holidays on the almost tideless Caribbean beaches, tidal range also controls the extent of landforms such as beaches, marshes and, in this area, mangrove swamps, and has a profound effect on tidal currents and sediment movements.

■ 2.2.6 Marine Ecosystems :

Covering more than 70 percent of the Earth's surface, with an average depth of 3,750m and a salinity averaging 3.5 percent (27 percent of which is sodium chloride), marine ecosystems are of singular ecological significance. Unlike land and freshwater ecosystems, the sea is continuous and is in continual circulation via the major surface currents, upwelling, seasonal turnover in the upper waters, waves, and tidal action. In waves, water oscillates back and forth without actually changing position. The concentration of nutrients in the ocean is low. Nitrates, phosphates, and other nutrients are measured in parts per billion in contrast to salts, such as

sodium chloride, which are measured in parts per thousand. After –hydrogen and oxygen which constitute about 96 percent of the Ocean's elements, the major elements (those constituting 0.3 to 2.9 percent) in descending order are chlorine, sodium, magnesium, sulfur, calcium, and potassium (Klemm et al. 1990).

The major zones in marine ecosystems are the littoral, neritic, pelagic, and benthic after Koromondy.

6.1.Littoral

The *littoral* or *intertidal zone* is the shoreline between land and open sea. It is subject to the physical violence of waves and tides and to fluctuations, sometimes extreme, of temperature, moisture, and light intensity. Along rocky shores, we find more sessile organisms (algae, barnacles, starfish, etc.) than in any other type of ecosystem (Sebens 1985). Along sandy shores, organisms are adapted by burrowing in or adhering to sand (eg. Ghost crabs, sand dollars, polychaetes). Mud flats, which occur in bays, harbor algae on the surface and often photosynthetic bacteria beneath the algae, along with an abundance of clams, worms, and crustacean. Coral reefs, fringes of coral forms around islands by colonial community, also may be regarded as littoral ecosystems. Because of their generally higher temperatures, depth of light penetration, and richer nutrient levels, coral reefs are among the most highly productive of all ecosystems (Goreau et al. 1979); Dubinsky 1990). Salt marshes are also intriguing ecosystems (Bertness 1992), as are wetlands because they are havens of biodiversity (Kulser et al. 1994).

6.2Neritic zone

The *neritic zone* is by a continental shelf, extending outward to its edge to a water depth of light penetration and the presence of nutrients washed from the land. This richness and productivity decrease gradually toward the deeper waters of the zone. Extensive algal communities of giant kelps, as well as of smaller uni and multi cellular forms, along with clams, snails, worms, and echinoderms dominate the bottom. Phytoplankton and zooplankton are relatively abundant and support some of the greatest fishing grounds in the world. Continental shelves are discussed extensively by Postma and Zijlstra (1988).

6.3 Pelagic zone

The *pelagic zone* is the open sea, constituting about 90 percent of the total ocean surface. Photosynthesis occurs in the surface-pelagic zone primarily by planktonic diatoms and dinoflagellates. Copepods and arrow worms are the major types of zooplankton, along with active swimming forms, such as shrimp, jellyfish, and ctenophores. Although the largest in size, this zone is low in nutrients and hence in both primary and secondary productivity. Nonetheless, summer blooms, which are absent in tropical waters but do occur in the Antarctic, are sufficient to support fin and blue whales; the latter are the largest animals alive, reaching lengths of 33m and a mass of 13.6×10^4 kg.

As Smith (1993) has noted, since most plankton live in the sunlit waters between the surface and 75 to 200 m in depth, most of the production occurs in about 5 percent of the ocean, by total volume of water. Phytoplankton account for over 95 percent of marine photosynthesis, and marine primary production may account for more than 30 percent of the global total.

Also, in spite of its relatively low nutrient content, the top millimeter of the ocean plays a significant role in the global regulation of carbon dioxide (MacIntyre 1974). The organisms in the pelagic zone below the level of light penetration are completely dependent on the rain of detritus from the upper regions and are wholly heterotrophic. Carnivorous copepods and other crustacean are abundant in the upper waters of this non photosynthetic zone. In deeper waters many animals have reduced vision, others, including many fish, are bioluminescent, and some deep-water fish have light producing organs.

6.4 Benthic zone

The *benthic zone* extends from the edge of the continental shelf to the deepest ocean trenches. Organisms here are heterotrophic, many being anchored in the muddy ooze made up largely of siliceous and calcareous shells of foraminifera, snails, radiolarians, and diatoms. Among these 'rooted' animals are sea lilies, sea fans, sponges, and brachiopods. Snails and clams embed in the mud while starfish, sea cucumbers, and sea urchins move on its surface.

Encompassing about 94 percent of the ocean bottom, the benthic zone is remarkably homogenous and stable in its physical and chemical parameters (Sanders and Hesslet 1969). However, volcanic areas, like that south of the island of Hawaii in the Hawaiian archipelago, that

are subject to underwater volcanic activity and the hot springs that are often associated with such areas, nourish intriguing forms of life, as well as laying down significant deposits of various ores (Edmond and Von Damm 1983).

■ 2.2.7 Deep-sea Deposits :

Sediments cover the ocean floor, except where the slopes of submarine mountains are too steep for the debris to withstand the pull of gravity. The average thickness of sediment is about 2,000 ft in the Atlantic Ocean and 1,000 ft in the Pacific Ocean. Many small submarine hills have been completely covered by sediments.

Approximately 3 billion tons of sediment are dumped into the oceans each year by the rivers of the world. Much of this material remains trapped on the continental shelf, but some is carried into the abyss. Deep-ocean sediment consists of meteoric dust, volcanic ash, some organic matter, and fine material carried seaward by winds.

Deposits on the abyssal plain are called oozes. At the ocean depths of 2,000 ft to 13,000 ft to 27,000 ft, these deposits contain mostly clay, with some quartz, mica, and iron stain. This material is so fine that individual grains can be identified only by use of X-ray. Deep-ocean clay contains little organic matter, because calcitic shells are dissolved in deep waters. Carbon dioxide, which is in solution in relatively large amounts, plus the great pressure of the water help to dissolve calcite. Shark teeth and fish ear bones are two of the few recognizable animal remains.

Shells of calcite and coverings of silica overlay about half the ocean floor. This material is called calcareous ooze or siliceous ooze, depending on which material is more abundant. Siliceous ooze, most of which comes from diatom coverings, is less abundant than calcareous ooze. However, some deposits of diatomaceous earth have been found on what is now land area. This proves that these deposits have been uplifted above sea level. Diatomaceous earth is a fine material which is used in filters and as an abrasive in toothpaste.

Meteoric dust, or "star dust", falls into the ocean and accumulates as part of ocean sediment. On land, such materials mingle with the soil and cannot be recognized. The meteoric origin of the dust in the ocean is quite apparent. Measuring this gives a clue to the total amount received by the earth. Other interesting materials found in the deep sea are manganese and iron nodules or lumps. Both manganese and iron are precipitated around cores of animal remains.

Manganese is necessary to the manufacture of steel. Since little manganese is found on land, perhaps eventually the steel industry will turn to nodules from the sea floor for its manganese.

■ 2.2.8 Self Assessment Questions

1. Describe the components of a wave. How do a wave experience transformation when it approaches into shallow water?
2. How do you estimate wave energy in shallow water?
3. Define wave refraction. Describe the mechanism of wave refraction.
4. Distinguish between Shore Normal and Long Shore Current. How do you measure the intensity of Long Shore Current?
5. Describe the mechanism for the formation of the Rip Cell? How do the Rip currents maintain balance between onshore and offshore region?
6. Where does the tide get energy from? Discuss the role of moon in the formation of tidal wave.
7. Classify tide. Describe the progress of Rotating tide with reference to the co tidal line.
8. Describe the composition of Marine Eco system with special reference to the discussion on littoral zone.
9. Describe the composition and characteristics of deep sea deposits.

■ 2.2.9 References.

- Carter RWG (1988), Coastal Environment, Academic Press Limited, 1-191.
- French W.P., (1997), Coastal and Estuarine Management pp.25-52 Routledge, London.
- Pethick, J. (1984), Introduction to Coastal Geomorphology, pp. 9-19, 21-38, 47-59.
- Nordstorm KF and Roman, CT, (1996), Estuarine Shore, pp. 1-9, Wiley.
- Sharma, R.C.and Vatal M,(1999) Oceanography for Geographers 287P

PART-I
Paper - I, Module - II
OCEANOGRAPHY AND HYDROLOGY
Unit – 03

MODULE STRUCTURE

- 2.3.0 : Introduction
- 2.3.1 : Objective
- 2.3.2 : Onshore and Offshore Oceanic Regions
- 2.3.3 : Geomorphology of Coastal Regions
- 2.3.4 : Coastal Ecology— Coastal Dunes, Mangroves and Coral Reefs
- 2.3.5 : Primary Productivity; Trophic Interrelationship
- 2.3.6 : Diversity Gradients and Biogeography
- 2.3.7 : Self Assessment Questions
- 2.3.8 : References

Paper - I, Module - II
OCEANOGRAPHY AND HYDROLOGY
Unit – 03

□ **2.3.0 INTRODUCTION :**

The present unit Module-II includes oceanic regions, coastal regions, Geomorphological features of the coastal regions and approaches of study on geomorphology of coastal regions, and finally the habitat diversity of coastal ecology. Sand dunes, Mangroves and coral Reefs act as natural buffers against marine hazards and harbour specific community of plants and animals under different physiographic settings of the coastal regions.

□ **2.3.1 OBJECTIVES :**

- i) To know the subdivision of oceanic regions using both morphological and wave process terminology and current activities.
- ii) To know the study approach on geomorphology of coastal regions.
- iii) To know the significance of coastal ecology and functions of major coastal habitats.

□ **2.3.2 ONSHORE AND OFFSHORE OCEANIC REGIONS :**

Δ **Marginal sea coasts :** This is a relatively uncommon coastal type and occurs where plate convergence takes place offshore, with a relatively wide continental shelf separating the plate boundary from the coast. In many ways, this coastal type is similar to trailing edge coasts, especially in the sedimentary features that may develop, but it suffers earthquakes more regularly and many experience limited tectonic movement. The Gulf of Mexico coasts are an example, with the volcanic islands of the Caribbean marking the zone of plate convergence.

Δ **Coastal terminology :** As with all scientific disciplines, the study of coasts has generated a vast lexicon of terms to describe landforms, processes, deposits, habitats and ecosystems. Hopefully readers will agree that one of the most valuable parts of this book is the fairly extensive glossary. I have included all the important coastal terms that one will come across in the literature, in the hope that students will not be discouraged by the sometimes bewildering array of terms that are used, not only to describe different attributes of coastal systems, but sometimes apparently the same attributes. Rarely, however, are these terms precisely synonymous, and important, if subtle, differences in meaning do exist. A command of coastal terminology serves to better understand coastal systems, and also to enhance the communication of the subject. I therefore encourage you not to pass over

unfamiliar terms, but to try and grasp their meaning both from the text in which it occurs, and with reference to the glossary. For when one goes beyond this introductory text to delve into the primary literature in scientific journals, which I thoroughly recommend, one quickly discovers that no allowance is given to readers with a limited vocabulary on the subject, and the value of spending time reading such articles is correspondingly diminished.

A useful example of similar yet subtly different terms involves the standard subdivision of the coastal zone. Two sets of terms are shown. However, one subdivides the coastal zone based on morphological changes (backshore, foreshore, inshore and offshore), whilst the other is based on the types of waves processes that operate in different parts of the coastal zone (swash zone, surf zone and breaker zone, which together comprise the nearshore zone). The same space may also be subdivided on sedimentological grounds, either according to the type of sediment, sedimentary structures and/or depositional processes that occur (Reading and Collinson 1996).

- the beach zone occurs between mean low water and the landward limit of wave activity and mainly possesses parallel bedded sand layers.
- the shoreface zone occurs between mean low water and the mean fairweather wave base and is characterized by wave swept sand ripples.
- The offshore-transition zone lies between the mean fair-weather wave base and the mean storm wave base and is dominated by sediment deposition during storms.
- The offshore zone lies below the mean storm wave base and is characterized by fine-grained sediment settling out of the water.

Δ **The Neritic Province** : The world's seas can be broadly divided into benthic (sea-bed) and pelagic (water-column) realms both of which can be further subdivided according to depth. The pelagic realm consists of oceanic and neritic provinces, the latter comprising the shallow (down to 100-200 m) coastal water overlying the continental shelf. Although accounting for only a comparatively small proportion (c. 5%) of the world's seas, neritic environments are often highly productive and many of the important commercial fisheries are located in these inshore waters. The euphotic (= epipelagic) zone is that part of the pelagic realm which is well illuminated. Its lower boundary varies according to the clarity of the water but is usually around 100-200 m. It grades into the aphotic zone which is permanently dark below about 1000 m.

□ 2.3.3 GEOMORPHOLOGY OF COASTAL REGIONS :

Δ **Introduction** : The study of landforms provides us with an exciting challenge. The proper study of mankind may be man, but the chance to study a subject which is both bigger and older than ourselves sets an intellectual hurdle which requires some flexibility to overcome. If we happened to

live for hundreds rather than tens of years then perhaps we would be able to notice landforms developing — rather like a speeded-up film of cloud formations or plant growth. Unfortunately we do not possess such geomorphological insight and consequently we must perform acrobatics with our scientific imagination in order to comprehend the enormous scales involved.

Yet there is one landscape which almost everyone recognizes as undergoing continuous change. The coastline changes; not only over centuries or decades, but in a matter of hours or minutes. This rapid development applies both to the form of the coastline — beach profiles may change quite significantly during a single day — and to the coastal processes — tidal variations being an obvious example. Thus the coastal geomorphologist has the great fortune to be involved in the study of a dynamic landform whose development can be observed directly.

It does not set out to treat such changes in an historical sense but to analyse the mechanisms which result in change. It also attempts to define the end product of such change — the equilibrium coastal landforms. In order to do this some considerable time must be spent in discussing the forces which drive the coastal processes — waves, tides and currents. This discussion will allow us to examine the processes of coastal landform development — the transport of sediments. Lastly we may apply this knowledge to a consideration of coastal landforms, their development and equilibrium states.

Such a study may provide an exciting intellectual challenge, but it may be questioned whether it should not be of some use at the same time. Yet coastal geomorphology is directly applicable to our lives; the world's coastline — some 440,000 km of it — encompasses only a small area of the total land surface — about 0.03 per cent if the coastal zone is regarded as about 100 m wide — but its importance to man is fundamental. A United Nations estimate suggests that 66 per cent of the world's population lives within a few kilometres of the coast, consequently food production, communications, settlement, even recreation, are concentrated here. However the coastline presents enormous problems for such intensive use; flooding, erosion, pollution and the continued threats posed by rising sea levels — all demand constant action in order to preserve man's investment. Such action depends on knowledge of the mechanism by which the natural coastal environment functions.

Δ Coastal classifications : Until quite recently, the almost universal approach of the coastal geomorphologist to his subject was to attempt a classification of coastal landforms. It is true that we should be able to recognize and name coastal features that classification does allow minor discrepancies between components to be put aside in favour of their more significant similarities. Yet classification does tend to describe rather than explain and the task of the coastal geomorphologist must be to understand the relationships between form and process, not merely to describe forms. Thus classifications have tended to retard the development of a truly scientific coastal geomorphology.

Consequently no overall classification will be used in this book; instead, as we have pointed out, our discussion will be concerned with the forces, processes and landforms of the coast. Yet there

are classifications which are based upon the controls of the coastal environment and which do provide some insight into the functioning of the coast as a whole. These genetic classifications are dominated by the realization that coastal forms are largely the product of sea-level variations. Several classifications use a division into submerged and emerged coasts (e.g. Johnson 1919; Valentin 1952). Such classifications suggest that submergence results in fjord or ria coastlines while emergence produces tidal flats or even barrier islands. There is no doubt that the sea-level changes do play a dominant role in coastal development but the complex variations in sea-level even over the past few thousand years makes identification of specifically emergent or submerged coasts extremely difficult. In fact, Johnson's (1919) classification places most of his observed coastal forms into a class labeled 'neutral' in order to avoid any such controversy.

Another basic group of classifications used structural controls to distinguish between various types of coastal landforms. Thus Bloom (1978) distinguished between bold and low coasts. Bold coasts are developed in resistant rock and the resultant forms reflect both the inherited sub-aerial topography and the effect of marine erosion in picking out lines of weakness in the rocks. Low coasts are developed on alluvial coastal plains and, rather confusingly, are also associated with recent sea-level variations.

More recently Inman and Nordstrom (1971) have developed a tectonic classification based on plate tectonic theory. They recognize four distinct coastal types : those on actively diverging plates (e.g. Red Sea coasts) those on zones of plate convergence (e.g. the island arc systems of Indonesia and Japan) those on major transform faults (e.g. the coast of southern California) and lastly those coasts developed on stable plate zone (e.g. coasts of India and Australia). This is a useful distinction and is reminiscent of the suggestion by Suess (1904) that coasts could be divided into Atlantic and Pacific types.

A third group of genetic classifications are based on coastal processes. Included here is Shepard's (1963) division into primary and secondary coasts. Primary coasts are those in essentially the same condition as they were left in at the end of the last sea-level change, that is relatively unaffected by marine processes. The form of these coasts reflects the sub-aerial processes which shaped the land surface before sea-level changes occurred. Secondary coasts have been altered considerably by marine processes such as erosion, accretion and organic deposition.

Also in this group are classifications by Davies (1980) and Tanner (1960). Both these authors classify coasts according to the level of energy inputs that they receive. Davies (1980) distinguishes between storm wave environments – primarily those in middle latitudes, and swell wave environments – found mainly in low latitudes. Tanner's (1960) classification into high, moderate and low-energy coasts is less geographical than that of Davies but is based on a similar premise.

The form and function of coastal landforms : Looked at in this light the function of the steady-

state coastal landform is to dissipate wave-energy. Changes in form are necessary only when a change of energy input occurs, in which case the energy will not be dissipated without net sediment transport occurring. Consequently the morphology begins to alter – until a new equilibrium or steady state is established. The time taken for such a re-establishment of steady state is referred to as the relaxation time (Chorley 1962).

This delicate balance between form and function is maintained by the processes of sediment transport and driven by forces of wave, tides and wind – a triple alliance.

Δ Time and space in coastal geomorphology : One of the more confusing aspects of coastal geomorphology is the extremely wide range of scale covered by the landforms. These include small-scale features such as beach cusps – no more than a few metres across, medium-scale features such as salt-marshes or sand dunes – several kilometres across, and large-scale features – the configuration of the coastline itself for instance – covering tens, even hundreds, of kilometres : a scale which may be called ‘capex and bays geomorphology’.

However these variations in the spatial scale are paralleled by changes in the temporal scale, that is, the relaxation time of each landform. Our present day shoreline was formed very recently – in geological terms – during the post-glacial sea-level rise which ended only 6000 years ago. The establishment of a shoreline does not mean that coastal landforms were also produced; of course the shoreline does delimit a series of large-scale capes and bays – but these are not equilibrium coastal landforms, merely the land-scape inherited from sub-aerial processes. Once such a shoreline is established however the process of landform development towards a steady state begins. Small-scale features, beach cusps or beach profiles, may reach a steady state with their environmental controls in a matter of hours or days. Salt-marshes take much longer – perhaps between 200 years to 1000 years to complete their relaxation time. Finally the coastal configuration, head-lands, bays, estuaries, and so on will take thousands of years to adjust to the environmental changes produced by the new sea-level.

Returning to the coastal landforms, it could be argued that ‘steady-time’ characterizes the rapid adjustments of the beach profile, while graded-time is applicable to the arcuate shapes developed by the beach plan. The capes and bays of the coastline however are the coastal equivalent of drainage basins and require thousands of years to reach a steady state – the cyclic time of Schumm and Lichty (1965). Similar hierarchies could be assigned to the estuarine environment where the mudflat surface, salt-marsh and estuary itself conform to the steady, graded and cyclic time scales.

Coastal geomorphologists have been slow in applying such terrestrial geomorphic models to their environment.

The importance of the nesting together of the various scales of landform lies in the insight that

this allows into the independent controls of landform development. A beach profile for example may respond to a given wave input during a single day but if the beach profile is monitored over a longer period – 10 years for example – it will be seen to respond to long-term variations which are taking place in the beach plan shape. These longshore changes will alter the angle of wave approach to the beach and thus wave inputs which were a totally independent control of beach profile in the short term become a dependent variable in the longer term.

The coastal geomorphologist must be aware of such scale interactions between the components of his environment before he begins any investigation. One of the major problems he faces in attempting to assess such interactions is the recognition of fossil elements in the landscape. At the largest time-space scale, cyclic time, the development of shoreline configuration may take thousands of years to attain a steady state. Since our present sea-level was established only 6000 years ago this implies that many large-scale features have not yet adjusted to these 'new' conditions. Consequently many of our present-day coastal landforms are responses, not to modern wave or tidal conditions, but to some previous environmental processes. Such fossil forms may have been inherited from a previous high sea-level, in which case they will be a response to marine processes; some may be a response to sub-aerial processes acting during a previous low sea-level.

|| 2.3.4 COASTAL ECOLOGY – COASTAL DUNES, MANGROVES AND CORAL REEFS :

Coastal areas can be regarded as the interface between three habitable media, namely earth, air and sea. However, this interface is composed of a series of boundaries or gradients whose dimensions range from a few nanometers or less, for example a water film around a sand grain, to a few kilometres or more, for example the windborne carriage of sea-spray inland. These boundaries are not constant but change in space and time, sometimes slowly, sometimes rapidly. Some changes such as tidal rise and fall are predictable and regular, others stochastic, that is more or less unpredictable or at random.

All coastal areas contain at least two obvious habitats, namely the maritime zone, which is home for many terrestrial animals and plants, and the sea itself. Although there are several areas in the world where water depths greater than 200 m are found close inshore, this book will restrict its marine sections to consideration of the ecology of the shallow-water zone overlying the continental shelf. A third obvious habitat, the intertidal zone, sometimes called the littoral, occurs along much of the world's coastlines, but even in areas without noticeable tides there is often a specialized fauna and flora at the sea – land interface.

Δ Sand dunes

The strandline : Coastal sand dunes, like saltmarshes, result from the stabilization of transported sediment by vegetation. They are formed from wind-blown sand rather than from water-moved silt. Although the development of dunes is wind-dependent, the earliest stage depends on tidal or

climatically induced water-level changes which enable establishment of a high level *drift* or *strandline*. Sand blown up the shore settles around this drift which contains seaweed or sea-grass debris and often some animal material. Drift lying on the shore promotes water retention in the underlying sediment and provides a supply of nutrients. In addition, temperature fluctuation beneath drift may be dampened to only a third of its original range. There are therefore good germination conditions for salt-tolerant seeds. In northern boreal climates strandline plants tend to germinate in late April and May after deposition of drift by the equinoctial spring tides.

One of the commonest strandline plant genera is *Atriplex*, like *Salicornia* a member of the Chenopodiaceae, a family specialization in coastal habitats. *Atriplex* species, some of which are common to North American and European Atlantic shores, reflect latitudinal and other climatic change; *A. arenaria* extends from Florida to New Hampshire but is replaced to the north by *A. glabriuscula* and *A. sabulosa*. In Europe the latter two species hardly penetrate the Baltic where *A. littoralis* is fairly common in the south and *A. longipes* in the north. *Salsola*, *Honkenya* and *Cakile* are also widely distributed. Strandline plants are said to have a high nitrogen requirement which is met by bacterial decomposition of the drift, although laboratory work with *Atriplex* hastate has demonstrated survival with very low nitrogen levels. The presence of drift line plants leads to further accumulation of sand, thus raising the beach level so that dune pioneer species can become established. In Europe *Agropyron* (*Elytrigia*) *junciforme* and *Elymus arenarium* are important species. *E. mollis* plays a similar role on some Canadian shores and *Uniola panaliculata* on the United States eastern seaboard. These grasses rapidly produced a vertical and lateral root and rhizome system which binds the sand together. Reproduction is by seed or from rhizome fragments. It is known that *Elymus* seeds have a water-soluble germination inhibitor so that growth is not initiated in dry periods. On germination a vertical seedling root is produced rapidly. This penetrates to deeper and moister sand layers. Lateral roots and then rhizomes are established which spreads the plant outwards.

Accumulations of drifted seaweed provide habitats for a variety of invertebrates. The marine component is limited but usually includes one or more amphipod species; 'terrestrial' species include seaweed-fly larvae, such as *Coelopa*, small oligochaetes, beetles and mites. The species composition varies with size and depth and, at least in Europe, there is a decline in diversity with increasing latitude (Remmert, 1964).

Dune ridge formation : The process of sand transport by wind was studied first in deserts but the same principles apply to beach dune sand. If surface wind is in excess of about 4.5 ms^{-1} some sand particles will be lifted from the sand surface. They then fall back, setting other particles into motion so that these appear to jump from the sand surface, a process known as *saltation*. Sand leaving the surface may enter a higher (and more laminar) air stream and thus be carried greater distances before falling. However, the presence of vegetation decreases air flow considerably. As a net result,

sand coming into the vicinity of pioneer growth is likely to settle and remain on the sediment surface. Embryonic dunes are formed providing that upward and lateral grass growth can keep pace with deposition; upward rates of 30 cm y^{-1} have been reported.

The pioneer grasses are unable to build high dunes probably because of moisture limitation. Formation of higher dunes is through the action of other grass species, in particular *Ammophila arenaria* and its congener *A. breviligulata* (both commonly known as marram grass). Various other species may be involved in dune building, for example in Japan a sedge and *Wedelia prostrata* (Compositae) are important. *Ammophila*, like the foredune grasses, can establish itself from seeds or from rhizome fragments. The plant is able to maintain an above-ground portion by leaf growth, providing accretion is not rapid. However, with increased deposition *Ammophila* keeps pace by producing long vertical stem (rhizome) internodes which root adventitiously near the sand surface. Horizontal rhizomes are also produced as the plants grow. Previous years' sand accretion rates can be measured from the buried vertical internode lengths since the nodes form at the level of each year's leaf production. Rates of up to 1 m y^{-1} have been found.

As mentioned, the growth of vegetation slows air flow so that surface wind speed may be reduced below the level necessary for erosion. However, as dunes grow upward they interfere with laminar wind flow. As air is forced upward by the dune face the wind speed increases to reach a maximum near the dune crest. After passing the crest, wind speed lessens and there is usually a lee or backface vortex. Sand carried off the forward slope will be deposited on the lee face. Thus the dune is moved slowly backwards. The rate of movement depends on wind speed, particle size and the degree of protection afforded by vegetation. Values of 9 m y^{-1} have been recorded in Finance, but more usual rates are $1\text{--}5 \text{ m y}^{-1}$. This type of dune movement tends to occur on exposed coasts and where the sand supply is limited: thus it is more characteristic of eroding than of accreting coastline. In an idealized dune system one might expect to find a series of small *foredunes* followed by a series of *dune* ridges and intervening hollows (stacks) parallel to the shoreline. If the direction of the strongest and of the prevailing winds coincides, the highest dunes lie some distance inland. If these winds are opposed with the prevailing wind being offshore the highest dunes will lie near the beach. Due to the uneven cover of vegetation, dune crests rarely erode evenly and as the wind is funneled into eroding areas an extensive 'blow-out' may be produced. This may have the effect of bending part of the main dune ridge backwards; the concave section thus produced may detach from the main ridge and travel backwards as a *parabolic dune*. Travelling dunes often bury vegetation the remains of which are uncovered later. Whole woods may be overwhelmed and later reappear as 'ghost-forests'.

On coasts with a greater sand supply and gentler winds, dunes can form and become stable *in situ*, new dunes forming to seaward as the coast builds outwards. Neither moving nor stable dunes are necessarily aligned at 90° to wind direction but, probably due to geostrophic forces, are aligned

at a somewhat lesser angle. Dune stability can result from vegetational succession, climatic change or coastal management. Stabilized dunes, sometimes referred to as fixed or grey dunes, as distinct from moving or yellow dunes, can persist for centuries. On the Dutch coast two main dune series are recognizable. One of these is pre-Roman and the other appears to have formed between the twelfth and seventeenth centuries. As dunes develop the organic content and nutrient level of the soil increases (that is, following an initial decline from tidal litter conditions). The first *Ammophila* dune 'soil' will have organic and nitrogen contents of less than 0.5 and 0.01% respectively. However, by the stage of grass-sward formation organic content exceeds 1%. Little is known about microfloral development in dunes, but some nitrogen fixation occurs and there is some association of the bacteria with plant roots. This is certainly the case with some dune scrub species such as *Hippophae* (sea-buckthorn) which has root nodules. Fungal mycorrhizae are also important. Recent work on Australian dunes has shown that fungi are commonly associated with colonizing plants, whereas it had previously been thought that they were important only in older base-deficient dune soils (Jehne and Thompson, 1981).

The water-holding capacity of dune soils increases as the organic content builds up. There is a considerable shortage of water in young dunes and this limits the number of species able to take part in colonization. In most dunes the primary source of water is rain, but once it has percolated through the dune to the underlying water table it is unavailable to the dune ridge plants. Three main strategies have evolved in response to this problem. (1) Dune annuals undergo their vegetative cycle in autumn and spring or other wet seasons and survive the dry periods as seeds. Perennials either (2) produce a deep rooting system and are thus able to draw water from a greater vertical zone, or (3) produce an extensive shallow shower-exploiting root system. Water availability is enhanced by diurnal temperature change which leads to dew formation, the dunes acting as large water condensers.

Water percolation has two important chemical effects on dune soils. Dune sands often have an initially high calcareous content (i.e. are base-rich) due to shell fragments, but this is gradually leached out so that after 300-400 years the top 1 cm of soil is base-deficient. This affects the plant species composition and a visual sign of this process is colonization of the dune by species, such as heathers, characteristic of more acid soil. Nutrients are also washed out of the soil and enter the general water table. Rain itself contains small but significant amounts of nitrate and phosphate which will help support the dune vegetation, especially lichens which are able to absorb nutrients directly into their thalli.

As dune soils develop they become suitable for further plant species. There is an initial increase of diversity through the addition of various lichens, bryophytes, grasses and herbs. Some of these may participate in *hummock* formation. This is a characteristic feature of many dunes where local erosion occurs – small raised clumps or islands persisting because of their binding vegetation. After

grass/herb layer formation, shrubby species such as buckthorn, rose and willow may become established. Some of these, for example *Salix repens* in Europe and *S. lasiolepis* in the United States, also help in hummock production. The final stage of natural dune succession is generally woodland or forest of, for example, *Abies*, *Pinus* or *Quercus*.

Dune slacks : Erosion of open sand surfaces may continue until the damp sand immediately above the water table is reached or until a layer of less erodable material such as shell or pebble is uncovered. Thus the hollows or slacks left after dune ridge formation tend to have moist or coarse floors. Erosion will not always proceed to such depths due to the sheltering effect of ridges or to sand redeposition. It is possible to distinguish three types of dune slack. These are (1) dry slacks which are generally grass-dominated; the water table lies 1 m or more below the surface and although there are a few deep-rooted species the majority of the vegetation is rain-dependent; (2) wet slacks in which the water table is shallow, moss growth is often extensive and blue-green prokaryotes such as *Nostoc* are common (these often provide good orchid habitat); (3) flooding slacks where the water table remains close to the surface in summer and covers the ground in wet seasons; here the vegetation is characterized by amphibious species such as *Polygonum* (knot-grass) or *Ranunculus*. Dune systems may develop a variety of more or less permanent water bodies. A recent study by Timms (1982) demonstrates decline in alkalinity and associated changes with distance from the foredunes.

Dune fauna : Dune vegetation provides a resource for terrestrial animals. Some dune meadows are grazed by domestic cattle but the dominant herbivore is often the rabbit. Grazing tends to maintain species diversity by providing a sward with some open areas for annual plants and by preventing scrub establishment. However, areas can be impoverished by overgrazing. A large range of invertebrates occurs in dune systems. Foredunes are the least suitable habitat because of their aridity and considerable temperature fluctuations (even in Britain the sand surface temperature may exceed 60°C). Some thick-shelled snails benefit from the high calcium availability and are able to escape the worst temperature stress by attaching to marram above the sand surface. The snail *Cepaea* has been the subject of several studies in population genetics. The insect community, particularly of beetles, wasps, flies and consequently, the spider fauna is extensive. Actively hunting forms such as wolf-spiders predominate. Spiders in their turn are hunted by other wasp and by some vertebrates.

Mangroves

Mangroves form dense thickets consisting of some of the few terrestrial plants capable of tolerating fully saline conditions. They occur along protected sedimentary shores especially in tidal lagoons, embayments and estuaries (McNae, 1968). Although they can penetrate far inland they are never totally isolated from the sea. The rigours of life at the interface between land and sea has produced a vegetation which is of relatively low diversity when compared with other tropical and indeed temperate plant communities. Their emergent, evergreen canopies are inhabited by a variety

of terrestrial animals including insects, birds and arboreal mammals, whilst their complex root systems, and the attendant thick silty sediments in which they grow, provide suitable habitats for many marine invertebrates. In addition, numerous fish and crustaceans move in and out of the mangrove system with the tides. Mats of filamentous algae (e.g. *Cladophora*, *Vaucheria*) present on the substratum trap fine sediments, and leafy lagae (*Bostrychia*, *Galidium*) attach directly to the mangrove roots.

The term mangrove has been used to describe both an ecological group of flowering halophytic shrubs and trees (upto 30 m high) belonging to several unrelated families, and the complete community or association of plants which fringe sheltered tropical shores. The term 'mangal' is sometimes used to describe the latter. It is estimated that some 60-75% of tropical coastlines may be fringed with mangroves (Walsh in Reimold and Queen, 1974). They are important to man as a source of timber, charcoal and tannin, their stabilizing properties protect coastlines from major erosional damage by tropical storms and they provide an important nursery ground for many commercially important fish and shellfish. For an account of the economic value of mangroves (Christensen 1983).

Geographical distribution

Although mangroves attain their maximum development and diversity along moist warm coastlines of the tropics where rainforests are the climax vegetation, some species (e.g. *Avicennia marina*) do extend northwards to Japan and Bermuda and southwards to Australia and northern New Zealand. Their geographical limits appear to be set by low temperature. Sensitivity to temperature varies with species, but mangroves generally cannot survive air temperatures below -4°C . Optimal climatic conditions occur where the temperature rarely falls below 20°C and where seasonal variations are $<5^{\circ}\text{C}$. Dispersal ability is also clearly an important factor determining geographical distribution. The capacity for some seedlings to disperse over long distances may account for the wide distribution of certain mangrove species (e.g. *Rhizophora mangle*)

Two main groups of mangroves are recognized, those from the Indo-Pacific region and those from the Atlantic. Only a few genera (no species) are present in both regions. Considerably more species and genera occur in the Indo-Pacific (60 species compared with 10 species in the Atlantic province). The greater diversity of species in the Indo-Pacific has led to the suggestion that mangroves may have spread outwards from the geographical region. This 'Centre of Origin Theory' however, is largely inconsistent with the fossil record (McCoy and Heck, 1976). Alternatively, the greater diversity could reflect the general lack of long-distance dispersal by the propagules of most mangrove species, together with the greater opportunity for speciation along the extensive coastline provided by the numerous geographically isolated islands in that biogeographical province.

Mangroves may be locally absent within the latitudinal range of the community type for various

reasons. The forces responsible for such regional distributional patterns fall into two broad categories: those which *prevent* propagules from becoming established (e.g. unfavourable soil conditions; insufficient protection against tidal currents; adverse shore profile) and those which *disrupt* established mangroves (e.g. fires; hurricanes; coastal erosion). Within any given geographical region local conditions will therefore favour some species but not others.

Zonation

One of the most distinctive features of mangrove vegetation is the occurrence of species or groups of species in discrete bands or zones running across the marine-terrestrial environmental gradient. Each zone is typically dominated by a single species. Although considerable emphasis has been focused on detailed descriptions of the species composition and arrangement of these zones in the major mangrove systems of the world, no universal zonation scheme such as that described for rocky shores has so far emerged. Species composition varies and even whole zones may be transposed between one area and another. In areas where mangroves are restricted in area (e.g. where the tidal range is small) no zonation patterns may be discernible. At the limits of mangrove distribution, few species may be represented, with consequential implications for zonation. Numerous physical and biological factors have been regarded as important determinants of mangrove zonation. These include the following.

1) Tidal regime

This is possibly the single most important factor in determining mangrove zonation. The number of consecutive days that a particular zone is exposed by the tides seems to be especially important for the establishment and subsequent survival of mangrove propagules. Aerial exposure dries the soil and leads to an increase in salt content. As very high salt concentrations develop in the upper soil layers, mangroves tend to be replaced progressively by halophytes like *Salicornia*, and in extreme cases a bare salt desert on the landward margins of the mangrove forest may develop. Tidal flushing also influences production, in particular by transporting oxygen and nutrients to the root zone and by removing toxic sulphides and excessive amounts of salt. Accretion and erosion are also related to tides.

2) Salinity

Evidence suggests that mangroves are facultative rather than obligate halophytes. Their tolerance of saline conditions, however, enables them to survive in environments which are unavailable to their principal competitors (e.g. tropical rain forests). It has been suggested, therefore, that salinity may simply serve as a 'competition eliminator' (Thom, 1967) and not a principal determinant of zonation. Salt concentration in the substratum is highly variable and depends on evaporation, tidal flushing and freshwater inputs from rainfall, subterranean and terrestrial run-off. Species vary in their tolerance

or adaptation to salinity (Bunt *et al.*, 1982) and each accordingly occupies the appropriate zone on the environmental gradient. Even within any given species, seedlings and established plants may differ in their relationship to salinity. Soil salinity influences growth (and therefore size) through its effect on photosynthesis and respiration rates. Hypersaline conditions lead to stunting. Salt spray may be important at the seaward margins of the mangrove swamp.

3) *Geomorphology and climate*

These are important largely through their effects on other factors (e.g. Semenuik, 1983). The nature of the sediment (e.g. degree of aeration, water content), the pattern of tidal inundation, drainage and soil salinity all depend to a large extent on local topography. Erosion can truncate or even alter zonation patterns. Mangroves are often most luxuriant in areas of high precipitation. Rainfall is especially important at the landward margins where aridity and high salt content of the soil during neap tides can potentially exclude mangroves. High temperature increases evaporation which in turn increases soil salinity; it also accelerates the rate of oxygen consumption by micro-organisms and can thus lead to anoxic conditions.

4) *Species interactions*

Studies on mangrove zonation have concentrated on the importance of the physical environment and much less is known about the biological determinants of zonation. Competitive ability is probably important in determining which species become established within any particular zone. Dense canopies of *Rhizophora* and *Ceriops* considerably reduced the amount of light reaching the understorey and prevent colonization of open-canopied forms like *Avicennia* whose sporelings are less tolerant of shade. Sometimes species share the same optimal environment and the outcome then depends on relative competitive ability. *Rhizophora* and *Laguncularia*, for example, achieve maximum growth in the intertidal zone in Florida (Ball, 1980). When these occur concurrently *Laguncularia* is eventually forced into suboptimal areas above high-water mark. Such competitive exclusion of one species by another may be widespread, though detailed investigations are generally wanting. The fauna may also be important in determining species distribution patterns. Beetle larvae (e.g. *Chrysobothrus*) feed on and kill *Rhizophora* mangle and the subsequent build-up of plant debris makes the environment more suitable for *Avicennia germinans*. Some crabs strip bark and consume mangrove seedlings. The isopod *Sphaeroma terebrans* bores into prop roots; damaged trees may be severely weakened and thus more easily eroded by waves and currents.

5) *Man's influence*

Stresses induced by man are usually more intense and non-selective than those which occur naturally. Mangroves are cut for their tannin properties, for construction timber and for charcoal. Excessive logging can lead to changes in species composition. Mangrove swamps act as sediment

traps – when destroyed this sediment may be carried out to sea where it can smother neighbouring coral reefs. Drainage schemes, thermal stress from power plants and the extensive use of defoliants (e.g. during the Vietnam War) all influence the local distribution patterns of mangroves.

6) *Stochastic (· chance) factors*

Sudden unpredictable catastrophes can befall any community and bring about gross changes in species composition. Hurricanes and lightning fires, apart from their direct destructive effects, open up the mangrove canopy and thus enable light-demanding species to become established. They may therefore serve to maintain diversity in much the same way as physical disturbance and predation prevent competitive monopolization on rocky shores. However, in Florida, areas of mangrove destroyed by hurricanes were subsequently invaded by dense stands of the fern *Acrostichum aureum* which acts as a barrier to recolonization by more light-demanding mangrove seedlings. The effects of hurricanes are variable and depend on wind direction and velocity, local topography and vegetation types; some species (e.g. *Avicennia germinans*) are especially resistant and it is easy to envisage that under a regime of regular hurricanes it could become a dominant species. There may be a degree of randomness in mangrove colonization, the species composition of seedlings in virgin areas of Florida everglades, for example, simply reflecting that of the propagules that ‘happened to arrive’ (Ball, 1980).

Several factors are therefore probably involved in determining mangrove zonation patterns. It seems unlikely that greater emphasis can be attached to any one particular factor since none of the physical factors is totally independent of the others. Changes in one factor lead to changes in other factors with inevitable consequences to the flora. The degree of importance attached to any one factor therefore probably varies from one area to another, depending on local conditions and species occurrence. However, one factor common to all mangrove communities is tidal inundation. The frequency and duration of inundation is often well correlated with zonation. Tides exert an over-riding influence through a multitude of intermediate factors, and these have a *collective* influence with a gradient roughly normal to the shoreline. Each species will therefore find its own particular place within this environmental gradient. The general pattern may be complicated by creeks, hollows or hummocks which produce *local* changes in environmental conditions and therefore corresponding changes in vegetation. Physical factors must also ultimately determine the relative competitive abilities of component species and it is probably a combination of physical and biological factors that determine the distinctive zonation patterns of mangroves. A multivariate approach to this problem is clearly long overdue.

Succession

One of the most controversial areas of mangrove ecology is whether zonation of species across the marine-terrestrial gradient constitutes true *succession*. Some authorities argue that zones do indeed

represent serial stages in the succession from an estuarine/mud-flat type of community to a climatic climax community (in this case a tropical rain forest) though any one of the later stages in the succession may be repeated before the succession eventually proceeds to the next stage. In some cases, physio-graphic conditions may not be suitable for transition to the tropical forest and a preclimax swamp forest effectively becomes the climax community. In this interpretation of zonation, the coastal zone is the pioneer stage and the more landward zones progressively later stages in the succession. The concept implies a gradual seaward movement of the system (hence the reputation of mangroves as land builders) and that mangrove vegetation has the ability to modify the habitat (e.g. by accumulating sediment) thus inducing subsequent vegetational changes. One community, therefore, succeeds another until an emergent forest is formed.

This classical view of zonation has been criticized largely on the grounds that it is inconsistent with geological data. An alternative explanation is that zonation represents a response of the mangrove ecosystem to external forces rather than a temporal sequence induced by the plants themselves. Mangroves only grow where sedimentary regimes are suitable, they therefore follow sedimentation rather than cause it, at least initially. However, once established, their net of prop roots acts as a sedimentary weir, slowing water movement and allowing any load of entrained particles to settle out. As with most controversial issues the truth probably lies somewhere between the two views. Certainly mangrove vegetation is sensitive to changes in the physical environment. Classical serial succession, however, may provide an appropriate generalization for many mangrove communities.

Adaptation of mangroves

Mangroves exhibit numerous physiological and structural adaptations to their unusual physical environment. Consequently they provide one of the most remarkable examples within the plant kingdom of convergent evolution between several taxonomically unrelated groups (12 or so genera from 8 families) living within the same habitat. Mangroves become established on sedimentary shores where wave action is minimal. The lack of vigorous water movement results in the gradual accumulation of fine-grained muds, a process which is subsequently enhanced by the mangroves themselves. Mangroves must, therefore, be adapted not only to salinity stress but also to growth in waterlogged, anoxic sediments.

All mangroves have evolved a system of shallow, laterally-spreading cable roots. This underground system is richly supplied with finer anchor and absorptive roots. Above ground roots are of two main types – pneumatophores (e.g. *Avicennia nitida*) and prop roots (e.g. *Rhizophora mangle*). Pneumatophores are negatively geotropic and extend upwards through the soil surface. They arise from the cable root system and may be branched or unbranched. The extensive development of pneumatophores enables species like *A. nitida* to grow in extremely anoxic sediments. Stilt or prop roots (= rhizophores) emerge from the trunk well above ground and curve downwards into the surface

mud. They provide firm anchorage and allow exploitation of particularly soft sediments. The type of root system will therefore to some extent serve to determine zonal position. Both types of roots contain numerous small pores or *lenticels* which are permeable to air but not to water. These communicate with an elaborate system of *aerenchyma* tissue and thereby provide rapid transport of gases to the underground root system. Oxygen within these intercellular passages is utilized when the roots are submerged. If the lenticels are experimentally coated with grease then the oxygen within the root system is rapidly depleted.

Halophytes in general are faced with the major problem of absorbing water against the osmotic pressure gradient that exists between the saline sediment water and the plant vascular system. Their survival therefore depends on their ability either to accumulate high concentrations of salt (e.g. *Avicennia marina*) which they subsequently eliminate through special salt-secreting glands in the leaves, or to exclude salt from the roots by a process of ultrafiltration at the cell membrane (e.g. *Bruguiera gymnorhiza*). In the latter case any salt that does enter the plant can be stored in old, succulent leaves which are periodically shed. The complex physiological mechanisms underlying these responses are reviewed by Scholander (1968).

The osmotic pressure gradient between soil water and plant sap controls the rate at which water is supplied to the leaf tissue in order to support transpirational water loss. When this supply is deficient (e.g. at higher salt concentrations), the stomata, which are set in pits on thick waxy leaves (xeromorphic adaptations) close in order to conserve water. This, however, also reduces the uptake of gaseous carbon dioxide, which in turn depresses photosynthesis and presumably therefore primary production. Succulence, associated with the presence of water storage tissue in the leaves, is a further adaptation to the *physiological dryness* of the mangrove habitat.

Mangroves therefore have several ways of controlling salt concentration and it is perhaps not surprising that different species will vary in their salinity tolerances, with consequential implications for local distribution patterns.

Many mangroves (e.g. *Rhizophora*, *Bruguiera*, *Ceriops*) are viviparous, having unusual 'seeds' that actually germinate whilst still attached to the tree (Rabinowitz, 1978). Each propagule produces a long torpedo-shaped fruit (= hypocotyls) and when these are released they plummet vertically into the mud and immediately become established. In some species, however, they float away on the tide, sometimes being transported over long distances by ocean currents. The dispersal properties of propagules also appear to be correlated with mangrove zonation. Some mangroves (e.g. *Avicennia*) are not viviparous but have seedlings that can withstand prolonged periods of tidal inundation by virtue of their ability to respire anaerobically.

Mangrove fauna

The fauna of mangroves, derived largely from adjacent terrestrial and marine habitats, has been

less well studied than the mangrove vegetation itself. Broad patterns of zonation can be discerned both horizontally through the swamp and vertically from the sediment to the emergent canopy. Vertical stratification depends mainly on tidal inundation and salinity. The canopy is largely free from tidal influences and supports a fauna (insects, birds, mammals) that is essentially of terrestrial origin. These species generally show no special adaptations for life in mangroves, though many of them do feed on the marine species below. They also contribute to the nutrient input into the mangrove ecosystem in the form of faecal material. Root-holes and clefts retain rainwater and provide a valuable microhabitat, especially for insects. Below the canopy salt-tolerant species appear. The distribution of these essentially marine species depends primarily on their resistance to water loss and the availability of food and suitable substrata. Most species exhibit distinct preferences for particular habitat types which also occur elsewhere, and their association with mangroves is therefore often considered to be rather fortuitous. McNae (1968) provides an interesting general account of the macrofauna of Indo-Pacific mangroves.

The marine component of the mangrove fauna is largely dominated by various mollusks, crustaceans and fish. Littorinid snails (e.g. *Littorina scabra*) and many sedentary invertebrates (e.g. oysters, ascidians, sponges, and barnacles) are especially prominent on the trunks and extensive root systems. The mangrove snail *Cerithidea* feeds on the mud surface at low tide but climbs up the mangrove trunks with the rising tide to escape predators, mainly crabs and fish. The snails appear to have an inbuilt biological rhythm, since their upward migration starts before the tide reaches them. Lucinid lamellibranches may be locally abundant in mangrove sediments. Prawns (e.g. *Upogebia*), fiddler crabs (e.g. *Uca*), tropical land crabs (e.g. *Cardisoma*) and ghost crabs (e.g. *Dotilla*) burrow into the soft sediments; their burrows enable oxygen to penetrate the sediment and thus ameliorate the generally anoxic conditions of mangrove muds. Many of these crabs have mouthparts specialized for straining particles of organic detritus from the mud. Some show partial adaptation for air-breathing, the walls of their gill chambers being highly vascularized and lung-like. *Sesarma eulime* and *S. meinerti* are desiccation-resistant species and occupy the drier areas of the landward margins; *S. guttata*, and *S. catenata* are confined to shaded areas due to their inability to resist water loss. Mudskippers (e.g. *Periophthalmus*), a group of unusual gobioid fish, climb over roots and walk on their modified pectoral fins, or 'skip' using their fins and tail. Their large, frog-like eyes, set on top of the head, function best in air, whilst vascularized sacs present in the mouth and gill chambers enable them to breathe air. Mullet (*Mugil*), swimming crabs (e.g. *Callinectes*) and penaeid prawns visit the mangrove at high tide. Snakes, lizards and crocodiles may also be present.

□ 2.3.5 PRIMARY PRODUCTIVITY; TROPHIC INTERRELATIONSHIPS :

Despite the dearth of quantitative studies, available evidence strongly suggests that mangroves are highly productive systems. Estimates indicate that net primary production may be around 350-

500 g C m⁻² y⁻¹; phytoplankton production in adjacent nutrient-deficient tropical waters, on the other hand, rarely exceeds 50-75 g C m⁻² y⁻¹.

Much of this primary production eventually enters the aquatic systems as plant debris (e.g. leaves, twigs, scales). This is attacked by micro-organisms and is subsequently consumed by a wide variety of detritivores which fall broadly into three functional groups : (1) grinders (= shredders) which chew large leaf particles : (2) deposit feeders, which select smaller particles deposited on the sediment surface; and (3) filter feeders, which strain fine particles from the water. The detrital food chain of the Florida red mangrove (*Rhizophora*) has been extensively studied (e.g. Odum and Heald, 1975). The cyclical pattern of this chain arises from the multiple re-use of detritus. Detritivores digest the micro-organisms present on the surfaces of the decaying plant material before returning finely macerated debris to the environment as faecal pellets. This is subsequently reworked by micro-organisms (= 'faecal loop') leading to further break-down, and the whole cycle is repeated. The final stages in the main food web are small carnivores (e.g. minnows) and higher carnivores (fish-eating birds and 'game' fish). Black mangroves (*Avicennia*) occur at higher tidal levels in the Florida mangrove swamp than does *Rhizophora*. Mosquito and midge larvae are the principal detritivores in this zone and these in turn are consumed by a variety of 'forage' fish.

Mangroves are interface ecosystems with no precise boundaries to distinguish them from neighbouring systems on which they depend (Lugo and Snedaker, 1974). Nutrient cycles are driven by various physical and biological factors that control the rate of import and export of inorganic and organic materials. Of these, tidal flushing, riverine input, accumulation and decomposition of plant detritus and certain activities of the fauna are probably the most important. Little is known about secondary production in mangroves. However, it does appear that substantial quantities of nutrient may be exported from the system to be consumed elsewhere; the extent to which this occurs probably varies with local topographical conditions.

Thus although mangroves occur only in a relatively narrow coastal strip (comparable in this respect to the macroalgae which often festoon sheltered rocky shores in cooler latitudes) they are extensively distributed through-out the tropics and their contribution to coastal productivity in waters which otherwise tend to be rather unproductive may be extremely important. Their high levels of primary production enables mangroves to support diverse faunal assemblages which include species of birds and game fish which are of direct commercial and aesthetic importance to man.

Δ Coral Reefs

For their sheer beauty, species richness and multiplicity of biological interactions, coral reefs have few equals in the natural world. They are also highly productive systems in waters which are otherwise low in nutrients and phytoplankton. Their complex, wave-resistant structures are built almost entirely by biological activity. The dominant reef-forming organisms are the scleractinian (= stony)

corals which grow by means of an accretionary exoskeleton, but the lithifying (cementing) properties of crustose coralline algae, aided by the precipitation of magnesium carbonate particles from the water, are also very important in maintaining the structural integrity of the reef framework, especially in shallow waters.

Most of the structure of a coral reef consists of dead skeletal limestone deposited over years of accretion; living coral tissue forms only a thin veneer of interconnecting polyps over the surface of this extensive matrix. In addition to stony corals and coralline algae many other calcareous algae (e.g. *Halimeda*) and lime-secreting invertebrates (e.g. sponges, hydrocorals, octocorals, mollusks) also contribute to the formation of the reef.

Corals occur throughout the oceans, but colonial reef-building (hermatypic) corals are confined to the tropics and subtropics where calcification rates are greatly enhanced. *Lophophelia* is a branching coral found in cooler waters. Solitary corals do occur on coral reefs but these tend to be much larger than their counterparts from higher latitudes. One of the characteristic features of hermatypic corals is the presence of unicellular algae (*Symbiodinium*) within the gastrodermal tissue. By their influence on coral growth and calcification rates these symbiotic zooxanthellae play a fundamental role in the reef-building process.

Δ Classification

Three major geomorphological types of coral reefs can be recognized – fringing reefs, barrier reefs and atolls. Fringing and barrier reefs, however, are not easily separated from each other and are sometimes grouped together as *coastal reefs*, since both occur adjacent to land masses which they effectively protect against the damaging effects of waves. Coastal reefs are common throughout the coral-reef zones in all oceans whereas atolls occur mainly in the Indo-Pacific region.

Fringing reefs grow seawards, leaving an ever-widening reef that between the actively growing edge of the reef and the shoreline. They are the commonest type of reef and occur for example in the Red Sea and along the northern coast of Jamaica. *Barrier reefs*, like fringing reefs, run parallel to the coastline but are separated from the land mass by a shallow lagoon. The Great Barrier Reef (actually a collection of reefs) off the Queensland coast of Australia is almost 2000 km long and covers an area of more than 207000 km². *Atolls* are crescent or ring-shaped reefs surrounding a central lagoon. Parts of the reef platform may emerge as an array of small islands. Atolls occur in deep water and are not associated with any obvious land mass. *Patch reefs* are small circular or irregular reefs that arise from the floor of lagoons behind barrier reefs or within atolls.

Darwin suggested that atolls were formed from fringing reefs which developed on the shores of slowly subsiding volcanic islands. As these islands sank, vigorous upward growth of coral at the periphery of the reef produced first a barrier type reef and then, as the island finally disappeared

beneath the sea, an atoll. Continued growth along the outer edge keeps the reef at the surface but calmer conditions, sedimentation and bioerosion on the inside discourage vigorous growth and a lagoon develops. Darwin's 'Subsidence Theory' was subsequently confirmed in 1953 when corings at Eniwetok atoll in the Marshall Islands revealed volcanic rock beneath more than 1000 metres of reef limestone. Darwin's theory conveniently links all three reefs types in an evolutionary sequence. However, it must not be assumed that all fringing or barrier reefs are gradually evolving into atolls. Indeed coastal reefs unlike atolls, probably have very diverse origins and histories. Atolls are exceedingly old (upto 60 million years) but many other reefs are much more recent, some dating back only to the last Ice Age (< 15000 years).

Δ Distribution and limiting factors

Coral reefs are widely distributed wherever there is a suitable climate and substratum on which they can grow. They occur principally within the 20° isotherm but attain their maximum development where the mean annual water temperature is 23-25°C. No significant reef development occurs below 18°C, though higher temperatures (up to 40°C) can be tolerated. The absence of reefs along much of the western coastlines of South America and Africa is due to the strong upwelling of cold water. Occasionally, as in the western Atlantic, warm surface currents enable coral reefs to extend into higher latitudes. Recent evidence (Johannes *et al.*, 1983) suggests that coral distribution may be set through the interaction of temperature and competition with macroalgae. Lower temperatures which are suboptimal for corals appear to favour macroalgal growth; thus competition and not temperature *per se* may effectively exclude corals from latitudes where calcification rates are still potentially high.

Adequate *light* is essential for the photosynthetic activity of the zooxanthellae which permeate the coral tissue. Since light intensity decreases exponentially with depth active reef building rarely occurs below about 20-30 m. the compensation depth for coral growth occurs where the light intensity is approximately 15-20% of the surface intensity. Coral reefs are thus largely restricted to the margins of continental land masses and oceanic islands. Some corals, however, are adapted to lower light intensities and soon die if exposed to high surface illumination.

Coral reefs are very sensitive to low salinity and sedimentation. Consequently they are absent wherever there is an excessive influx of silt-laden fresh water even though conditions might otherwise be favourable. The absence of reefs along much of the tropical coastline of the Atlantic coast of South America for example is due to the influence of large rivers such as the Amazon and Orinoco. Optimal conditions occur in the Red Sea, around offshore islands in the Caribbean and Indo-Pacific regions and along the east coast of Australia, areas where terrestrial run-off is minimal or locally restricted. Coral reefs are tolerant of hypersaline conditions and in the Persian Gulf flourish at salinities exceeding 40%. Turbidity reduces light penetration whilst settling silt particles smother the feeding structures of the delicate coral polyps. Some corals, however, especially those with larger

polyps, produce moving sheets of mucus which effectively carry any sediment off the surface of the coral; these species are therefore more tolerant of turbid conditions.

Coral reefs thrive in regions of strong water movement. *Wave action* prevents sedimentation and keeps the water well oxygenated. Although the rigid calcareous skeletons of most hermatypic corals are particularly resistant to wave-shock, severe storms do occasionally cause extensive damage and are thus a major source of change in many coral-reef communities. Corals are unable to withstand prolonged *aerial exposure* and rarely grow much above the low spring-tide level.

Δ Reef morphology and zonation

In response to steep gradients in the physical environment, coral reefs, like rocky shores, are conspicuously zoned as a series of broadly overlapping bands running roughly parallel to the shore. Reef morphology is exceedingly complex, but whilst detailed structure and composition varies considerably from one locality to another the basic pattern in all types of reefs is similar. The pattern for atolls, although generally more complex than other reef systems, will serve as an example.

Basically an atoll consists of a sheltered central lagoon separating windward and leeward reefs on which there may be one or more low islands formed from consolidated coral debris. The windward *reef front* is characterized by vigorous coral growth in the form of massive finger-like spurs or buttresses running normal to the reef edge and extending down to the *reef terrace* at a depth of about 15-18 m. These spurs, which may be up to 300 m long depending on the slope of the roof, alternate with deep grooves or surge channels. This spur and groove formation effectively dissipates the destructive energy of the waves by channeling water up the grooves and over the sea ward edge of the reef. As the water washes back down the channels it is loaded with sediment and coral debris from the reef above; this erodes the grooves still further. The spurs and walls of the channels are optimal sites for coral growth and it is in this region that maximum accretion occurs. Dominant reef-forming corals (e.g. *Acropora*, *Montastrea*, *Pocillopora*, *Porites*) are abundant in this buttress zone. Sometimes corals from adjacent spurs coalesce over the surge channels forming canyons or tunnels which open through distinctive 'blow holes' on the reef crest.

The lower limit of the windward reef front is marked by the reef terrace below which the *reef slope* extends down, sometimes vertically, to the sea-floor. Although living corals occur down to depths of around 50-70 m their growth is here much less prolific and the reef slope is typically colonized by more delicately branching and plate-like forms. These corals (e.g. *Agaricia*) have large surface areas and are thus well adapted for intercepting the reduced amounts of light that penetrate to these deeper regions. Even individual species (e.g. *Montastrea annularis*) are known to change their morphology in response to differing light levels. Moreover, flattened plate-like forms are also less likely to roll down the reef slope to lethal depths should they become dislodged by waves or undermined by the biological activity (= bioerosion) of various reef boring organisms (e.g. fungi,

sponges, bivalves). Lower down the reef calcareous sponges, octocorals and non-hermatypic stony corals become increasingly important.

Immediately behind the buttress zone the seaward crest is dominated by crustose coralline algae (e.g. *Lithothamnion*, *Porolithon*). These form a smooth *algal ridge* which receives the full impact of the breaking waves. The crest is dissected by smooth-sided surge channels. Algal ridges are widespread on Indo-Pacific reefs but rarely occur on Atlantic reefs. A rich invertebrate fauna including crabs, shrimps and gastropods finds refuge in the subsurface cavities of the ridge where it is effectively protected from waves and predators. Zoanthids (anemone-like anthozoans) and vermetid gastropods may also be abundant in this zone whilst various urchins (e.g. *Diadema*) 'burrow' into the calcareous surface. Where wave action is less severe, reef crests may be dominated by robust corals such as *Millepora* (a hydrocoral) and *Acropora* whose heavy spreading branches form a seaward-projecting thicket.

Behind the algal ridge is the broad *reef flats* (back reef) which is very variable in character. Changes in salinity, temperature and sedimentation are most extreme in this part of the reef. This results in a large number of different habitat types and this zone is often variously subdivided. Reef rubble dislodged by waves sometimes forms a well-defined boulder zone at the outer edge of the reef flat. Conditions for coral growth are most favourable near the algal ridge where waves spill over the reef crest. Elsewhere on the reef flat reduced water circulation, accumulation of sediment and tidal exposure prevent extensive coral growth, though some massive species (e.g. *Porites*) may occur, especially in the deeper regions. Hard surfaces are frequently colonized by calcareous algae and unconsolidated sediments by sea-grasses, each community supporting a rich assemblage of invertebrates. In protected environments on the inner reef flats benthic foraminiferans may be abundant.

Inside the reef flat is the lagoon. This is rarely more than 60 m deep and the floor is typically covered with sediments derived largely from erosion of the windward reef. These sediments support numerous in faunal species whilst extensive beds of sea-grasses (e.g. *Thalassia*) and calcareous algae (e.g. *Halimeda*) may also be present. Reduced water movement and increased sediment create conditions which are generally unfavourable for extensive coral growth, though reefs varying in size from small knolls to larger patch reefs do occasionally rise from the lagoon floor wherever suitable conditions prevail.

The leeward reef is essentially similar to the windward reef though it is narrower and generally lacks the algal ridge, buttresses and surge channels. Extensive coral growth occurs at the seaward edge where, in addition to massive species, faster-growing branching species occur in the less wave-swept conditions. This region is rich in invertebrates (especially urchins) and fish.

Whilst lacking the complexity of atolls, the morphology and zonation of coastal reefs along

windward and leeward shores of land masses is essentially similar to that described respectively for windward and leeward reefs of atolls. In barrier reefs the inner margin of the lagoon is of course the coastline itself.

Δ Predation

The importance of predators on coral reefs is less well documented than for rocky shores. Many small predators (e.g. gastropods, crabs) feed on coral polyps or coral mucus, but their overall effect on community structure is probably minimal. The two major groups of coral predators are starfish and fish. *Acanthaster planci* is a large starfish which feeds on live coral polyps. Normally it occurs at moderate densities and its presence may even promote diversity since it feeds selectively on competitively dominant species. However, since about 1960 populations of *Acanthaster* have increased dramatically and have caused extensive damage to many Indo-Pacific reefs. No-one knows for certain what has caused these population expansions. One suggestion is that the systematic removal by shell collectors of the giant triton *Charonia tritonis*, a natural predator of *Acanthaster*, may be partially responsible (Endean, 1977); other explanations include pollution and natural population fluctuations. Following destruction by *Acanthaster*, reefs often become colonized sequentially by filamentous algae, calcareous algae and octocorals. Several years may elapse before new corals become established and it may take up to 40 years for complete recovery.

A striking feature of coral reefs is the presence of myriads of brightly coloured fish which swim in and around the coral heads.

Δ Physical disturbance

Coral reefs often appear to be exceedingly stable and persistent systems. Nevertheless, they can suffer extensive destruction, particularly from the effects of tropical storms. Hurricanes and cyclones often overturn coral heads, destroy most branching forms and even disrupt spur and groove formation. Colonies are seldom completely obliterated however and reef recovery is facilitated by vegetative growth of surviving portions. It is estimated that some 25-30 years may be required for reefs to recover completely from the effects of particularly destructive hurricanes. Prolonged periods of tidal emersion, especially when these coincide with extreme temperatures or heavy rainfall, can also result in extensive damage. Delicate branching species are less resistant to desiccation than are massive corals.

Although natural disturbances can be extremely damaging, moderate levels of disturbance may actually be beneficial to some corals by a process of thinning. The net effect is to open up space on the reef, thereby disrupting competitive exclusion and promoting coexistence. By transporting living fragments into new habitats storms may also serve as an effective means of coral dispersal. Coral reefs are thus strongly controlled by physical disturbance and in this regard are remarkably similar to many rocky intertidal communities.

II 2.3.6 DIVERSITY GRADIENTS AND BIOGEOGRAPHY

Coral reefs are extremely diverse, their complex three-dimensional structure providing suitable habitats for numerous invertebrates and fish. It is estimated that over 3000 species of invertebrates occur on certain reefs within the Great Barrier Reef complex. Polychaetes belonging to 103 species (13-1400 individuals) have been extracted from approximately 2.5 kg of dead coral rubble. One reason for this diversity is that the coral reef environment is a mosaic of different habitat types – wave-swept hard surfaces, sheltered sediments and carpets of algae and sea-grasses can often all be found within a relatively small area. Another reason is the relative stability of coral reefs over evolutionary time which has allowed ample opportunity for the coevolution of many specialized, finely-tuned interspecies relationships.

Two separate provinces of coral reef biotas can be recognized, one in the Indo-Pacific, the other in the Atlantic. The Atlantic province has significantly fewer coral species than the vast, yet relatively homogenous Indo-Pacific province (62 species in 36 genera and 700 species in 80 genera respectively). Coralline algae are generally less important in Atlantic reefs where many invertebrates (e.g. tridacnid clams, giant sea anemones and their commensal fish) are also absent. Horny gorgonian corals (sea-whips and sea-fans), are particularly conspicuous on Atlantic reefs, whereas soft corals (alcyonaceans) are more prominent in the Indo-Pacific. *Acropora* is a dominant reef builder in both provinces but whilst it is represented by 150 species in the Indo-Pacific, only 3 species occur on Atlantic reefs. In both provinces the number of coral genera declines as one moves away from the centres of distribution. Since temperature is strongly correlated (positively) with coral diversity, the latter also declines more rapidly along the N-S than along the E-W axis, particularly in the Indo-Pacific.

Indo-Pacific reefs are much older than those in the Atlantic. The latter are generally situated on shallow platforms cut by wave action during the Pleistocene period when sea-level was much lower than it is today. Most Atlantic reefs are therefore of fairly recent origin dating back only to the last Ice Age. Evidence suggests that the rate of evolution in the Indo-Pacific may be approximately double that in the Atlantic; this, together with the greater age and environmental stability of Indo-Pacific reefs, probably accounts for the difference in species diversity between these two provinces. Some of this difference could also be accounted for by the much greater size of the Indo-Pacific (an 'area effect').

Maximum coral diversity occurs at depths of about 15-20 m where light conditions and water movement are most favourable. On some Indo-Pacific reef slopes, coral species gradually replace each other with increasing depth. On Caribbean reefs, on the other hand, fewer species are present over the depth gradient and many of these simply respond to decreasing light levels by altering their gross morphology.

□ **2.3.7 SELF ASSESMENT QUESTIONS**

- a) **Divide oceanic regions using both morphology and wave process terminology. Describe the characters of offshore oceanic regions.**
- b) **Describe the ecology of coastal sand dunes.**
- c) **Write a short notes on :Coastal regions;**
Mangroves or Saltmarshes;
Coral Reefs.

□ **2.3.8 REFERENCES :**

Simon K. Haslett (2000). *Coastal systems*, Routledge, London page 218.

John Pethick (1984). *Introduction to coastal Geomorphology*. Arnold Hanimen, London. pp. 235

P J S Boaden and R Seed (1985). *An Introduction to Coastal Ecology*. Blackie USA : Chapman and Hall New York. pp. 218.

PART – I
Paper-I : Module- II
Unit – 04

Module Structure

- 2.4.1.1 Definition ;
- 2.4.1.2 Application of Hydrology to Environmental Problems
- 2.4. 2 Hydrological cycle, Global and Basin Hydrology;
- 2.4.2.1 Hydrological cycle
- 2.4.2.2 Cycle within the Basin
- 2.4.2.3 Global Hydrological Cycle
- 2.4.3 Estimation and Measurement of Hydrological Parameters
- 2.4.3.1 Colour, Taste, and Oudor
 - ▲ 2.4.3.1.1 Colour;
 - ▲ 2.4.3.1.2 Oudor;
 - ▲ 2.4.3.1.3 Taste;
- 2.4.3.2 Alkalinty
- 2.4.3.3 Temperature
 - ▲ 2.4.3.3.1 Factors of Temperature Variation
 - ▲ 2.4.3.3.2 Effects of Temperature variation
 - ▲ 2.4.3.3.3 Temperature variation: range and causes
 - ▲ 2.4.3.3.4 Determining temperature
- 2.4.3.4 Turbidity
 - ▲ 2.4.3.4.1 Meaning of the term
 - ▲ 2.4.3.4.2 Determining turbidity
- 2.4.3.5 Salinity
 - ▲ 2.4.3.5.1 Definition and source
 - ▲ 2.4.3.5.2 Characteristics of Water Affected by Salinity
 - ▲ 2.4.3.5.3 Sources of oceanic salt
 - ▲ 2.4.3.5.4 Determining salinity
 - ▲ 2.4.3.5.5 Desalinisation of water

■ 2.4.3.6 pH

▲ 2.4.3.6.1 Definition of pH

▲ 2.4.3.6.2 Explanation

▲ 2.4.3.6.3 The pH scale

▲ 2.4.3.6.4 Determining pH value of water-

▲ 2.4.3.6.5 Importance of pH value of water

■ 2.4.4 Study of trace elements and dissolved gases in water

▲ 2.4.4.1 Trace elements

▲ 2.4.4.1.1 Definition

▲ 2.4.4.1.2 Forms and occurrence

▲ 2.4.4.1.3 Range of concentration

▲ 2.4.4.1.4 Iron : Range of concentration and sources

▲ 2.4.4.1.5 Determining trace element concentrations

■ 2.4.4.2 Dissolved Oxygen

▲ 2.4.4.2.1 Range

▲ 2.4.4.2.1 Concentration

▲ 2.4.4.2.3 Biochemical Oxygen Demand

▲ 2.4.4.2.4 Chemical Oxygen Demand

■ 2.4.5 Self Assessment Questions

■ 2.4.6 Reference

Paper-I : Module- II
Unit – 04

■ **2.4.1.1 Definition ;**

Hydrology deals with the attributes of water, its distribution, availability and conservation for future generation. It encompasses the occurrence, distribution, movement, and properties of the waters of the earth and their environmental relations. Closely allied fields include geology, climatology, meteorology and oceanography. As water is considered as the basic resource on earth and as all the living beings demand water for metabolism and other physiological actions, washing and maintaining hygiene through making the pollutants mobile, the study of hydrology becomes relevant. Water, being the most dynamic and mobile resource on earth, thus become most sensitive resource demanding greater attention and study in scientific manner (Vissman et.al.,1989).

■ **2.4.1.2 Application of Hydrology to Environmental Problems**

It is true that humans cannot exist without water, it is also true that water, if mismanaged, or during times of deficiency (droughts), or times of surplus (floods), can be life threatening. Furthermore, there is no aspect of environmental concern that does not relate in some way to water. Land, air, and water are all interrelated as are water and all life forms. Accordingly, the spectrum of issues requiring an understanding of hydrologic processes is almost unlimited. As water becomes more scarce and as competition for its use expands, the need for improved water management will grow. And to provide water for the world's expanding population, new industrial developments, food production, recreational demands, and for the preservation and protection of natural systems and other purposes, it will become increasingly important for us to achieve a thorough understanding of the underlying hydrologic processes with which we must contend. This is the challenge to the water engineers, planners, policymakers, lawyers, economists, and others who must guarantee that future generations will be served with an abundant water supply and that a quality environment will prevail. (Vissman et.al.,1989).

■ **2.4.2 Hydrological cycle, Global and Basin Hydrology :**

▲ **2.4. 2.1 Hydrological cycle**

Hydrological cycle is a continuous process by which water is transported from ocean to atmosphere ,to the land and back to the sea. The evaporation of inland water and its subsequent precipitation over land before returning to the ocean is one example of a sub-cycle where the above definition is not fulfilled. The energy required for the water transfer from one sphere to another in connection to the change of state from liquid to vapour or solid is supplied from solar energy. The quality of water is changed during the passage through cycle as for example the saline water become fresh when it enters into the atmosphere in gaseous state Vissman et al(1989).

The precipitation mainly rain is partly absorbed by the vegetation and is called as inception and the other part is infiltrated and the rest flows on the surface. The infiltrated water is either absorbed by the soil or flows downward as gravitational flow and ultimately contributed to the ground water storage. This ground water again emerges on the surface through spring or withdrawal by man. The surface flow some times flow on the surface as sheet flow and gradually concentrates to the channels and thus rivers are developed and ultimately meets the ocean. Some times part of this water is arrested on the surface in the lowlands, ponds, lakes or swamps in the form of depression storage.

The evaporation may happen from every section continuously either at high or at slow rate. This may happen from ice surface, water bodied both inland and ocean, from leaf surface or from the surface of a barren land.

The complete water cycle is global in nature. World water problems require studies on regional, national, international, continental and Global scales. The total supply of fresh water is limited and very small in comparison to the salt water. The water flowing in one country can not be used by other at the same time. Water is highly mobile in nature that means the water content and moisture can not stay at one position and in one form for long. On an average, the residence time for moisture in the atmosphere is 8.2 days which means water is highly mobile and so rational use can only get maximum utility from water.

Although the concept of hydrological cycle is simple, the phenomenon is enormously complex and intricate. It is not just a large cycle, but rather is composed of many interrelated cycles of continental, regional and local extent. Although the total volume of water in the global hydrological cycle remains essentially constant, the distribution of this water is continually changing on continents, in region, and within local drainage basins.

▲ 2.4.2.2 Cycle within the Basin

Drainage basin can be considered as the complete hydrological unit within which all the components of global hydrological components and their systematic interactions are present at a local scale. The understanding of that relation help in the planning for rational management and augmentation of the water use. The natural supply from precipitation within the basin is definitely marked by the water divide and so the amon for a certain rainfall intensity is constant and can be measured. The other components of the cycle is also vivid and measurable. The division of the precipitation into infiltration, inception, evaporation and depression storage is conspicuous like global hydrological cycle.

▲ 2.4.2.3 Global Hydrological Cycle

In true sense the complete hydrological cycle is global in nature as the total supply of water in the global scale is finite and is distributed among the different component of cycle. Though for the convenience of analysis we generally discuss about regional, local and even basin cycle. The global annual water balance can be discussed in respect to the total annual precipitation on land taken as 100. it can be seen that evaporation from the land surface consumes 61 % of the precipitation, the remaining 39% forming runoff to the ocean, mostly as surface water. Evaporation from the oceans is 424 of which 385 directly returns to ocean in the form of precipitation. The remaining 39% of the atmospheric moisture over ocean blows towards land to compensate the surface runoff (Chow et al,1985).

		Ocean	Land
Precipitation	Km ³ /Yr mm/Yr	458000 1270	119000 800
Evaporation	Km ³ /Yr mm/Yr	505000 1400	72000 484
Runoff to Ocean			
Rivers	Km ³ /Yr	-	44700
Groundwater	Km ³ /Yr	-	2200
Total Runoff	Km ³ /Yr	-	47000
	mm/yr	-	316

■ 2.4.3 Estimation and Measurement of Hydrological Parameters

■ 2.4.3.1 Colour, Taste, and Oudor

Colour, taste, and oudor are important measurements for determining drinking water quality. Along with turbidity, colour, taste, and oudor are important from the standpoint of aesthetics. If water looks coloured, smells bad, or tastes swampy, people will instinctively avoid using it, even though it might be perfectly safe from the public health aspect. Colour, taste, and oudor problems in drinking water are often caused by organic substances such as algae or humic compounds, or by dissolved compounds such as iron.

▲ 2.4.3.1.1. Colour :

Colour can be measured visually by comparison with potassium chloroplatinate standards or by scanning at different spectrophotometric wavelengths. Turbidity interferes with colour determinations. So, the samples are filtered or centrifuged to remove suspended material.

▲ 2.4.3.1.2 Oudor :

Oudor is measured by successive dilutions of the sample with oudor free water until the oudor is no longer detectable. (Oudor-free water is prepared by passing distilled, deionized water through an activated charcoal filter.) This test is obviously subjective and depends entirely on the olfactory senses of the tester. Panels of testers are used to compensate for variations in individual perceptions of oudor.

▲ 2.4.3.1.3 Taste :

Taste is evaluated using three methods: the flavor threshold test (FTT), the flavor rating assessment (FRA), and the flavor profile analysis (FPA). For the FTT, water samples are diluted with increasing amounts of reference water until a panel of taste testers concludes that there is no perceptible flavor. In the FRA, a panel of testers is asked to rate the flavor from very favorable to very unfavorable. The oldest, and most useful, of the taste tests is the FPA, which measures both taste and odor of a water sample in comparison to taste and odor reference standards. The intensity of specific taste and odors are described on a 12-point, ranging from no taste or odor (0) to taste or odor (12) (Schwab et al 2002).

■ 2.4.3.2 ALKALINTY

Alkalinity measures the buffering capacity of the water against changes in pH. Water that has a high alkalinity can accept large doses of acids or bases without altering the pH significantly. Waters with low alkalinity, such as rainwater or distilled water, can experience a drop in the pH with only a minor addition of an acid or base.

In natural waters much of the alkalinity is provided by the carbonate/bicarbonate buffering system. Carbon dioxide (CO_2) dissolves in water to form carbonic acid (H_2CO_3), which dissociates and is in equilibrium with bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}) ions:

If an acid is added to the water, the hydrogen ion concentration is increased, and this combines with both the carbonate and bicarbonate ions, driving the equilibrium to the left, releasing carbon dioxide into the atmosphere. As long as there is bicarbonate and carbonate, the

added hydrogen ions will be absorbed by readjustment of the equilibrium equation. Only when all of the carbonate and bicarbonate ions are depleted will the addition of acid cause a drop in Ph.

The amount of bicarbonate in water is supplemented by naturally occurring carbonates, such as CaCO_3 (limestone) that dissolve as acidic rain water comes into contact with watershed soils or the stream bed. The CaCO_3 dissolves to form calcium bicarbonate ($\text{Ca}(\text{HCO}_3)_2$), which dissociates and increases the bicarbonate concentration in the water :



Alkalinity is determined by measuring the amount of acid needed to lower the pH.

In a water samples to a specific endpoint; the results are usually reported in standardized units as milligrams CaCO_3 per liter. Poorly buffered water may have alkalinities lower than 40 mg CaCO_3/L while water sampled from a stream flowing through a limestone or "karst" region may have alkalinities greater than 200 mg CaCO_3/L .

■ 2.4.3.3 TEMPERATURE

▲ 2.4.3.3.1 Factors of Temperature Variation

Temperature of water varies with latitude, altitude, season, time of the day, air circulation, cloud cover and the flow and depth of water. Lakes and reservoirs may also exhibit vertical stratification of temperature within the water column. In case of river water temperature increases with increasing distance from source (Rodda, 1976).

▲ 2.4.3.3.2 Effects of Temperature variation

Temperature affects a large number of physical, chemical and biological processes in water bodies and therefore, the concentration of many variables. As water temperature increases, i) the rate of chemical reactions generally increases together with the evaporation and volatilisation of substances from the water. ii) Increased temperature also decreases the solubility of gases in water such as O_2 , CO_2 , N_2 , CH_4 and others. iii) The metabolic rate of aquatic organisms is also related to temperature and, in warm waters, iv) respiration rate increases leading to v) increased oxygen consumption and v) increased decomposition of organic matter. Vi) Growth "rates also increases leading to vii) increased water turbidity and algal booms, when nutrient conditions are suitable.

Most aquatic animals, being cold blooded, are very sensitive to temperature changes. Most fish species can detect as small changes as 0.05°C and an increase of 5°C can completely disrupt their physiological functioning.

3.3.3 Process of Heating Cooling

Water gets temperature from solar radiation which is affected by latitude, i.e. angle of incidence and duration of day and turbidity factors. According to Kimball if in clear sky the isolation becomes 100%, it is reduced to 22% in overcast condition. The convection of ocean water brings temperature from ocean bottom to the surface and is distributed throughout the ocean body. This amounts to $0.95 \times 10^{-3} \text{ gm Cal/cm}^2/\text{min}$ to $0.152 \times 10^{-3} \text{ gm Cal/cm}^2/\text{min}$. The kinetic energy collected from the blowing wind may raise the temperature up to 0.2°C up to the depth of 50m. the condensation of water releases latent heat which is added to water at its surface.

The cooling is mainly done by re-radiation of heat energy, convection of temperature to the cool air above and consumption of latent heat for evaporation. According to Angstrom 90% of the water temperature is used for evaporation.

▲ 2.4.3.3.3 Temperature variation: range and causes

Surface waters occur usually within the temperature range from 0°C to 30°C although waters of hot springs may reach 40° C or more. These temperatures fluctuate seasonally with minimum occurring during winter or wet periods, and maximum in the in the summer or dry seasons, particularly in shallow waters. Ground water usually maintains a constant temperature which, for surfacial aquifers, is normally close to the mean annual sir. temperature. However, deep aquifers have higher temperatures due to the earth's thermal gradient.

▲ 2.4.3.3.4 Determining temperature

Temperature should be measured in *situ*, using a thermometer or thermistor. As temperature has an influence on so many other aquatic variables and processes, it is important to always include it in a sampling regime, and to take and record it during collecting of water samples. For detailed understanding of biological and chemical processes in water in water bodies; it is often necessary to take a series of temperature measurements throughout the depth of the water. This can be done with a recording thermistor linked to a pressure transducer, directly reading temperature with depth

■2.4.3.4 Turbidity

▲ 2.4.3.4.1 Meaning of the term

Turbidity results from the scattering and absorption of incident light by suspended particles present in water. The type and concentration of Suspended matter controls the extent of turbidity. In water, suspended matter generally consists of silt, clay, fine particles of organic and inorganic matter, soluble organic compounds, plankton and other microscopic organisms. Such particles may vary in size from approximately 10 nm to 0.1 nm in diameter

▲ 2.4.3.4.2 Determining turbidity

Turbidity should be measured in field but, if necessary, samples can be stored in the dark for not more than 24 hours Settling during storage and changes in pH leading to precipitation can affect the results during storage. Visual determination of turbidity of an open water body is usually done using the Secchi disk—a white disk of standard size (30 cm). The disk is lowered in water and the depth at which it disappears from sight of the observer, the so-called Secchi depth, is recorded. In extremely transparent water Secchi depths can be as high as 80 m. More sophisticated instruments generally compare the length of a light path through the sample against a standard solution and measure it in Jackson Turbidity Units (JTU). However, the most reliable method of turbidity determination uses nephelometry (light scattering by suspended particles) by means of a turbidity meter which gives values in Nephelometric Turbidity Units (NTU). Normal values range from 1 to 1,000 NTU and levels can be increased by the presence of organic matter pollution, other effluents or run-off with a high suspended matter content.

■ 2.4.3.5 Salinity

▲ 2.4.3.5.1 Definition and source

Salinity is the measure of the concentration of salts dissolved in water. It is the ratio of weight of dissolved material to that of sample water. 1902 International Oceanographic congress defined salinity as the total dissolved solids in one Kilogram of sea water when all carbonate is converted to oxides, bromine and iodine are replaced by chlorine and organic matter is oxidized to 480°C . The amount of salinity varies from 34 to $37\text{ }^{\circ}\text{‰}$.

The salinity of the oceans has resulted from constant addition of dissolved ions by rivers from terrestrial sources.

▲ 2.4.3.5.2 Characteristics of Water Affected by Salinity

The important physical and chemical characters of the soil are affected by the salinity. The density, rate of thermal expansion, refractive index, temperature to maximum density, temperature to freezing point, temperature to boiling point, rate of evaporation and intensity of scattering radiation etc. are affected by salinity.

▲ 2.4.3.5.3 Sources of Oceanic Salt

From the beginning of the earth's history, ocean water dissolves earth surface matter and those are concentrated in the sea. The washing of shores by waves are also another way for salt concentration. The dissolved salt drained into ocean through rivers is responsible for concentration of oceanic salt.

▲ 2.4.3.5.4 Determining salinity

Water is considered potable if it contains less than 0.5 grams of salt per a kilogram of sample. The simplest way of determining this is to evaporate a kilogram of water and measure the salts which are left behind in grams and express the results in g kg⁻¹ or ‰. This approach, however, is not practical for routine measurements and hence is replaced by a relation based on the chlorinity titration. In this method, chlorine content of the sample is determined first by titration with silver nitrate, following which the total salt content of the sample may be resolved on the basis of the rule of constant proportions: $S = 1.805 \text{ Cl}\% + 0.03$.

Apart from g kg⁻¹, salinity of a water sample is also expressed as the nondimensional figure of **practical salinity** (S). S is determined by measuring the electrical conductivity of water using salinometer. Since electrical conductivity is affected by changes in temperature and pressure, ideally salinometer should be used *in situ*.

Infractometer is another instrument used to measure the refractive index. Salinity is determined by a comparison between the Refractive Index of an unknown sample and that of a known standard sample.

▲ 2.4.3.5.5 Desalinisation of water

- To meet the ever-increasing demands for fresh water, especially in arid and semiarid areas, much research has gone into finding efficient methods of removing salt from sea water and terrestrial brackish waters.

Three of the processes involve evaporation followed by condensation of the resultant steam and are known as multiple-effect evaporation, • vapour-compression distillation and • flash evaporation. • Freezing is an alternate method, based on the different freezing points of fresh and salt water. • Electrodialysis is also used to desalt brackish waters.

■ 2.4.3.6 pH

▲ 2.4.3.6.1 Definition of pH

Puissance de Hydrogen or pH is a numerical system, developed in 1909 by a Swedish chemist, Sorensen, for expressing acidity or basicity. Technically, pH may be defined as either of the following:

- The logarithm, to the base 10, of the reciprocal of concentration of hydrogen ion, i.e., $\text{pH} = \log_{10} \times 1/\text{H}^+$
- The negative logarithm, to the base 10, of concentration of hydrogen ion, i.e., $\text{pH} = -\log_{10} \times 1/\text{H}^+$

▲ 2.4.3.6.2 Explanation

Water is a neutral substance with equal concentrations of H^+ and OH^- ions. The ionisation constant of water is 10^{-14} at 24°C . As the concentration (c) of H^+ and OH^- are equal, the value for each must be equal to 10^{-7} . Considering only the concentration of H^+ , because of its role in determining acidity, we have $\text{H}^+ = 10^{-7} = 0.0000001 \text{ gL}^{-1}$.

Since water is neutral, any substance with pH value of 7 is said to be neutral, while any value below this is acidic and above, basic.

▲ 2.4.3.6.3 The pH scale

The pH scale runs from 0 to 14—i.e. from very acidic to very basic. It should be noted that the scale is logarithmic, not linear. This means that at pH 7, the cH^+ is 10 times less than that of pH 6, 100 times less than that of pH 5 or 1000 times less than that of pH 4 and so on.

▲ 2.4.3.6.4 Determining pH value of water-

Ideally, pH should be determined *in situ*, or immediately after the sample is taken since, a number of natural factors can influence it. A rough indication of pH can be obtained chlorimetrically with indicator dyes. Accurate measurement of pH is usually undertaken electrometrically with an electrode probe. Many of the available electrometric pH meters are suitable for field use and continuous measurement and recording. If field measurement becomes impossible, samples must be transported to the laboratory in completely full and tightly closed bottles with out slightest contamination. As pH is temperature dependent, the water temperature must also be measured in order to determine the pH value accurately.

▲ 2.4.3.6.5.Importance of pH value of water

The pH is an important variable in water quality assessment as it influences many biological and chemical processes within a water body and all processes associated with water supply and treatment. At a given temperature, pH indicates the intensity of the acidic and basic character of the solution and is controlled by the dissolved chemical compounds and the biochemical processes in it. In unpolluted water, pH is principally controlled by the balance between the carbon dioxide, carbonate and bicarbonate ions as well as other natural compounds. Changes in the pH can indicate the presence of certain effluents, particularly when continuously-measured and recorded together with the conductivity of a water body. Diurnal variations in pH can be caused by the photosynthesis and respiration cycles of algae in eutrophic waters. The pH of most natural waters is between 6.0 and 8.5, although lower values can occur in dilute waters high in organic content, and higher values in eutrophic waters, groundwater brines and salt lakes.

■ 2.4.4 Study of trace elements and dissolved gases in water

■ 2.4.4.1 Trace elements

▲2.4.4.1.1 Definition

Trace elements are the chemical elements, chiefly metals, present in water in concentrations typically below micromolar level. The ability of a water body to support aquatic life, as well as its sustainability for other uses depends on many trace elements — also called micronutrients. Some metals eg, Mn, Zn and Cu present in trace concentrations are important for the physiological functions of living tissue and regulate many biochemical processes.

▲ 2.4.4.1.2 Forms and occurrence

Metals in natural waters can exist in dissolved, colloidal and suspended forms. The proportions of these forms vary for different metals and for different water bodies.

Generally, trace amounts of metals are always present in freshwaters (as well as seawaters) from the weathering of rocks and soils. In addition, particularly in developed countries, industrial wastewater discharges and mining are major sources of metals into freshwater. Significant amounts of metals also enter the surface waters in sewage as well as from atmospheric deposition (e.g., lead). Lead is still widely used as an additive in petroleum for automobiles and is emitted to the atmosphere in the exhaust gases, thereby entering the hydrological cycle.

▲ 2.4.4.1.3 Range of concentration

The assessment of metal pollution is an important aspect of most water quality assessment programs. The commonly included elements include Al, Cd, Cr, Cu, Fe, Hg, Mn, Ni, Pb and Zn. The concentrations of different metals in waters varies over a wide range (0.1 to $0.001 \mu\text{g l}^{-1}$) and can rise to concentrations which are dangerous for human health in some water bodies influenced by human activities.

▲ 2.4.4.1.4 Iron : Range of concentration and sources

Among metals, the iron is often considered a major element present in water. Its occurrence in aqueous solution is dependent on environmental conditions, especially oxidation and reduction. Flowing surface water, that is fully aerated, commonly contains not more than a few of uncompleted dissolved iron at equilibrium in the pH range 6.6 to 8.5. In groundwater, however, much higher levels can occur. In ground waters with a pH value of 6 to 8, ferrous iron (Fe^{++}) concentrations can be as high as 50 mg l^{-1} and concentrations between 1 to 10 mg l^{-1} are common.

The iron originates by solution at sites of either reduction of ferric hydroxides or oxidation of ferrous sulphide and the process is strongly influenced by microbial activity. Reduced ground water is clear when first brought from a well but becomes cloudy and then orange in colour, as oxidation immediately occurs with the precipitation of ferric hydroxide. High iron concentrations in groundwater are widely reported from developing countries like India, where iron is often an important water quality issue.

▲ 2.4.4.1.5 Determining trace element concentrations

In water, some 50 to 99.9 per cent of the total metal present is usually absorbed into suspended sediments. Consequently, monitoring and assessment programs generally include the determination of both total (unfiltered) and dissolved (filtered through a $0.45 \mu\text{m}$ filter) concentrations of metals.

■ 2.4.4.2 DISSOLVED OXYGEN

▲ 2.4.4.2.1 Range

One of the most important measures of water quality is dissolved oxygen. Oxygen, although poorly soluble in water, is fundamental to aquatic life. Without free dissolved and most invertebrates can not live. Dissolved oxygen is inversely proportional to temperature, and the maximum amount of oxygen that can be dissolved in water at 0°C is 14.6 mg/L . The saturation value decreases rapidly with increasing water temperature, as shown in Table 5-1.

▲ 2.4.4.2.2 Estimation

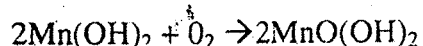
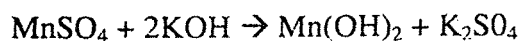
The amount of oxygen dissolved in water is usually measured either with an oxygen probe or by iodometric titration. The latter method, known as the *Winkler* test.

Table 5-1. Solubility of Oxygen in Water

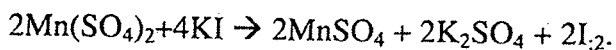
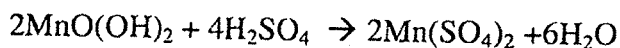
Water temperature (°C)	Saturation concentration of oxygen in water (mg/L)
0	14.6
2	13.8
4	13.1
6	12.5
8	11.9
10	11.3
12	10.8
14	10.4
16	10.0
18	9.5
20	9.2
22	8.8
24	8.5
26	8.2
28	8.0
30	7.6

was developed about 100 years ago and is the standard against which all other measurements are compared. The chemical reactions of the Winkler test are as follows:

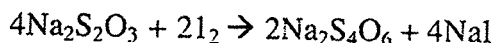
Manganous sulfate (MnSO_4) and a mixture of potassium hydroxide and potassium iodide (KOH and KI) are added to a water sample. If there is no oxygen present, the MnSO_4 will react with the KOH to form a white precipitate, manganous hydroxide (Mn(OH)_2). If oxygen is present, the Mn(OH)_2 will react further to form a brown precipitate, manganic oxide (MnO(OH)_2):



Sulfuric acid is added, which dissolves the manganic oxide and, in conjunction with the KI added earlier, forms iodine (I_2), which imparts a yellowish orange color to the sample:



The quantity of iodine is measured by titrating with sodium thiosulfate ($Na_2S_2O_3$) until the orange color from I_2 is no longer apparent:



Starch is added near the end of the titration because it turns deep purple in the presence of I_2 and gives a more obvious color endpoint for the test.

The quantity of $MnO(OH)_2$ formed in the second step is directly proportional to the available dissolved oxygen, and the amount of iodine formed in the second step is directly proportional to the $MnO(OH)_2$. Therefore, the titration measures a quantity of iodine directly related to the original dissolved oxygen concentration. Disadvantages of the Winkler test include chemical interferences and the inconvenience of performing a wet chemical test in the field. These two disadvantages can be overcome by using a dissolved oxygen electrode, or probe.

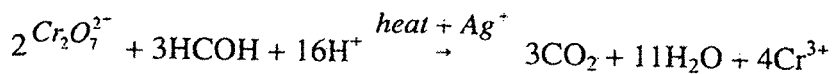
▲ 2.4.4.2.3. Biochemical Oxygen Demand

The rate of oxygen use is commonly referred to as biochemical oxygen demand (BOD). Biochemical oxygen demand is not a specific pollutant, but rather a measure of the amount of oxygen required by bacteria and other micro-organisms engaged in stabilizing decomposable organic matter over a specified period of time.

▲ 2.4.4.2.4. Chemical Oxygen Demand

One problem with the BOD test is that it takes 5 days to run. If the organic compounds were oxidized chemically instead of biologically, the test could be shortened considerably. Such oxidation can be accomplished with the chemical oxygen demand (COD) test. Because nearly all organic compounds are oxidized in the COD test, while only some are decomposed during the BOD test, COD results are always higher than BOD results. One example of this is wood pulping waste, in which compounds such as cellulose are easily oxidized chemically (high COD) but are very slow to decompose biologically (low BOD).

The standard COD test uses a mixture of potassium dichromate and sulfuric acid to oxidize the organic matter ($HCOH$), with silver (Ag^+) added as a catalyst. A simplified example of this reaction is illustrated below, using dichromate ($Cr_2O_7^{2-}$) and hydrogen ions (H^+):



A known amount of a solution of $K_2Cr_2O_7$ in moderately concentrated sulfuric acid is added to a measured amount of sample, and the mixture is boiled in air. In this reaction, the oxidizing agent, hexavalent chromium (Cr^{VI}), is reduced to trivalent chromium (Cr^{III}). After boiling, the remaining Cr^{VI} is titrated against a reducing agent, usually ferrous ammonium sulfate.

ale. The difference between the initial amount of Cr^{VI} added to the sample and the Cr^{VI} remaining after the-organic matter has been oxidized is proportional to the chemical oxygen demand.

■ 2.4.5 Self Assessment Questions :

1. Define Hydrology. How do you associate environmental problems with Hydrology?
2. How are the Environmental problems solved by allocation of hydrological Knowledge?
3. Describe Hydrological Cycle? How does the quality of water get changed while passing through the hydrological cycle?
4. What is the implication of Basin hydrology? Describe global water balance.
5. How do the amount of Dissolved Oxygen change with change in temperature? What are the methods practiced for estimating Dissolved Oxygen? What do you mean by BOD and COD?
6. What are the importance of analysing colour, taste and odor of water? What are the standard methods of their estimation?
7. Define Turbidity. How do you measure it?
8. Define Salinity. What characters of water gets changed with change in salinity?
9. How is the salinity Estimated? What are the methods of desalinization?
10. Define pH. How do the trace elements occur in water?
11. Describe how the amount of dissolved Oxygen differ according to temperature.
12. Define BOD and COD.

■ 2.4.6 References

- Andrews. W.A., Moore. DK and Le Roy. A.C. (1972) : A Guide to the Study of Environmental Pollution, Prentice-Hall. Englewood Cliffs: 260p [A.B]
- Chapman, D. (Ed: 1992) : Water Quality Assessments, Chapman & Hall, London. 585p.
- Chow, Maidment, and Mays (1985), Applied Hydrology, Mc Grawhill, pp2-17.
- Harvey, H.G. (1976) : Atmosphere and Ocean: Our Fluid environment. The Artemis Press, Ltd. Sussex: 143p.
- Schwab, O.G., Fangmier D.D, Elliot, W.J. and Frevert, R.K. (2002) Soil and Water Conservation Engineering, John Wiley and Sons Inc.
- Vissman, Lewis and Knap (3-13), Hydrology, pp 3-13.

M.A. / M. Sc. PART – I
Paper - I : Module - II
Unit – 05

Module Structure

- 2.5.1 Ground water
 - 2.5.1.1 Adsorbed water
 - ▲ 2.5.1.1.1 Hygroscopic Moisture
 - ▲ 2.5.1.1.2 Soil Moisture Film
 - 2.5.1.2 Free water (Ground Water)
 - ▲ 2.5.1.2.1 Ground water Zones-
 - 2.5.1.2.1.1 Capillary fringe;
 - 2.5.1.2.1.2 Ground water zone
 - 2.5.1.2.1.3 Aquifers
 - 2.5.1.3 Ground Water flow;
- 2.5.2 Wet Land
 - 2.5.2.1 Inland Wetlands (National Sites)
 - 2.5.2.2 Forest Wetland
 - 2.5.2.3 Wetland conservation
 - ▲ 2.5.2.3.1 Issues
 - ▲ 2.5.2.3.2 Inventory
 - ▲ 2.5.2.3.3 Classification
 - ▲ 2.5.2.3.4 Integrated approach
 - ▲ 2.5.2.3.5 Socioeconomic Issues
 - ▲ 2.5.2.3.6 Technological Advancements
 - ▲ 2.5.2.3.7 Prioritization of Inland wetlands for conservation
 - 2.5.2.3.7.1 Criteria
 - 2.5.2.4 Wetland Ecology
- 2.5.3 Watershed
 - 2.5.3.1 Importance of drainage basins or watershed
 - ▲ 2.5.3.1.1 As geopolitical boundaries
 - ▲ 2.5.3.1.2 In Hydrology
 - ▲ 2.5.3.1.3 In Ecology
 - ▲ 2.5.3.1.4 In resource management
 - 2.5.3.2 Major Watersheds in India
- 2.5.4 Self Assessment Questions
- 2.5.5 References

M.A. / M. Sc. PART – I
Paper - I : Module - II
Unit – 05

■ **2.5.1 Ground water**

Water remains in the soil in either solid, liquid and or gaseous state. Water of the soil can occur in the form of aqueous vapour, in the form of film adsorbed to the surface of the soil particle, as free (Gravitational) water, or water in bulk and as water in the solid state of aggregation, as for example, the ice.

■ **2.5.1.1 Adsorbed water**

Water which is adsorbed by the particle may occur either in the form of Hygroscopic Moisture or Soil Moisture Film.

▲ **2.5.1.1.1 Hygroscopic Moisture**

All mineral matter in the soil is covered with a very thin film of moisture. Soils which appear quite dry nevertheless contain an amount of moisture – the so called Hygroscopic Moisture. In technical literature Hygroscopic Moisture is called also as adsorbed moisture, contact moisture, or surface bound moisture. This form of soil moisture is in a very dense state and surrounds the surface of the individual soil particle as a very thin hull of film of water. Soil obtain it from either atmosphere, or from the aqueous vapour present surrounding them and held by adhesion. Hygroscopic moisture does not have any association with ground water and so does not respond to the fluctuation of ground water table. Hygroscopic Moisture can be removed by drying the particles at 105° C temperature. The change in the thickness of the Hygroscopic Moisture film may occur through an electrolyte, a change in temperature, a change in the air moisture content, a change in the pore water pressure and a change in the mechanical pressure on the soil.

▲ **2.5.1.1.2 Soil Moisture Film**

Film moisture forms in the soil by the condensation of aqueous vapour or remains there after the removal of the bulk of free water. The film moisture is attached to the surface of the soil particles as a hull or film upon the layer of hygroscopic moisture film. The film moisture is held by the molecular forces of considerable interstices but not as a large as in the case of the Hygroscopic Moisture Film. The film moisture is set in motion or migration

by application to the soil system an external energy potential called as primary potential- a thermal or electrical potential for example. The moisture of soil then migrates from places of higher temperature or higher electric potential to the lower.

■ 2.5.1.2 Free water (Ground Water)

Ground water is that kind of gravity water which fills up the pore spaces of soil and the soil is saturated and the water content by volume is equal to the volume of the voids. This water obeys the laws of hydraulics. The upper surface of the zone of full saturation of the soil, at which the ground water is subjected to the atmospheric pressure, is called ground water table (GWT). The ground water table is not a fixed surface plane, but is undulating, tending to parallel the undulations and contours of the surface topography. Ground water table fluctuates in different places-and –even in the same place- according to the amount of precipitation and to the season of the year; because of withdrawal of it for water supply, other industrial purposes, or lowering of the ground water table.

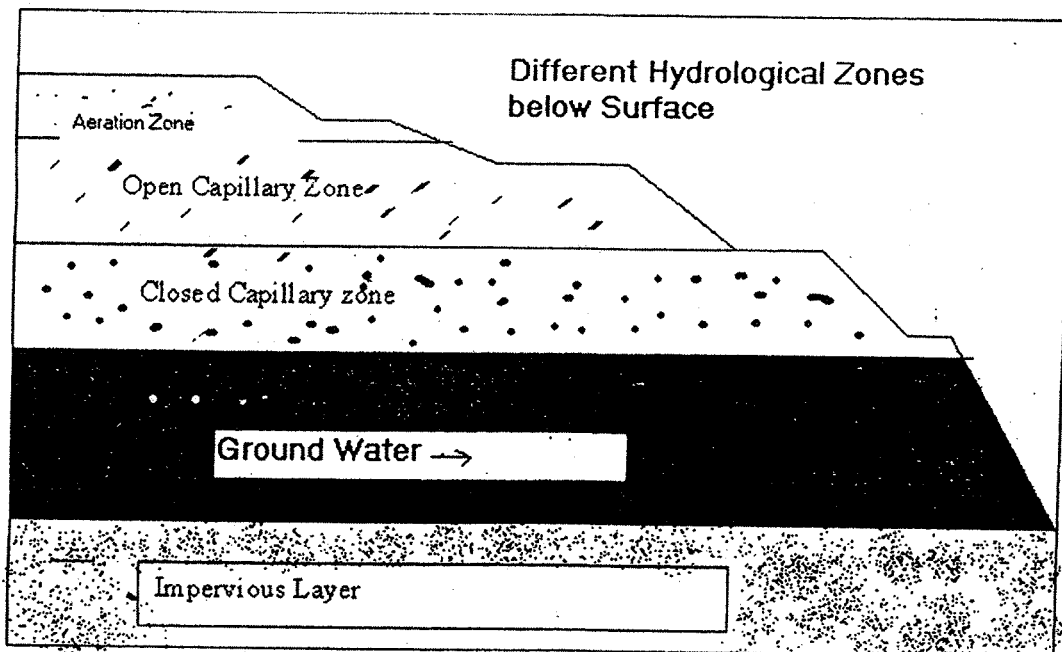
▲ 2.5.1.2.1 Ground water Zones-

● 2.5.1.2.1.1 Capillary fringe;

Just above the ground water table there lies capillary zone where water moves up and down through the capillaries. Capillary, in general is a phenomena of the rise or depression of the liquids in tubes having a bore so fine as to be comparable in diameter with a hair (capillus). The rise takes place and the liquid is held by means of a force called the surface tension force at the top of the water column in a capillary tube. Capillary movement in the soil is the movement of soil moisture through the minute pores between the soil particles. The minute pores serve as the capillary tubes in saturated condition through which the soil moisture rises above the ground water table. For the capillary-rise in the soil to exist the voids should be completely and uninterruptedly filled with water. The capillary saturated zone between the GWT and the aerated zone above is called the closed capillary fringe. This contains no air and the thickness of this zone depends on the fineness of the particles. The larger will be the diameter of particle, smaller will be the rise of capillary and less will be the thickness of capillary fringe. Capillary water cannot be drained by normal drainage but can be controlled by controlling the ground water table. It can be drained when the actual amount of moisture is excess to that retained by surface tension.

● 2.5.1.2.1.2 Ground water zone

Below closed capillary zone, the main saturation zone lies, above the impermeable rock. It is the ground water zone experiencing a flow according to hydraulic gradient. This water if confined to the depressions in pockets of moraine clay located below ground surface in permeable sand is called as perched ground water. The amount of ground water accumulated depends on the season, rate of evaporation from the depression in the direction of ground surface and freezing (if occurs).



● 2.5.1.2.1.3 Aquifers

The body of highly permeable rock lying in the zone of saturation is the aquifer. An aquifer whose upper surface co-insides with the water table and therefore is in contact with the atmosphere is an unconfined aquifer where as the aquifer that is bounded by aquiclude (impermeable units) is the confined aquifer.

■ 2.5.1.3 Ground Water flow;

Ground water exists within the aquifers and the flow can be similar between flow in pipes under pressure and in fully saturated confined aquifers. The flow of the ground water with a free surface is also similar to that in an open channel. A major difference is the geometry of a ground water system flow channel as compared with common hydraulic pipe flow or channel system. The problem can easily be recognized by envisioning a discharging cross section composed of a number of small

openings, each with its own geometry, orientation, and size so that flow velocity from each pore varies in both magnitude and direction.

Darcy's Law for explaining the flow of fluid through a permeable bed is stated as

$$Q = -KA(dh/dx)$$

A = Total Cross Sectional Area,

K = Hydraulic Conductivity of the material

Q = flow across the control area A

dh/dx = Hydraulic gradient

Darcy's law is applicable to a number of application in hydrology particularly at Renold number 1 or less. Deviation from Darcy's law is experienced from the Renold number as low as 2, depending on grain size and shape and the density of fluid.

■ 2.5.2 Wet Land

India has a wide variety of wetlands like marshes, open water bodies, mangroves, swamps and salt marshes which are essential for environmental balance.

They help to retain water during dry season, thus helping in maintaining the water table. During floods, they act to reduce flood levels and to trap suspended solids and nutrients in the lakes. The wetlands have suffered alarming decline because of the rapidly expanding human population, large scale changes in land use and land cover and implementation of development projects. Wetland may broadly be categorized into inland, mangrove and forest wetland

■ 2.5.2.1 Inland Wetlands (National Sites)

The wetlands occurring in the land area with out any opening to sea is called the inland wetland. These wetlands are mainly riverine and associated to flood plain and other topographic depression. List of wetlands identified under national wetland conservation programme (After Ministry of Environment & Forests, 1990) is given below -

Sr. No.	Name of Wetlands	State/UT	Sr. No.	Name of Wetlands	State/UT
1	Wullar	Jammu and Kashmir	35	Rasik Beel	W. Bengal
2	Tso Morari	Jammu and Kashmir	36	Nawabganj	Uttar Pradesh
3	Tisgul Tso	Jammu and Kashmir	37	Sandi	Uttar Pradesh
4	Renuka	Himachal Pradesh	38	Lakh Bahoshi	Uttar Pradesh
5	Pong Dam	Himachal Pradesh	39	Samaspur	Uttar Pradesh
6	Chandratal	Himachal Pradesh	40	Suitanpur	Haryana
7	Harike	Punjab	41	Bhindawas	Haryana
8	Ropar	Punjab	42	Magadhi	Karnataka
9	Kanjli	Punjab	43	Gudavi Bird Sanctuary	Karnataka
10	Chilka	Orissa	44	Bonal	Karnataka
11	Kabar	Bihar	45	Hidkal and Ghataprabha	Karnataka
12	Sambhar	Rajasthan	46	Kaliveli	Tamilnadu
13	Kolleru	Andhra Pradesh	47	Pallikarni	Tamilnadu
14	Loktak	Manipur	48	Great Rann of Kachh	Gujarat
15	Ashtamudi	Kerala	49	Thol Bird Sanctuary	Gujarat
16	Sasthamkotta	Kerala	50	Khijadiya Bird Sanctuary	Gujarat
17	Ujani	Maharashtra	51	Little Rann of Kachh	Gujarat
18	Nalsarovar	Gujarat	52	Pariej	Gujarat
19	Deepar Beel	Assam	53	Wadhvana	Gujarat
20	Rudrasagar	tripura	54	Nanikakrad	Gujarat
21	Hokersar	Jammu and Kashmir	55	Barna	Madhya Pradesh
22	Mansar-Surinsar	Jammu and Kashmir	56	Yashwant Sagar	Madhya Pradesh
23	Pangong Tsar	J and K (Ladakh)	57	Wetland of Ken River	Madhya Pradesh
24	East Calcutta	West Bengal	58	National Chambal Sanctuary	Madhya Pradesh
25	Sunderbans	West Bengal	59	Ghatigaon	Madhya Pradesh
26	Point Calimer	Tamil Nadu	60	Ratapani	Madhya Pradesh
27	Kottuli	Kerala	61	Denwa Tawa wetland	Madhya Pradesh
28	Palak	Mizoram	62	Kanha Tiger Reserve	Madhya Pradesh
29	Tamdil	Mizoram	63	Pench Tiger Reserve	Madhya Pradesh
30	Barilla	Bihar	64	Sakhyasagar	Madhya Pradesh
31	Kusheshwar Asthan	Bihar	65	Dihaila	Madhya Pradesh
32	Ban Ganga Jhilmil Tal	Uttaranchal	66	Ranjitsagar	Jammu and Kashmir
33	Rewalsar	Himachal Pradesh	67	Govindgarh	Madhya Pradesh
34	Ahiron Beel	W. Bengal	68	Udhwa	Jharkhand

■ 2.5.2.2 Forest Wetland

Fresh water Wetlands - Myristica Swamps of Western Ghats

The Western Ghats is among the thirty-four global biodiversity hotspots of the world. From among the incredible diversity of ecosystems, there is rare fresh water wetlands found in Western Ghats with unique assemblage of floral and faunal biodiversity. Popularly known as Myristica swamps, these freshwater swamp forest ecosystem is confined to low altitudes characterized by slow flowing streams of Western Ghats river systems. Trees belonging to a primitive family of angiosperms; "Myristicaceae" are the dominant tree species of this unique wetland ecosystem. Most of Myristicaceae family members are capable of producing aerial roots when faced with unfavorable conditions. Hence the Myristica species found in these exceptional wetlands have adaptations such as knee roots for the anaerobic condition of the swamps, and stilt root for anchoring the tree species in damp soil. Other than Myristicaceae family species members of other families such as Celastraceae, Diptereocarpaceae, Anacardiaceae, Xanthophyllaceae and others are significant part of this swamp community. Ground vegetation is not dense and the dominant ground vegetation is belonging to Araceae, Zingiberace and Acanthaceae families. Some of these swamps are considered even as sacred groves and have thus ecoheritage value. These swamps having variety of microhabitats which provides favorable conditions for survival and procreation of many annelids, arthropods, molluscs, fishes, amphibians, reptiles, birds and mammals. Many of these animals found in the Myristica swamps are endemic and some are on the red-list of IUCN. More importantly this wetlands also play a critical role in water storing and maintaining ground water level. Myristica swamps are critically endangered ecosystem in Western Ghats.(Roby and Nair 2006) About 53 patches of Myristica swamps have been recorded from the Kulathupuzha, Anchal forest ranges and Shendurny Wildlife Sanctuary of the Kollam and Thiruvananthapuram districts of Southern Kerala. A total of 65 tree species and 72 species of shrub-herb combine were recorded from the Myristica swamps. Two species of earthworms, three species of crabs, ten species of fishes, thirty four species of amphibians, thirty three species of reptiles, fifty eight species of birds and twenty one species of mammals have been recorded as present in the swamps. In addition to this annelids and arthropods are also being recorded. (Source: Kerala Forest Research Institute (KFRI)).

■ 2.5.2.3 Wetland Conservation

The UNCED conference on environment and development as well as global conservation organizations, including RAMSAR convention, have identified the aquatic biodiversity to be the most threatened of all biodiversities. There is, therefore, an urgent and felt need to conserve the aquatic biodiversity including

the ecosystem processes. Wetlands cover an area of about 5.5 million ha, of which 1.5 million ha enjoy complete protection and 1.6 million ha partial protection. India being a developing country supporting the second largest population in the world, having mainly agrarian economy, has a significant impact on all natural resources including that of wetlands. It is in this context, an inventory of the important wetlands is essential. The directory of Asian Wetlands (1989) lists 93 Wetlands of International importance in India. Information on the type and extent of wetlands is lacking in India. Hence, this information is a baseline requirement for forming protected area network and conservation.

However, so far in India there has been no systematic attempt to evolve conservation preserves analogous to terrestrial protected areas. While many developing countries such as China and south-east Asian countries have progressed substantially in formulating an action plan for conserving aquatic biodiversity, similar task of identifying such reserves for India is lacking. While a country like UK could designate 161 Ramsar sites, obviously India being a mega diverse country will have many more than the 20 sites presently identified by the national wetland programme.

▲ 2.5.2.3.1 Issues

There are obviously a number of issues involved in evolving wetland conservation and monitoring programme. There are three major components. The first involves use of spatial technology tools viz, primarily use of remote sensing data to obtain actual location of wetland, inventory and classification of spatial data. The second issue relates to use of information on species distribution of birds, plants and socioeconomic issues related to the wetlands. The final issue involves integration of the data and information from the above two.

▲ 2.5.2.3.2 Inventory

The Ministry of Environment and Forests, Govt. of India, conducted the first nation wide survey of wetlands of India in 1980's. This was done by a survey whose results were summarized in Wetlands of India- A Directory (Anon, 1990) while this inventory is a good beginning; it is obvious that more was needed for conservation work. SAC and MoE&F have conducted an inventory of indian wetlands on the basis of Visual interpretation of Satellite Data. They have mapped wetlands which are 50 ha or above in size. These maps convey the spatial extent of the wetlands but do not give any information about the exact nature of Fauna and Flora and water quality. Thus these maps are not sufficient to assign conservation values to the wetlands. Various smaller wetlands required mapping due to their biological importance. Thus maps of larger scale and greater information regarding the biological communities sustained by these wetlands are essential for classification and evolving long-term conservation strategies.

▲ 2.5.2.3.3 Classification

The only serious attempts made to classify Indian wetlands on biological basis were done in late 1970's and early 1980's based on herbaceous and forested vegetations. Gopal and Krishnamurthy (1993) also suggested the classification scheme for Wetlands of South Asia based on vegetation types. But none of the above classifications took in to consideration other biological parameters like faunistic assemblages. Wetlands were also not taken into consideration while planning the protected area network for wildlife in India (Rodgers and Panwar, 1988), which was based on vegetation and land area. Thus there is a strong need to classify wetlands according to their biological components for conservation needs. A classification system has been proposed in the ongoing Inland wetland project for classification of wetlands according to their types and levels of turbidity and aquatic vegetation. This classification proved effective for classification of wetlands using satellite remote sensing data.

▲ 2.5.2.3.4 Integrated approach

Conservationists and all stakeholders in wetland sustainability believe that wetland conservation strategy requires an extensive participatory programme. The ongoing wetland mapping and conservation programme under MoEn&F, UNDP and SACON aims at developing such a user-friendly effective protected area network of wetlands in the country. Understandably, this is a project of very large magnitude and requires an enormous manpower and inter-institutional collaboration. A large number of individuals from NGO's, academic institutions and professionals have come together to make this possible. From amongst the various participant institutions and individuals, state level coordinators have been identified, who will coordinate the fieldwork in their respective states. The field data enables prioritization of a set of wetlands from each state and constitute the countrywide network of inland wetlands.

Use of existing spatial information

The maps produced by SAC, Ahmedabad forms a baseline data at a reconnaissance level. Thus the inputs regarding area of wetlands-Pre/Post Monsoon, presence or absence of aquatic vegetation and qualitative information on turbidity levels are welcome additions to the MoEn&F questionnaire survey. Database of irrigation department and government agencies involved in major and minor irrigation and fisheries will also be a valuable input in such studies. Efforts should be made to integrate such data as inputs to management and conservation plans.

Use of existing information of species distribution

The database of the past species distribution of birds and animals of the country done by Asian Waterfowl Count and other organizations should be integrated and their spatial distribution plotted for proper understanding of the temporal distribution of the species.

▲ 2.5.2.3.5 Socioeconomic Issues

Socio-economic values of wetlands for prioritisation of biological conservation sites are a very important aspect. The location of these wetlands, their nearest town or city, its biogeographic region, its present legal conservation status, and major use of the wetland to the local population are important inputs for conservation values.

▲ 2.5.2.3.6 Technological Advancements

The Inland Wetland inventory for the entire country requires to be carried out by making use of available data and also by fresh data generation using data on remote sensing. Remote sensing data has proved to be economic means for inventory, monitoring and management of wetlands. Satellite Data in association with geographical information system provide cost and time effective tool for mapping and formulation of conservation and management plans. Using Remote sensing data the spatial mapping is done and biologically important parameters are attached to it using GIS. Thus, by using digital remote sensing data it will be possible to map, analyse data and attain information at any scale for all wetlands, according to the management and conservation requirements. For reconnaissance survey satellite data of 20m resolution is very effective while for larger scale studies use of planned Cartosat data will be a better option.

In order to test this approach and formulate a viable programme of wetland inventors, a nationwide project has been initiated. This includes digital image processing of satellite data to identify the type of wetlands, turbidity levels and presence and absence of aquatic vegetation in various districts through out the country. The database of the work carried out by SAC has been used to create a digital inventory of wetlands using GIS. Using this database of centroid locations of wetlands, district and state wise wetland coverages were generated. Associated attribute data were also attached to these coverage. Keeping in view the gigantic amount of data customized graphical user interfaces (GUI) were developed to automate the process of coverage generation. These coverage and inputs of local individuals and institutions were used to select wetlands for fieldwork for their biological and conservational values. Database of the distribution of fauna and floral assemblages of the selected wetlands will also be converted into coverage for spatial depiction and related attributes attached to them using GIS. These inputs will be attached to the

digitally processed remote sensing data for further analysis. Thus, all the information will be readily available at all scales to assign conservation value to the wetlands. The importance of the wetlands will depend on whether it is a good, rare and unique wetland type, its socio economic and cultural value, importance of the threatened and endemic species of plants and animals found in these wetlands. They also depend on the importance of the wetland in maintaining the genetic or ecological diversity of the area, whether the site is of crucial value to certain plants or animal species at a critical stage of their life history and the importance of it regarding the resident and migrating waterfowl populations. Such an effort requires first hand insights into problems and prospects using the modern spatial technology tools.

In order to achieve viable wetland conservation in the country a great deal of effort has to go in systematizing the inventories at desired spatial and temporal scales. Substantial extent of synergy can be established between various institutions and organizations involved in natural resource management, conservation and organisation involved in research and development. A permanent wetland monitoring authority analogous to forest monitoring organisations should be set up.

▲ 2.5.2.3.7 Prioritization of Inland wetlands for conservation

Although as many as 16 parameters such as hydrology, birds, vegetation, fishes, and socioeconomic attributes were chosen for prioritization, criteria on birds, size of the wetland and socioeconomic were only considered for arriving at the conservation value assessment, as all other data were not available uniformly for most wetlands. Yet this exercise is again, first of its kind to be attempted at a national level. It must be added that there is ample scope to qualitatively and quantitatively improve upon this methodology and its application in any future attempt of this nature.

Major criteria identified for prioritizing the wetlands were birds as recognised by the Ramsar Convention Bureau (1971,1996). Size and economic use of the wetlands were also used. The specific criteria related to birds used were:

● 2.5.2.3.7.1 Criteria

- » total number of species,
- » total number of birds (abundance),
- » number of species with 1% biogeographic population,
- » number of threatened species,
- » number of threatened birds,
- » number of rare or near threatened species,
- » number of species of raptors

In all, 655 inland wetlands were prioritized. Ranks were given for 604 wetlands according to the availability of the above data. Of the 604 sites surveyed, 167 were identified as top-ranking sites followed

by 173 in the second rank and 264 in the third and fourth ranks together. Distribution of the prioritized wetlands is given in

Altogether 199 wetlands fulfilled more than one criteria to be qualified for declaration as Ramsar site (Table below). A few more wetlands which are currently included in the fourth rank are potential Ramsar sites. These sites, according to the AWC data are potential, but could not be assessed during year 2002, because of the drought. Distribution of the proposed Ramsar sites in each state is given in the following table

State	Total sites	Sites with bird count	1st rank	2nd rank	3rd & 4th ranks	Ramsar sites proposed
Tamil Nadu	49	49	5	10	34	5
Kerala	4	4	1	3	0	0
Karnataka	75	75	11	21	43	21
Goa	2	2	1	1	0	1
Andhra Pradesh	55	55	10	10	35	20
Orissa	7	7	4	1	2	4
Madhya Pradesh & Chhattisgarh	48	41	11	12	18	6
Maharashtra	33	33	11	9	13	11
Gujarat	45	45	16	10	19	21
Rajasthan	47	47	8	9	30	8
Haryana	4	4	4	0	0	4
Punjab	13	13	3	5	5	3
Jammu & Kashmir	29	22	9	11	2	7
Himachal Pradesh	4	3	1	1	1	0
Uttar Pradesh & Uttaranchal	50	50	14	17	19	25
Bihar & Jharkhand	42	6	3	3	0	3
West Bengal	25	25	7	8	10	13
Sikkim	4	4	2	2	0	2
Meghalaya	7	7	3	2	2	3
Assam	75	75	25	28	22	24
Arunachal Pradesh	18	18	8	3	7	8
Nagaland	3	3	2	1	0	3
Manipur	6	6	5	1	0	4
Mizoram	3	3	1	2	0	1
Tripura	7	7	2	3	2	2
Total	655	604	167	173	264	199

Action needs to be taken on priority to prepare detailed proposal for the 199 wetlands to be brought under Ramsar sites. The remaining 438 wetlands prioritized for conservation, may be recognized as wetlands of national and local importance and, the entire 655 wetlands, may be brought under a **National Network of Wetland Conservation Areas**. Those wetlands while have 'low' use for the local people living around may be declared as protected areas, while those with medium and high use may be declared as Community Reserves to be managed by stakeholders with the participation of wildlife experts and managers.(Vijayan V.S.et al 2004)

■ 2.5.2.4 Wetland Ecology

One of the long standing needs for conservation of the wetland biodiversity in the country is to comprehensively prioritize them. To begin with the process of prioritization, a tentative list of wetlands were identified by referring to the data on the wetlands from SAC (1998), from Asian Waterfowl count for the last 12 years (Anon 1988,1990,1994,1995,1996, Mundkur and Taylor 1993; Perennou *et al.* 1994; Lopez and Mundkur 1997 and from the BNHS and Wetland International). Thus a list of wetlands to be studied was prepared on the basis on earlier data and also based on the information of the local people. Thus the process of initial prioritization, is a participatory exercise.

Each team was to collect data on waterfowl, vegetation cover, socioeconomic use of the wetlands by the villagers around, and also collect fish for assessing contamination level. However, data on vegetation and fishes were collected only from some of the wetlands.

Endemic wetlands plants, Threatened fishes and Threatened turtles

Sl	State	Endemic plants	Threatened fishes	Threatened turtles
1	Tamil Nadu	46	35	4
2	Kerala	65	37	3
3	Karnataka	64	15	2
4	Goa	17	15	0
5	Andhra Pradesh	13	19	2
6	Orissa	6	22	6
7	Madhya Pradesh & Chhattisgarh	20	16	7

8	Maharashtra	69	19	3
9	Gujarat	11	12	1
10	Rajasthan	9	17	3
11	Haryana	0	15	1
12	Punjab	0	15	5
13	Jammu & Kashmir	2	20	2
14	Himachal Pradesh	0	19	0
15	Uttar Pradesh	3	34	10
16	Bihar & Jharkhand	5	19	9
17	West Bengal	5	32	11
18	Assam	2	35	10
19	North-Eastern States*	5	34	9
20	Total number of species	114	102	16

* Arunachal Pradesh, Meghalaya, Nagaland, Manipur, Mizoram, Sikkim, Tripura

Threatened birds observed in the major wetlands of India

Sl	Common Name	Scientific Name	No. of birds	No. of sites (n=591)	No. of states (n=25)
1	Baer's Pochard	<i>Aythya baeri</i>	378	7	4
2	Baikal Teal	<i>Anas formosa</i>	35	1	1
3	Blacknecked Crane	<i>Grus nigricollis</i>	*	5	3
4	Bristled Grass Warbler	<i>Chaetornis striatus</i>	11	2	2
5	Dalmatian Pelican	<i>Pelecanus crispus</i>	396	20	4
6	Finn's Weaver	<i>Ploceus megarhynchus</i>	1	1	1
7	Greater Adjutant	<i>Leptoptilos dubius</i>	95	2	2
8	Greater Spotted Eagle	<i>Aquila clanga</i>	173	36	8
9	Imperial Eagle	<i>Aquila heliaca</i>	3	2	2
10	Indian Skimmer	<i>Rynchops albicollis</i>	2	1	1

11	Longbilled Vulture	<i>Gyps indicus</i>	228	46	3
12	Lesser Adjutant	<i>Leptoptilos javanicus</i>	298	4	8
13	Lesser Kestrel	<i>Falco naumanni</i>	2	3	2
14	Marbled Teal	<i>Marmaronetta angustirostris</i>	69	4	2
15	Masked Finfoot	<i>Helipais personata</i>	8	1	1
16	Pallas's Fishing Eagle	<i>Haliaeetus leucoryphus</i>	70	23	7
17	Sarus Crane	<i>Grus antigone</i>	1,812	76	7
18	Sociable Lapwing	<i>Vanellus gregarius</i>	5	3	1
19	Spoonbilled Sandpiper	<i>Eurynorhynchus pygmeus</i>	14	2	2
20	Spotbilled Pelican	<i>Pelecanus philippensis</i>	1,485	59	8
21	Spotted Greenshank	<i>Tringa guttifer</i>	76	8	6
22	Swamp Francolin	<i>Francolinus gularis</i>	25	1	1
23	Whitebellied Heron	<i>Ardea insignis</i>	*	1	1
24	Whitebacked Vulture	<i>Gypus bengalensis</i>	451	38	10
25	Whitewinged Wood Duck	<i>Cairina scutulata</i>	*	7	3
26	Wood Snipe	<i>Gallinago nemoricola</i>	*	1	1
Total			5,637		

* Number not available; Status of species based on BirdLife International 2001

■ 2.5.3 Watersheds

A watershed is an area of land within which all precipitation drains into a single major body of water. Put another way, it is the geographic area that feeds a stream, creek, wetland or aquifer. Small watersheds feed into larger watersheds, which feed into even larger river basins, which eventually feed into oceans and seas.

Each watershed is very unique, and can supply and retain water in very different ways. Some watersheds may experience times of drought, while nearby watersheds remain fertile and healthy. Much of this depends upon aquifers, the natural underground water storage tanks which are cyclically recharged by rainfall and snow. Aquifers affect soil moisture and support all life living within the watershed. **Watershed** may be defined as a **drainage divide**, **water divide**, or simply **divide** is the separation between neighbouring drainage basins (catchments). In hilly country, the divide lies along topographical peaks and ridges, but in flat country or on a high plateau (especially where the ground is marshy) the

divide may be invisible – just a more or less notional line on the ground on either side of which falling raindrops will start a journey to different rivers, and even to different sides of a country or continent. Drainage divides are important geographical, and often also political, boundaries. Roads (such as ridgeways) and rail tracks often follow divides to minimise grades (gradients) and the need for bridges. In British Isles usage, a **watershed**, the line between drainage basins – *shedding* is an old term for splitting or dividing, so it is the line which divides the water, however in North America **watershed** is used instead for the drainage basin itself (Wikipedia, 2008).

The committee on runoff of the American Geophysical Union treats the micro unit within a river basin as the Watershed which is defined as those having within their perimeter a long overland flow rather than that of a surface channel, so that, hydrologically there is no surface storage effect in the watershed and the effect of overland flow distinctly dominates in the total runoff. Hence the watershed area remains sensitive to the high sensitive storms rainfalls of short duration as well as to the land management systems (Pal, S K.,).

■ 2.5.3.1 Importance of drainage basins

▲ 2.5.3.1.1 As geopolitical boundaries

Drainage basins have been important historically in determining boundaries, particularly in regions where trade by water has been important. For example, the English crown gave the Hudson's Bay Company a monopoly on the Indian Trade in the entire Hudson Bay watershed, an area called Rupert's Land. The company later acquired the North American watershed of the Arctic Ocean (the North-Western Territory). These lands later became part of Canada as the Northwest Territories, making up the vast majority of Canada's land area. Today, bioregional democracy can include agreements of states in a particular drainage basin to defend it. These include the Great Lakes Commission.

▲ 2.5.3.1.2 In hydrology

In hydrology, the drainage basin is a logical unit of focus for studying the movement of water within the hydrological cycle, because the majority of water that discharges from the basin outlet originated as precipitation falling on the basin. A portion of the water that enters the groundwater system beneath the drainage basin may flow towards the outlet of another drainage basin because groundwater flow directions do not always match those of their overlying drainage network. Measurement of the discharge of water from a basin may be made by a stream gauge located at the basin's outlet.

Rain gauge data is used to measure total precipitation over a drainage basin, and there are different ways to interpret that data. If the gauges are many and evenly distributed over an area of uniform precipitation, using the arithmetic mean method will give good results. In the Thiessen polygon method, the watershed is divided into polygons with the rain gauge in the middle of each polygon assumed to be representative for the rainfall on the area of land included in its polygon. These polygons are made by drawing lines between gauges, then making perpendicular bisectors of those lines form the polygons. The isohyetal method involves contours of equal precipitation are drawn over the gauges on a map. Calculating the area between these curves and adding up the volume of water is time consuming.

▲ 2.5.3.1.3 In ecology

Drainage basins are important elements to consider also in ecology. As water flows over the ground and along rivers it can pick up nutrients, sediment, and pollutants. Like the water, they get transported towards the outlet of the basin, and can affect the ecological processes along the way as well as in the receiving water source.

Modern usage of artificial fertilizers, containing nitrogen, phosphorus, and potassium, has affected the mouths of watersheds. The minerals will be carried by the watershed to the mouth and accumulate there, disturbing the natural mineral balance.

▲ 2.5.3.1.4 In resource management

Because drainage basins are coherent entities in a hydrological sense, it has become common to manage water resources on the basis of individual basins. In the U.S. state of Minnesota, governmental entities that perform this function are called watershed districts. In New Zealand, they are called catchment boards. Comparable community groups based in Ontario, Canada, are called conservation authorities. In North America this function is referred to as watershed management. In Brazil, the National Policy of Water Resources, regulated by Act n° 9.433 of 1997, establishes the drainage basin as territorial division of Brazilian water management.

■ 2.5.3.2 Major watersheds in India

The name, the drainage area, annual runoff and ground water potential has been noted below. (Pal, S.K.)

Sl.No.	Watersheds	Catchment Area km ²	Annual Runoff In m ha	Ground water	
				Potential In m ha	Utilization In m ha
1.	Ganga	1098600	55.01	18.50	10.639
2.	Brahmaputra	187100	42.20	0.88	1.802
3.	Barak and others	78,200	17.50	0.37	0.66
4.	Godavari	312800	11.54	8.33	1.91
5.	Indus	321000	7.69	4.93	1.118
6.	Mahanadi	141600	7.07	6.64	1.06
7.	Krishna	259000	5.78	5.28	1.76
8.	Subarnarekha, baitrani and Brahmani	81000	4.35	4.09	0.651
9.	Narmada	98800	4.01	2.96	0.710
10.	Pennar and others(Between Krishna & Kaveri)	14490	2.53	1.79	0.998
11	Tapi	66900	1.97	1.46	0.406
12	Kaveri	87900	1.86	2.06	0.49
13	East flowing rivers (Btm Mahanadi & Godavari)	49700	1.72	1.61	0.405
14	Luni and other rivers of Sourashtra	321900	1.23	0.19	2.019
15.	3 Mahi including Dhadar	37600	1.18	0.87	0.253
16.	Tanbra pami	35100	0.90	1.17	0.11
17.	Sabarmati	21700	0.37	0.27	0.065

(Source; Pal.S. K.)

■ 2.5.4 Self Assessment Questions

1. What do you mean by capillary? Define closed capillary zone.
2. How do ground water move and what are the factors for ground water movement? Discuss with the help of darcy's law.
3. Define watershed. Discuss the importances of watershed with special reference to hydrological values?
4. Discuss the ecological importance of wetland.
5. What do you mean by prioritisation of wetland and what should be the criteria for prioritisation?
6. What is the role of RS & GIS for conserving wetland?
7. What are the steps to be taken for conservation of wetland?

■ 2.5.5 References

- Anon. 1988. Asian waterfowl census. IWRB, AWB.
- Anon. 1990. A Directory of wetlands in India. Ministry of Environment, Govt. of India, New Delhi.
- Anon. 1994. Republic of the Philippines - Conservation priority protected area project. World Bank/GEF Washington D.C.
- Anon. 1995. Integrated study for sustainable development of Kasaragode district of Kerala. Kerala State Land use Board, Thiruvananthapuram.
- Birdlife International. 2001. Threatened birds of Asia. BirdLife International, Cambridge.
- Environmental Information System run website <http://www.wetlandofindia.org>
- Murck, Skinner and Porter (1997) Environmental Engineering, pp.377-400
- Lopez A. and Mundkur T. 1997. The Asian waterfowl census (1994-1996): Results of the coordinated waterbird census and an overview of the status of the wetlands in Asia. Wetlands International.
- Mundkur T. and Taylor. 1993. Asian waterfowl census. 1993. AWB, Kuala Lumpur, Malaysia and IWRB, Slimbridge, UK.
- Pal, S.K. Physical geography of India pp.65-77
- Perennou C., Mundkur T., Scott D., Follested and Kvenild A.L. 1994. The Asian waterfowl census 1987-91: Distribution and status of Asian waterfowl. AWB Publication no.86/IWRB Special Publication No. 24. Asian Wetland Bureau, Kuala Lumpur, and International Waterfowl and Wetlands Research Bureau, Slimbridge.
- SAC (Space Application Centre). 1998. Wetlands of India. Project Report: RSAM/SAC/RESA/PR/01/98.
- Vijayan V.S., Prasad S.N., Vijayan L. and Muralidharan S. 2004. Inland Wetlands of India-Conservation Priorities. Salim Ali Centre for Ornithology and Natural History, Coimbatore. pp. xxiv + 532

“Learner’s Feed-back”

After going through the Modules/ Units please answer the following questionnaire.

Cut the portion and send the same to the Directorate.

To
The Director
Directorate of Distance Education,
Vidyasagar University,
Midnapore- 721 102.

1. The modules are : (give ✓ in appropriate box)
☐ Easily understandable; ☐ very hard; ☐ partially understandable.
2. Write the number of the Modules/Units which are very difficult to understand :
.....
.....
.....
3. Write the number of Modules/ Units which according to you should be re-written :
.....
.....
.....
4. Which portion/page is not understandable to you? (Mention the page no. And Portion)
.....
.....
.....
5. Write a short comment about the study material as a learner.
.....
.....
.....
.....
.....

.....
(full Signature of the Learner)

Date :

Enrolment No.

Phone/Mobile No.



Published by : Directorate of Distance Education,
Vidyasagar University, Midnapore 721 102, W.B.
Printed by : S P Communications Pvt. Ltd., 31B, Raja Dinendra Street, Kolkata - 700 009
E-mail : spcomkol@gmail.com