Geography

Models and Measurement

Ritika Gautam

Geography: Models and Measurement

Geography: Models and Measurement

Ritika Gautam

VIDYA BOOKS

Published by Vidya Books,
305, Ajit Bhawan,
21 Ansari Road,
Daryaganj, Delhi 110002

Ritika Gautam
ISBN: 978-93-5429-598-0

Contents

1

Introduction

Traditionally, geographers have been viewed the same way as cartographers and people who study place names and numbers. Although many geographers are trained in toponymy and cartology, this is not their main preoccupation. Geographers study the spatial and temporal distribution of phenomena, processes and features as well as the interaction of humans and their environment. As space and place affect a variety of topics such as economics, health, climate, plants and animals; geography is highly interdisciplinary.

> *"...mere names of places...are not geography...know by heart a whole gazetteer full of them would not, in itself, constitute anyone a geographer. Geography has higher aims than this: it seeks to classify phenomena (alike of the natural and of the political world, in so far as it treats of the latter), to compare, to generalize, to ascend from effects to causes, and, in doing so, to trace out the laws of nature and to mark their influences upon man. This is 'a description of the world'—that is Geography. In a word Geography is a Science—a thing not of mere names but of argument and reason, of cause and effect."*— William Hughes, 1863.

Geography as a discipline can be split broadly into two main subsidiary fields: human geography and physical geography. The former largely focuses on the built environment and how humans create, view, manage, and influence space. The latter examines the natural environment and how organisms, climate, soil, water, and landforms produce and interact. The difference between these approaches led to a third field, environmental geography, which combines physical and human geography and looks at the interactions between the environment and humans.

Branches of Geography

Physical Geography

Physical geography (also known as geosystems or physiography) is one of the

two major subfields of geography. Physical geography is that branch of natural science which deals with the study of processes and patterns in the natural environment like the atmosphere, biosphere and geosphere, as opposed to the cultural or built environment, the domain of human geography.

Within the body of physical geography, the Earth is often split either into several spheres or environments, the main spheres being the atmosphere, biosphere, cryosphere, geosphere, hydrosphere, lithosphere and pedosphere. Research in physical geography is often interdisciplinary and uses the systems approach.

Sub-fields

Physical Geography can be divided into several sub-fields, as follows:

Geomorphology

Geomorphology is the scientific study of landforms and the processes that shape them. Geomorphologists seek to understand why landscapes look the way they do, to understand landform history and dynamics, and to predict future changes through a combination of field observations, physical experiments, and numerical modelling. Geomorphology is practiced within geography, geology, geodesy, engineering geology, archaeology, and geotechnical engineering, and this broad base of interest contributes to a wide variety of research styles and interests within the field.

Overview

The surface of Earth is modified by a combination of surface processes (that sculpt landscapes) and geologic processes (that cause tectonic uplift and subsidence). Surface processes comprise the action of water, wind, ice, fire, and living things on the surface of the Earth, along with chemical reactions that form soils and alter material properties, the stability and rate of change of topography under the force of gravity, and other factors, such as (in the very recent past) human alteration of the landscape. Many of these factors are strongly mediated by climate. Geologic processes include the uplift of mountain ranges, the growth of volcanoes, isostatic changes in land surface elevation (sometimes in response to surface processes), and the formation of deep sedimentary basins where the surface of Earth drops and is filled with material eroded from other parts of the landscape. The Earth surface and its topography therefore are an intersection of climatic, hydrologic, and biologic action with geologic processes.

The broad-scale topographies of Earth illustrate this intersection of surface and subsurface action. Mountain belts are uplifted due to geologic processes. Denudation of these high uplifted regions produces sediment that is transported and deposited elsewhere within the landscape or off the coast. On progressively smaller scales, similar ideas apply, where individual landforms evolve in response to the balance of additive processes (uplift and deposition) and subtractive processes (subsidence and erosion). Often, these processes directly affect each other: ice sheets, water, and

sediment are all loads that change topography through flexural isostasy. Topography can modify the local climate, for example through orographic precipitation, which in turn modifies the topography by changing the hydrologic regime in which it evolves. Many geomorphologists are particularly interested in the potential for feedbacks between climate and tectonics mediated by geomorphic processes.

In addition to these broad-scale questions, geomorphologists address issues that are more specific and/or more local. Glacial geomorphologists investigate glacial deposits such as moraines, eskers, and proglacial lakes, as well as glacial erosional features, to build chronologies of both small glaciers and large ice sheets and understand their motions and effects upon the landscape. Fluvial geomorphologists focus on rivers, how they transport sediment, migrate across the landscape, cut into bedrock, respond to environmental and tectonic changes, and interact with humans. Soils geomorphologists investigate soil profiles and chemistry to learn about the history of a particular landscape and understand how climate, biota, and rock interact. Other geomorphologists study how hillslopes form and change. Still others investigate the relationships between ecology and geomorphology. Because geomorphology is defined to comprise everything related to the surface of Earth and its modification, it is a broad field with many facets.

Practical applications of geomorphology include hazard assessment (such as landslide prediction and mitigation), river control and stream restoration, and coastal protection.

History

With some notable exceptions, geomorphology is a relatively young science, growing along with interest in other aspects of the Earth Sciences in the mid 19th century. This section provides a very brief outline of some of the major figures and events in its development.

Ancient Geomorphology

Perhaps the earliest one to devise a theory of geomorphology was the polymath Chinese scientist and statesman Shen Kuo (1031-1095 AD). This was based on his observation of marine fossil shells in a geological stratum of a mountain hundreds of miles from the Pacific Ocean. Noticing bivalve shells running in a horizontal span along the cut section of a cliffside, he theorized that the cliff was once the pre-historic location of a seashore that had shifted hundreds of miles over the centuries. He inferred that the land was reshaped and formed by soil erosion of the mountains and by deposition of silt, after observing strange natural erosions of the Taihang Mountains and the Yandang Mountain near Wenzhou. Furthermore, he promoted the theory of gradual climate change over centuries of time once ancient petrified bamboos were found to be preserved underground in the dry, northern climate zone of *Yanzhou*, which is now modern day Yan'an, Shaanxi province.

Early Modern Geomorphology

The first use of the word geomorphology was likely to be in the German language when it appeared in Laumann's 1858 work. Keith Tinkler has suggested that the word came into general use in English, German and French after John Wesley Powell and W. J. McGee used it in the International Geological Conference of 1891.

An early popular geomorphic model was the *geographical cycle* or the *cycle of erosion*, developed by William Morris Davis between 1884 and 1899. The cycle was inspired by theories of uniformitarianism first formulated by James Hutton (1726–1797). Concerning valley forms, uniformitarianism depicted the cycle as a sequence in which a river cuts a valley more and more deeply, but then erosion of side valleys eventually flatten the terrain again, to a lower elevation. Tectonic uplift could start the cycle over. Many studies in geomorphology in the decades following Davis' development of his theories sought to fit their ideas into this framework for broad scale landscape evolution, and are often today termed "Davisian". Davis' ideas have largely been superseded today, mainly due to their lack of predictive power and qualitative nature, but he remains an extremely important figure in the history of the subject.

In the 1920s, Walther Penck developed an alternative model to Davis', believing that landform evolution was better described as a balance between ongoing processes of uplift and denudation, rather than Davis' single uplift followed by decay. However, due to his relatively young death, disputes with Davis and a lack of English translation of his work his ideas were not widely recognised for many years.

These authors were both attempting to place the study of the evolution of the Earth's surface on a more generalized, globally relevant footing than had existed before. In the earlier parts of the 19th century, authors-especially in Europe-had tended to attribute the form of landscape to local climate, and in particular to the specific effects of glaciation and periglacial processes. In contrast, both Davis and Penck were seeking to emphasize the importance of evolution of landscapes through time and the generality of Earth surface processes across different landscapes under different conditions.

Quantitative Geomorphology

While Penck and Davis and their followers were writing and studying primarily in Western Europe, another, largely separate, school of geomorphology was developed in the United States in the middle years of the 20th century. Following the early trailblazing work of Grove Karl Gilbert around the turn of the 20th century, a group of natural scientists, geologists and hydraulic engineers including Ralph Alger Bagnold, John Hack, Luna Leopold, Thomas Maddock and Arthur Strahler began to research the form of landscape elements such as rivers and hillslopes by taking systematic, direct, quantitative measurements of aspects of them and investigating the scaling

of these measurements. These methods began to allow prediction of the past and future behaviour of landscapes from present observations, and were later to develop into what the modern trend of a highly quantitative approach to geomorphic problems. Quantitative geomorphology can involve fluid dynamics and solid mechanics, geomorphometry, laboratory studies, field measurements, theoretical work, and full landscape evolution modelling. These approaches are used to understand weathering and the formation of soils, sediment transport, landscape change, and the interactions between climate, tectonics, erosion, and deposition.

Contemporary Geomorphology

Today, the field of geomorphology encompasses a very wide range of different approaches and interests. Modern researchers aim to draw out quantitative "laws" that govern Earth surface processes, but equally, recognize the uniqueness of each landscape and environment in which these processes operate. Particularly important realizations in contemporary geomorphology include:

1) that not all landscapes can be considered as either "stable" or "perturbed", where this perturbed state is a temporary displacement away from some ideal target form. Instead, dynamic changes of the landscape are now seen as an essential part of their nature.
2) that many geomorphic systems are best understood in terms of the stochasticity of the processes occurring in them, that is, the probability distributions of event magnitudes and return times. This in turn has indicated the importance of chaotic determinism to landscapes, and that landscape properties are best considered statistically. The same processes in the same landscapes does not always lead to the same end results.

Processes

Modern geomorphology focuses on the quantitative analysis of interconnected processes. Modern advances in geochronology, in particular cosmogenic radionuclide dating, optically stimulated luminescence dating and low-temperature thermochronology have enabled us for the first time to measure the rates at which geomorphic processes occur. At the same time, the use of more precise physical measurement techniques, including differential GPS, remotely sensed digital terrain models and laser scanning techniques, have allowed quantification and study of these processes as they happen. Computer simulation and modelling may then be used to test our understanding of how these processes work together and through time.

Geomorphically relevant processes generally fall into (1) the production of regolith by weathering and erosion, the transport of that material, and its eventual deposition. Although there is a general movement of material from uplands to lowlands, erosion, transport, and deposition often occur in closely-spaced tandem all across the landscape. The nature of the processes investigated by geomorphologists is strongly dependent

on the landscape or landform under investigation and the time and length scales of interest. However, the following non-exhaustive list provides a flavour of the landscape elements associated with some of these.

Primary surface processes responsible for most topographic features include wind, waves, chemical dissolution, mass wasting, groundwater movement, surface water flow, glacial action, tectonism, and volcanism. Other more exotic geomorphic processes might include periglacial (freeze-thaw) processes, salt-mediated action, or extraterrestrial impact.

Fluvial Processes

Rivers and streams are not only conduits of water, but also of sediment. The water, as it flows over the channel bed, is able to mobilize sediment and transport it downstream, either as bed load, suspended load or dissolved load. The rate of sediment transport depends on the availability of sediment itself and on the river's discharge.

Rivers are also capable of eroding into rock and creating new sediment, both from their own beds and also by coupling to the surrounding hillslopes. In this way, rivers are thought of as setting the base level for large scale landscape evolution in nonglacial environments. Rivers are key links in the connectivity of different landscape elements.

As rivers flow across the landscape, they generally increase in size, merging with other rivers. The network of rivers thus formed is a drainage system and is often dendritic, but may adopt other patterns depending on the regional topography and underlying geology.

Aeolian Processes

Aeolian processes pertain to the activity of the winds and more specifically, to the winds' ability to shape the surface of the Earth. Winds may erode, transport, and deposit materials, and are effective agents in regions with sparse vegetation and a large supply of unconsolidated sediments. Although water and mass flow tend to mobilize more material than wind in most environments, aeolian processes are important in arid environments such as deserts.

Soil, regolith, and rock move downslope under the force of gravity via creep, slides, flows, topples, and falls. Such mass wasting occurs on both terrestrial and submarine slopes, and has been observed on Earth, Mars, Venus, Titan and Iapetus.

Ongoing hillslope processes can change the topology of the hillslope surface, which in turn can change the rates of those processes. Hillslopes that steepen up to certain critical thresholds are capable of shedding extremely large volumes of material very quickly, making hillslope processes an extremely important element of landscapes in tectonically active areas. On Earth, biological processes such as burrowing or tree throw may play important roles in setting the rates of some hillslope processes.

Glacial Processes

Glaciers, while geographically restricted, are effective agents of landscape change. The gradual movement of ice down a valley causes abrasion and plucking of the underlying rock. Abrasion produces fine sediment, termed glacial flour. The debris transported by the glacier, when the glacier recedes, is termed a moraine. Glacial erosion is responsible for U-shaped valleys, as opposed to the V-shaped valleys of fluvial origin.

The way glacial processes interact with other landscape elements, particularly hillslope and fluvial processes, is an important aspect of Plio-Pleistocene landscape evolution and its sedimentary record in many high mountain environments. Environments that have been relatively recently glaciated but are no longer may still show elevated landscape change rates compared to those that have never been glaciated. Nonglacial geomorphic processes which nevertheless have been conditioned by past glaciation are termed paraglacial processes. This concept contrasts with periglacial processes, which are directly driven by formation or melting of ice or frost.

Tectonic Processes

Tectonic effects on geomorphology can range from scales of millions of years to minutes or less. The effects of tectonics on landscape are heavily dependent on the nature of the underlying bedrock fabric that more less controls what kind of local morphology tectonics can shape.

Earthquakes can, in terms of minutes, submerge large extensions creating new wetlands. Isostatic rebound can account for significant changes over thousand or hundreds of years, and allows erosion of a mountain belt to promote further erosion as mass is removed from the chain and the belt uplifts. Long-term plate tectonic dynamics give rise to orogenic belts, large mountain chains with typical lifetimes of many tens of millions of years, which form focal points for high rates of fluvial and hillslope processes and thus long-term sediment production.

Features of deeper mantle dynamics such as plumes and delamination of the lower lithosphere have also been hypothesised to play important roles in the long term (> million year), large scale (thousands of km) evolution of the Earth's topography. Both can promote surface uplift through isostasy as hotter, less dense, mantle rocks displace cooler, denser, mantle rocks at depth in the Earth.

Igneous Processes

Both volcanic (eruptive) and plutonic (intrusive) igneous processes can have important impacts on geomorphology. The action of volcanoes tends to rejuvenize landscapes, covering the old land surface with lava and tephra, releasing pyroclastic material and forcing rivers through new paths. The cones built by eruptions also build substantial new topography, which can be acted upon by other surface processes.

Biological Processes

The interaction of living organisms with landforms, or biogeomorphologic processes, can be of many different forms, and is probably of profound importance for the terrestrial geomorphic system as a whole. Biology can influence very many geomorphic processes, ranging from biogeochemical processes controlling chemical weathering, to the influence of mechanical processes like burrowing and tree throw on soil development, to even controlling global erosion rates through modulation of climate through carbon dioxide balance. Terrestrial landscapes in which the role of biology in mediating surface processes can be definitively excluded are extremely rare, but may hold important information for understanding the geomorphology of other planets, such as Mars.

Scales in Geomorphology

Different geomorphological processes dominate at different spatial and temporal scales. Moreover, scales on which processes occur may determine the reactivity or otherwise of landscapes to changes in driving forces such as climate or tectonics. These ideas are key to the study of geomorphology today.

To help categorize landscape scales some geomorphologists might use the following taxonomy:

- 1st-Continent, ocean basin, climatic zone (~10,000,000 km^2)
- 2nd-Shield, e.g. Baltic Shield, or mountain range (~1,000,000 km^2)
- 3rd-Isolated sea, Sahel (~100,000 km^2)
- 4th-Massif, e.g. Massif Central or Group of related landforms, e.g., Weald (~10,000 km^2)
- 5th-River valley, Cotswolds (~1,000 km^2)
- 6th-Individual mountain or volcano, small valleys (~100 km^2)
- 7th-Hillslopes, stream channels, estuary (~10 km^2)
- 8th-gully, barchannel (~1 km^2)
- 9th-Meter-sized features

Overlap with other Fields

There is a considerable overlap between geomorphology and other fields. Deposition of material is extremely important in sedimentology. Weathering is the chemical and physical disruption of earth materials in place on exposure to atmospheric or near surface agents, and is typically studied by soil scientists and environmental chemists, but is an essential component of geomorphology because it is what provides the material that can be moved in the first place. Civil and environmental engineers are concerned with erosion and sediment transport, especially related to canals, slope stability (and natural hazards), water quality, coastal environmental management,

transport of contaminants, and stream restoration. Glaciers can cause extensive erosion and deposition in a short period of time, making them extremely important entities in the high latitudes and meaning that they set the conditions in the headwaters of mountain-born streams; glaciology therefore is important in geomorphology.

Hydrology

Hydrology is the study of the movement, distribution, and quality of water on Earth and other planets, including the hydrologic cycle, water resources and environmental watershed sustainability. A practitioner of hydrology is a hydrologist, working within the fields of earth or environmental science, physical geography, geology or civil and environmental engineering.

Domains of hydrology include hydrometeorology, surface hydrology, hydrogeology, drainage basin management and water quality, where water plays the central role. Oceanography and meteorology are not included because water is only one of many important aspects.

Hydrological research can inform environmental engineering, policy and planning.

History of Hydrology

Hydrology has been a subject of investigation and engineering for millennia. For example, about 4000 B.C. the Nile was dammed to improve agricultural productivity of previously barren lands. Mesopotamian towns were protected from flooding with high earthen walls. Aqueducts were built by the Greeks and Ancient Romans, while the history of China shows they built irrigation and flood control works. The ancient Sinhalese used hydrology to build complex irrigation works in Sri Lanka, also known for invention of the Valve Pit which allowed construction of large reservoirs, anicuts and canals which still function.

Marcus Vitruvius, in the first century B.C., described a philosophical theory of the hydrologic cycle, in which precipitation falling in the mountains infiltrated the Earth's surface and led to streams and springs in the lowlands. With adoption of a more scientific approach, Leonardo da Vinci and Bernard Palissy independently reached an accurate representation of the hydrologic cycle. It was not until the 17th century that hydrologic variables began to be quantified.

Pioneers of the modern science of hydrology include Pierre Perrault, Edme Mariotte and Edmund Halley. By measuring rainfall, runoff, and drainage area, Perrault showed that rainfall was sufficient to account for flow of the Seine. Marriotte combined velocity and river cross-section measurements to obtain discharge, again in the Seine. Halley showed that the evaporation from the Mediterranean Sea was sufficient to account for the outflow of rivers flowing into the sea.

Advances in the 18th century included the Bernoulli piezometer and Bernoulli's equation, by Daniel Bernoulli, the Pitot tube. The 19th century saw development in

groundwater hydrology, including Darcy's law, the Dupuit-Thiem well formula, and Hagen-Poiseuille's capillary flow equation.

Rational analyses began to replace empiricism in the 20th century, while governmental agencies began their own hydrological research programs. Of particular importance were Leroy Sherman's unit hydrograph, the infiltration theory of Robert E. Horton, and C.V. Theis's Aquifer test/equation describing well hydraulics.

Since the 1950s, hydrology has been approached with a more theoretical basis than in the past, facilitated by advances in the physical understanding of hydrological processes and by the advent of computers and especially Geographic Information Systems (GIS).

Hydrologic Cycle

The central theme of hydrology is that water circulates throughout the Earth through different pathways and at different rates. The most vivid image of this is in the evaporation of water from the ocean, which forms clouds. These clouds drift over the land and produce rain. The rainwater flows into lakes, rivers, or aquifers. The water in lakes, rivers, and aquifers then either evaporates back to the atmosphere or eventually flows back to the ocean, completing a cycle. Water changes its state of being several times throughout this cycle.

Overview

Branches of Hydrology

Chemical hydrology is the study of the chemical characteristics of water.

Ecohydrology is the study of interactions between organisms and the hydrologic cycle. Hydrogeology is the study of the presence and movement of ground water. Hydroinformatics is the adaptation of information technology to hydrology and water resources applications.

Hydrometeorology is the study of the transfer of water and energy between land and water body surfaces and the lower atmosphere. Isotope hydrology is the study of the isotopic signatures of water. Surface hydrology is the study of hydrologic processes that operate at or near Earth's surface.

Drainage basin management covers water-storage, in the form of reservoirs, and flood-protection. Water quality includes the chemistry of water in rivers and lakes, both of pollutants and natural solutes.

Related Topics

Oceanography is the more general study of water in the oceans and estuaries.

Meteorology is the more general study of the atmosphere and of weather, including precipitation as snow and rainfall.

Limnology is the study of lakes. It covers the biological, chemical, physical, geological, and other attributes of all inland waters (running and standing waters, both fresh and saline, natural or man-made).

Applications of Hydrological

- Determining the water balance of a region.
- Determining the agricultural water balance.
- Designing riparian restoration projects.
- Mitigating and predicting flood, landslide and drought risk.
- Real-time flood forecasting and flood warning.
- Designing irrigation schemes and managing agricultural productivity.
- Part of the hazard module in catastrophe modelling.
- Providing drinking water.
- Designing dams for water supply or hydroelectric power generation.
- Designing bridges.
- Designing sewers and urban drainage system.
- Analyzing the impacts of antecedent moisture on sanitary sewer systems.
- Predicting geomorphological changes, such as erosion or sedimentation.
- Assessing the impacts of natural and anthropogenic environmental change on water resources.
- Assessing contaminant transport risk and establishing environmental policy guidelines.

Hydrological Measurements

Measurement is fundamental for assessing water resources and understanding the processes involved in the hydrologic cycle. Because the hydrologic cycle is so diverse, hydrologic measurement methods span many disciplines: including soils, oceanography, atmospheric science, geology, geophysics and limnology, to name a few. Here, hydrologic measurement methods are organized by hydrologic sub-disciplines. Each of these subdisciplines is addressed briefly with a practical discussion of the methods used to date and a bibliography of background information.

Quantifying groundwater flow and transport:

- Aquifer characterization
  - — Flow direction
    - * Piezometer-groundwater pressure and, by inference, groundwater depth
    - * Conductivity, storativity, transmisivity
    - * Geophysical methods

- Vadose zone characterization
  - Infiltration
    - Infiltrometer-infiltration
  - Soil moisture
    - Capacitance probe-soil moisture
    - Time domain reflectometer-soil moisture
    - Tensiometer-soil moisture
    - Solute sampling
    - Geophysical methods

Quantifying surface water flow and transport:

- Direct and indirect discharge measurements
  - Stream gauge-stream flow
  - Tracer techniques
  - Chemical transport
  - Sediment transport and erosion
  - Stream-aquifer exchange

Quantifying exchanges at the land-atmosphere boundary:

- Precipitation
  - Bulk rain events
    - Disdrometer-precipitation characteristics
    - Radar-cloud properties, rain rate estimation, hail and snow detection
    - Rain gauge-rain and snowfall
    - Satellite-rainy area identification, rain rate estimation, land-cover / land-use, soil moisture
    - Sling psychrometer-humidity
  - Snow, hail and ice
  - Dew, mist and fog
- Evaporation
  - from water surfaces
  - Evaporation-Symon's evaporation pan
  - from plant surfaces
  - through the boundary layer
- Transpiration
  - Natural ecosystems

  - — Agronomic ecosystems
- Momentum
- Heat flux
  - — Energy budgets

Uncertainty Analysis

Remote sensing of hydrologic processes:

- Land based sensors
- Airborne Sensors
- Satellite sensors

Water Quality

- Sample collection
- In-situ methods
- Physical measurements (includes sediment concentration)
- Collection of samples to quantify Organic Compounds
- Collection of samples to quantify Inorganic Compounds
- Analysis of aqueous Organic Compounds
- Analysis of aqueous Inorganic Compounds
- Microbiological sampling and analysis

Integrating Measurement and Modelling

- Budget analyses
- Parameter estimation
- Scaling in time and space
- Data assimilation
- Quality control of data — see for example Double mass analysis

Hydrological Prediction

Observations of hydrologic processes are used to make predictions of the future behaviour of hydrologic systems (water flow, water quality). One of the major current concerns in hydrologic research is “Prediction in Ungauged Basins” (PUB), i.e. in basins where no or only very few data exist.

Statistical Hydrology

By analysing the statistical properties of hydrologic records, such as rainfall or river flow, hydrologists can estimate future hydrologic phenomena, assuming the characteristics of the processes remain unchanged. When making assesments of how often relatively rare events will occur, analyses are made in terms of the return period

of such events. Other quantities of interest include the average flow in a river, in a year or by season.

These estimates are important for engineers and economists so that proper risk analysis can be performed to influence investment decisions in future infrastructure and to determine the yield reliability characteristics of water supply systems. Statistical information is utilised to formulate operating rules for large dams forming part of systems which include agricultural, industrial and residential demands.

Hydrological Modelling

Hydrological models are simplified, conceptual representations of a part of the hydrologic cycle. They are primarily used for hydrological prediction and for understanding hydrological processes. Two major types of hydrological models can be distinguished:

- Models based on data. These models are black box systems, using mathematical and statistical concepts to link a certain input (for instance rainfall) to the model output (for instance runoff). Commonly used techniques are regression, transfer functions, and system identification. The simplest of these models may be linear models, but it is common to deploy non-linear components to represent some general aspects of a catchment's response without going deeply into the real physical processes involved. An example of such an aspect is the well-known behaviour that a catchment will respond much more quickly and strongly when it is already wet than when it is dry.
- Models based on process descriptions. These models try to represent the physical processes observed in the real world. Typically, such models contain representations of surface runoff, subsurface flow, evapotranspiration, and channel flow, but they can be far more complicated. These models are known as deterministic hydrology models. Deterministic hydrology models can be subdivided into single-event models and continuous simulation models.

Recent research in hydrological modelling tries to have a more global approach to the understanding of the behaviour of hydrologic systems to make better predictions and to face the major challenges in water resources management.

Hydrologic Transport

Water movement is a significant means by which other material, such as soil or pollutants, are transported from place to place. Initial input to receiving waters may arise from a point source discharge or a line source or area source, such as surface runoff. Since the 1960s rather complex mathematical models have been developed, facilitated by the availability of high speed computers. The most common pollutant classes analysed are nutrients, pesticides, total dissolved solids and sediment.

Glaciology

Glaciology (from Middle French dialect is the study of glaciers, or more generally ice and natural phenomena that involve ice.

Glaciology is an interdisciplinary earth science that integrates geophysics, geology, physical geography, geomorphology, climatology, meteorology, hydrology, biology, and ecology. The impact of glaciers on people includes the fields of human geography and anthropology. The discoveries of water ice on the Moon, Mars and Europa add an extraterrestrial component to the field, as in "astroglaciology".

Overview

Areas of study within glaciology include glacial history and the reconstruction of past glaciation. A glaciologist is a person who studies glaciers. Glaciology is one of the key areas of polar research. A glacier is an extended mass of ice formed from snow falling and accumulating over the years and moving very slowly, either descending from high mountains, as in valley glaciers, or moving outward from centres of accumulation, as in continental glaciers.

Types

There are two general categories of glaciation which glaciologists distinguish: *alpine glaciation*, accumulations or "rivers of ice" confined to valleys; and *continental glaciation*, unrestricted accumulations which once covered much of the northern continents.

- Alpine-ice flows down the valleys of mountainous areas and forms a tongue of ice moving towards the plains below. Alpine glaciers tend to make the topography more rugged, by adding and improving the scale of existing features such as large ravines called *cirques* and ridges where the rims of two cirques meet called aretes.
- Continental-an ice sheet found today, only in high latitudes (Greenland / Antarctica), thousands of square kilometres in area and thousands of meters thick. These tend to smooth out the landscapes.

Zones of Glaciers

- Accumulation, where the formation of ice is faster than its removal.
- Wastage or Ablation, where the sum of melting and evaporation (sublimation) is greater than the amount of snow added each year.

Movement

Ablation : wastage of the glacier through sublimation, ice melting and iceberg calving.

Ablation zone : Area of a glacier in which the annual loss of ice through ablation exceeds the annual gain from precipitation.

Arete : an acute ridge of rock where two cirques abut.

Bergshrund : crevasse formed near the head of a glacier, where the mass of ice has rotated, sheared and torn itself apart in the manner of a geological fault.

Cirque, corrie or cwm : bowl shaped depression excavated by the source of a glacier.

Creep : adjustment to stress at a molecular level.

Flow : movement (of ice) in a constant direction.

Fracture : brittle failure (breaking of ice) under the stress raised when movement is too rapid to be accommodated by creep. It happens for example, as the central part of a glacier moves faster than the edges.

Horn : spire of rock formed by the headward erosion of a ring of cirques around a single mountain. It is an extreme case of an Arete.

Plucking/Quarrying : where the adhesion of the ice to the rock is stronger than the cohesion of the rock, part of the rock leaves with the flowing ice.

Tarn : a lake formed in the bottom of a cirque when its glacier has melted.

Tunnel valley : The tunnel is that formed by hydraulic erosion of ice and rock below an ice sheet margin. The tunnel valley is what remains of it in the underlying rock when the ice sheet has melted.

Glacial Deposits

Stratified

Outwash sand/gravel : from front of glaciers, found on a plain.

Kettles : block of stagnant ice leaves a depression or pit.

Eskers : steep sided ridges of gravel/sand, possibly caused by streams running under stagnant ice.

Kames : stratified drift builds up low steep hills.

Varves : alternating thin sedimentary beds (coarse and fine) of a proglacial lake. Summer conditions deposit more and coarser material and those of the winter, less and finer.

Unstratified

Till-unsorted : (glacial flour to boulders) deposited by receding/advancing glaciers, forming moraines, and drumlins.

Moraines : (Terminal) material deposited at the end; (Ground) material deposited as glacier melts; (lateral) material deposited along the sides.

Drumlins : smooth elongated hills composed of till.

Ribbed moraines : large subglacial elongated hills transverse to former ice flow.

Biogeography

Biogeography is the study of the distribution of species (biology), organisms, and ecosystems in space and through geological time. Organisms and biological communities vary in a highly regular fashion along geographic gradients of latitude, elevation, isolation and habitat area. Knowledge of spatial variation in the numbers and types of organisms is as vital to us today as it was to our early human ancestors, as we adapt to heterogeneous but geographically predictable environments. Biogeography is an integrative field of inquiry that unites concepts and information from ecology, evolutionary biology, geology, and physical geography. Modern biogeographic research combines information and ideas from many fields, from the physiological and ecological constraints on organismal dispersal to geological and climatological phenomena operating at global spatial scales and evolutionary time frames.

The patterns of species distribution across geographical areas can usually be explained through a combination of historical factors such as: speciation; extinction; continental drift; glaciation, and associated variations in sea level, river routes, and habitat; and river capture; in combination with the geographic constraints of landmass areas and isolation; and the available ecosystem energy supplies.

Over periods of ecological changes, biogeography includes the study of plant and animal species in: their past and/or present living *refugium* habitat; their interim living sites; and/or their survival locales. As writer David Quammen put it, "...biogeography does more than ask *Which species?* and *Where*. It also asks *Why?* and, what is sometimes more crucial, *Why not?*."

Modern biogeography often employs the use of Geographic Information Systems (GIS), to understand the factors affecting organism distribution, and to predict future trends in organism distribution. Often mathematical models and GIS are employed to solve ecological problems that have a spatial aspect to them.

Biogeography is most keenly observed on the world's islands. These habitats are often much more manageable areas of study because they are more condensed than larger ecosystems on the mainland. Islands are also ideal locations because they allow scientists to look at habitats that new species have only recently colonized and can observe how they disperse throughout the island, the success they achieve in these places, and they can then apply this information to similar mainland habitats. Islands are very diverse in their biomes, ranging from the tropical to arctic climates. This diversity in habitat allows for a wider range of species study in different parts of the world.

One scientist who recognized the importance of these geographic locations was Charles Darwin, who remarked in his journal "The Zoology of Archipelagoes will be well worth examination".

History

The scientific theory of biogeography grows out of the work of Alexander von Humboldt (1769–1859), Hewett Cottrell Watson (1804–1881), Alphonse de Candolle (1806–1893), Alfred Russel Wallace (1823–1913), Philip Lutley Sclater (1829–1913) and other biologists and explorers.

Wallace studied the distribution of flora and fauna in the Amazon Basin and the Malay Archipelago in the mid-19th century. Wallace and Sclater saw biogeography as a source of support for the theory of evolution. Key findings, such as the sharp difference in fauna either side of the Wallace Line, can only be understood in this light. Otherwise, the field of biogeography would be seen as a purely descriptive one.

The publication of *The Theory of Island Biogeography* by Robert MacArthur and E.O. Wilson in 1967 showed that the species richness of an area could be predicted in terms of such factors as habitat area, immigration rate and extinction rate.

This added to the long-standing interest in island biogeography. The application of island biogeography theory to habitat fragments spurred the development of the fields of conservation biology and landscape ecology.

Classic biogeography has been expanded by the development of molecular systematics, creating a new discipline known as phylogeography. This development allowed scientists to test theories about the origin and dispersal of populations, such as island endemics. For example, while classic biogeographers were able to speculate about the origins of species in the Hawaiian Islands, phylogeography allows them to test theories of relatedness between these populations and putative source populations in Asia and North America.

Paleobiogeography

Paleobiogeography goes one step further to include paleogeographic data and considerations of plate tectonics. Using molecular analyses and corroborated by fossils, it has been possible to demonstrate that perching birds evolved first in the region of Australia or the adjacent Antarctic (which at that time lay somewhat further north and had a temperate climate). From there, they spread to the other Gondwanan continents and Southeast Asia-the part of Laurasia then closest to their origin of dispersal – in the late Paleogene, before achieving a global distribution in the early Neogene. Not knowing the fact that at the time of dispersal, the Indian Ocean was much narrower than it is today, and that South America was closer to the Antarctic, one would be hard pressed to explain the presence of many “ancient” lineages of perching birds in Africa, as well as the mainly South American distribution of the suboscines.

Paleobiogeography also helps constrain hypotheses on the timing of biogeographic events such as vicariance and geodispersal, and provides unique information on the

formation of regional biotas. For example, data from species-level phylogenetic and biogeographic studies tell us that the Amazonian fish fauna accumulated incrementally over a period of tens of millions of years, principally by means of allopatric speciation, and in an arena extending over most of the area of tropical South America (Albert & Reis 2011). In other words, unlike some of the well-known insular faunas (Galapagos finches, Hawaiian drosophilid flies, African rift lake cichlids), the species-rich Amazonian ichthyofauna is not the result of recent adaptive radiations. For freshwater organisms, landscapes are divided naturally into discrete drainage basins by watersheds, episodically isolated and reunited by erosional processes. In regions like the Amazon Basin with an exceptionally low (flat) topographic relief, the many waterways have had a highly reticulated history over geological time. In such a context stream capture is an important factor affecting the evolution and distribution of freshwater organisms. Stream capture occurs when an upstream portion of one river drainage is diverted to the downstream portion of an adjacent basin. This can happen because of tectonic uplift (or subsidence), natural damming as a result of a landslide, or by headward or lateral erosion of the watershed between adjacent basins.

Classification

Biogeography is a synthetic science, related to geography, biology, soil science, geology, climatology, ecology and evolution.

Some fundamental concepts in biogeography include:

- evolution – change in genetic composition of a population
- extinction – disappearance of a species
- dispersal – movement of populations away from their point of origin, related to migration
- geodispersal – the erosion of barriers to biotic dispersal and gene flow, that permit range expansion and the merging of previously isolated biotas
- range and distribution
- endemic areas
- vicariance – the formation of barriers to biotic dispersal and gene flow, that tend to subdivide species and biotas, leading to speciation and extinction

Climatology

Climatology is the study of climate, scientifically defined as weather conditions averaged over a period of time, and is a branch of the atmospheric sciences. Basic knowledge of climate can be used within shorter term weather forecasting using analog techniques such as the El Niño – Southern Oscillation (ENSO), the Madden-Julian Oscillation (MJO), the North Atlantic Oscillation (NAO), the Northern Annualar

Mode (NAM), the Arctic oscillation (AO), the Northern Pacific (NP) Index, the Pacific Decadal Oscillation (PDO), and the Interdecadal Pacific Oscillation (IPO). Climate models are used for a variety of purposes from study of the dynamics of the weather and climate system to projections of future climate.

History

The earliest person to hypothesize the concept of climate change may have been the medieval Chinese scientist Shen Kuo (1031–95). Shen Kuo theorized that climates naturally shifted over an enormous span of time, after observing petrified bamboos found underground near Yanzhou (modern day Yan'an, Shaanxi province), a dry climate area unsuitable for the growth of bamboo.

Early climate researchers include Edmund Halley, who published a map of the trade winds in 1686, after a voyage to the southern hemisphere. Benjamin Franklin, in the 18th century, was the first to map the course of the Gulf Stream for use in sending mail overseas from the United States to Europe. Francis Galton invented the term *anticyclone*. Helmut Landsberg led to statistical analysis being used in climatology, which led to its evolution into a physical sciences.

Different Approaches

Climatology is approached in a variety of ways. Paleoclimatology seeks to reconstruct past climates by examining records such as ice cores and tree rings (dendroclimatology). Paleotempestology uses these same records to help determine hurricane frequency over millennia. The study of contemporary climates incorporates meteorological data accumulated over many years, such as records of rainfall, temperature and atmospheric composition. Knowledge of the atmosphere and its dynamics is also embodied in models, either statistical or mathematical, which help by integrating different observations and testing how they fit together. Modelling is used for understanding past, present and potential future climates. Historical climatology is the study of climate as related to human history and thus focuses only on the last few thousand years.

Climate research is made difficult by the large scale, long time periods, and complex processes which govern climate. Climate is governed by physical laws which can be expressed as differential equations. These equations are coupled and nonlinear, so that approximate solutions are obtained by using numerical methods to create global climate models. Climate is sometimes modeled as a stochastic process but this is generally accepted as an approximation to processes that are otherwise too complicated to analyse.

Indices

Scientists use climate indices based on several climate patterns (known as modes of variability) in their attempt to characterize and understand the various climate

mechanisms that culminate in our daily weather. Much in the way the Dow Jones Industrial Average, which is based on the stock prices of 30 companies, is used to represent the fluctuations in the stock market as a whole, climate indices are used to represent the essential elements of climate. Climate indices are generally devised with the twin objectives of simplicity and completeness, and each index typically represents the status and timing of the climate factor it represents. By their very nature, indices are simple, and combine many details into a generalized, overall description of the atmosphere or ocean which can be used to characterize the factors which impact the global climate system.

El Niño – Southern Oscillation

El Niño-Southern Oscillation (ENSO) is a global coupled ocean-atmosphere phenomenon. The Pacific ocean signatures, El Nino and La Niña are important temperature fluctuations in surface waters of the tropical Eastern Pacific Ocean. The name El Nino, from the Spanish for "the little boy", refers to the Christ child, because the phenomenon is usually noticed around Christmas time in the Pacific Ocean off the west coast of South America. La Nina means "the little girl". Their effect on climate in the subtropics and the tropics are profound. The atmospheric signature, the Southern Oscillation (SO) reflects the monthly or seasonal fluctuations in the air pressure difference between Tahiti and Darwin. The most recent occurrence of El Nino started in September 2006 and lasted until early 2007. ENSO is a set of interacting parts of a single global system of coupled ocean-atmosphere climate fluctuations that come about as a consequence of oceanic and atmospheric circulation. ENSO is the most prominent known source of inter-annual variability in weather and climate around the world. The cycle occurs every two to seven years, with El Nino lasting nine months to two years within the longer term cycle, though not all areas globally are affected. ENSO has signatures in the Pacific, Atlantic and Indian Oceans.

In the Pacific, during major warm events, El Nino warming extends over much of the tropical Pacific and becomes clearly linked to the SO intensity. While ENSO events are basically in phase between the Pacific and Indian Oceans, ENSO events in the Atlantic Ocean lag behind those in the Pacific by 12–18 months. Many of the countries most affected by ENSO events are developing countries within tropical sections of continents with economies that are largely dependent upon their agricultural and fishery sectors as a major source of food supply, employment, and foreign exchange. New capabilities to predict the onset of ENSO events in the three oceans can have global socio-economic impacts. While ENSO is a global and natural part of the Earth's climate, whether its intensity or frequency may change as a result of global warming is an important concern. Low-frequency variability has been evidenced: the quasi-decadal oscillation (QDO). Inter-decadal (ID) modulation of ENSO (from PDO or IPO) might exist. This could explain the so-called protracted ENSO of the early 1990s.

Madden–Julian Oscillation

The Madden–Julian Oscillation (MJO) is an equatorial travelling pattern of anomalous rainfall that is planetary in scale. It is characterized by an eastward progression of large regions of both enhanced and suppressed tropical rainfall, observed mainly over the Indian and Pacific Oceans. The anomalous rainfall is usually first evident over the western Indian Ocean, and remains evident as it propagates over the very warm ocean waters of the western and central tropical Pacific. This pattern of tropical rainfall then generally becomes very nondescript as it moves over the cooler ocean waters of the eastern Pacific but reappears over the tropical Atlantic and Indian Oceans. The wet phase of enhanced convection and precipitation is followed by a dry phase where convection is suppressed. Each cycle lasts approximately 30–60 days. The MJO is also known as the 30–60 day oscillation, 30–60 day wave, or intraseasonal oscillation.

North Atlantic Oscillation (NAO)

Indices of the NAO are based on the difference of normalized sea level pressure (SLP) between Ponta Delgada, Azores and Stykkisholmur/Reykjavik, Iceland. The SLP anomalies at each station were normalized by division of each seasonal mean pressure by the long-term mean (1865–1984) standard deviation. Normalization is done to avoid the series of being dominated by the greater variability of the northern of the two stations. Positive values of the index indicate stronger-than-average westerlies over the middle latitudes.

Northern Annular Mode (NAM) or Arctic Oscillation (AO)

The NAM, or AO, is defined as the first EOF of northern hemisphere winter SLP data from the tropics and subtropics. It explains 23% of the average winter (December–March) variance, and it is dominated by the NAO structure in the Atlantic. Although there are some subtle differences from the regional pattern over the Atlantic and Arctic, the main difference is larger amplitude anomalies over the North Pacific of the same sign as those over the Atlantic. This feature gives the NAM a more annular (or zonally symmetric) structure.

Northern Pacific (NP) Index

The NP Index is the area-weighted sea level pressure over the region 30N–65N, 160E–140W.

Pacific Decadal Oscillation (PDO)

The PDO is a pattern of Pacific climate variability that shifts phases on at least inter-decadal time scale, usually about 20 to 30 years. The PDO is detected as warm or cool surface waters in the Pacific Ocean, north of 20° N. During a "warm", or "positive", phase, the west Pacific becomes cool and part of the eastern ocean warms; during a "cool" or "negative" phase, the opposite pattern occurs. The mechanism by

which the pattern lasts over several years has not been identified; one suggestion is that a thin layer of warm water during summer may shield deeper cold waters. A PDO signal has been reconstructed to 1661 through tree-ring chronologies in the Baja California area.

Interdecadal Pacific Oscillation (IPO)

The Interdecadal Pacific Oscillation (IPO or ID) display similar sea surface temperature (SST) and sea level pressure patterns to the PDO, with a cycle of 15–30 years, but affects both the north and south Pacific. In the tropical Pacific, maximum SST anomalies are found away from the equator. This is quite different from the quasi-decadal oscillation (QDO) with a period of 8–12 years and maximum SST anomalies straddling the equator, thus resembling ENSO.

Models

Climate models use quantitative methods to simulate the interactions of the atmosphere, oceans, land surface, and ice. They are used for a variety of purposes from study of the dynamics of the weather and climate system to projections of future climate. All climate models balance, or very nearly balance, incoming energy as short wave (including visible) electromagnetic radiation to the earth with outgoing energy as long wave (infrared) electromagnetic radiation from the earth. Any unbalance results in a change in the average temperature of the earth.

The most talked-about models of recent years have been those relating temperature to emissions of carbon dioxide. These models predict an upward trend in the surface temperature record, as well as a more rapid increase in temperature at higher latitudes.

Models can range from relatively simple to quite complex:

- A simple radiant heat transfer model that treats the earth as a single point and averages outgoing energy
- this can be expanded vertically (radiative-convective models), or horizontally
- finally, (coupled) atmosphere–ocean–sea ice global climate models discretise and solve the full equations for mass and energy transfer and radiant exchange

Differences with Meteorology

In contrast to meteorology, which focuses on short term weather systems lasting up to a few weeks, climatology studies the frequency and trends of those systems. It studies the periodicity of weather events over years to millennia, as well as changes in long-term average weather patterns, in relation to atmospheric conditions. Climatologists, those who practice climatology, study both the nature of climates – local, regional or global – and the natural or human-induced factors that cause climates to change. Climatology considers the past and can help predict future climate change.

Phenomena of climatological interest include the atmospheric boundary layer, circulation patterns, heat transfer (radiative, convective and latent), interactions between the atmosphere and the oceans and land surface (particularly vegetation, land use and topography), and the chemical and physical composition of the atmosphere.

Use in Weather Forecasting

A more complicated way of making a forecast, the analog technique requires remembering a previous weather event which is expected to be mimicked by an upcoming event. What makes it a difficult technique to use is that there is rarely a perfect analog for an event in the future. Some call this type of forecasting pattern recognition, which remains a useful method of observing rainfall over data voids such as oceans with knowledge of how satellite imagery relates to precipitation rates over land, as well as the forecasting of precipitation amounts and distribution in the future. A variation on this theme is used in Medium Range forecasting, which is known as teleconnections, when you use systems in other locations to help pin down the location of another system within the surrounding regime. One method of using teleconnections are by using climate indices such as ENSO-related phenomena.

Meteorology

Meteorology is the interdisciplinary scientific study of the atmosphere. Studies in the field stretch back millennia, though significant progress in meteorology did not occur until the eighteenth century. The nineteenth century saw breakthroughs occur after observing networks developed across several countries. After the development of the computer in the latter half of the twentieth century breakthroughs in weather forecasting were achieved.

Meteorological phenomena are observable weather events which illuminate and are explained by the science of meteorology. Those events are bound by the variables that exist in Earth's atmosphere; temperature, air pressure, water vapour, and the gradients and interactions of each variable, and how they change in time. Different spatial scales are studied to determine how systems on local, region, and global levels impact weather and climatology.

Meteorology, climatology, atmospheric physics, and atmospheric chemistry are sub-disciplines of the atmospheric sciences. Meteorology and hydrology compose the interdisciplinary field of hydrometeorology. Interactions between Earth's atmosphere and the oceans are part of coupled ocean-atmosphere studies. Meteorology has application in many diverse fields such as the military, energy production, transport, agriculture and construction.

History

The beginnings of meteorology can be traced back in ancient India to 3000 B.C.E, such as the Upanishads, contain serious discussion about the processes of cloud

formation and rain and the seasonal cycles caused by the movement of earth round the sun. Varahamithra's classical work, Brihatsamhita, written about 500 A.D., provides a clear evidence that a deep knowledge of atmospheric processes existed even in those times. In 350 BC, Aristotle wrote *Meteorology*. Aristotle is considered the founder of meteorology. One of the most impressive achievements described in the *Meteorology* is the description of what is now known as the hydrologic cycle. The Greek scientist Theophrastus compiled a book on weather forecasting, called the *Book of Signs*. The work of Theophrastus remained a dominant influence in the study of weather and in weather forecasting for nearly 2,000 years. In 25 AD, Pomponius Mela, a geographer for the Roman Empire, formalized the climatic zone system. Around the 9th century, Al-Dinawari, a Kurdish naturalist, writes the *Kitab al-Nabat* (*Book of Plants*), in which he deals with the application of meteorology to agriculture during the Muslim Agricultural Revolution. He describes the meteorological character of the sky, the planets and constellations, the sun and moon, the lunar phases indicating seasons and rain, the *anwa* (heavenly bodies of rain), and atmospheric phenomena such as winds, thunder, lightning, snow, floods, valleys, rivers, lakes, wells and other sources of water.

Research of Visual Atmospheric Phenomena

In 1021, Ibn al-Haytham (Alhazen) wrote on the atmospheric refraction of light. He showed that the twilight is due to atmospheric refraction and only begins when the Sun is 19 degrees below the horizon, and uses a complex geometric demonstration to measure the height of the Earth's atmosphere as 52,000 *passuum* (49 miles (79 km)), which is very close to the modern measurement of 50 miles (80 km). He also realized that the atmosphere also reflects light, from his observations of the sky brightening even before the Sun rises.

St. Albert the Great was the first to propose that each drop of falling rain had the form of a small sphere, and that this form meant that the rainbow was produced by light interacting with each raindrop. Roger Bacon was the first to calculate the angular size of the rainbow. He stated that the rainbow summit can not appear higher than 42 degrees above the horizon. In the late 13th century and early 14th century, Theodoric of Freiberg and Kamal al-Dîn al-Fârisî continued the work of Ibn al-Haytham, and they were the first to give the correct explanations for the primary rainbow phenomenon. Theoderic went further and also explained the secondary rainbow In 1716, Edmund Halley suggests that aurorae are caused by "magnetic effluvia" moving along the Earth's magnetic field lines.

Instruments and Classification Scales

In 1441, King Sejongs son, Prince Munjong, invented the first standardized rain gauge. These were sent throughout the Joseon Dynasty of Korea as an official tool to assess land taxes based upon a farmer's potential harvest. In 1450, Leone Battista

Alberti developed a swinging-plate anemometer, and is known as the first *anemometer*. In 1607, Galileo Galilei constructs a thermoscope. In 1611, Johannes Kepler writes the first scientific treatise on snow crystals: "Strena Seu de Nive Sexangula (A New Year's Gift of Hexagonal Snow)". In 1643, Evangelista Torricelli invents the mercury barometer. In 1662, Sir Christopher Wren invented the mechanical, self-emptying, tipping bucket rain gauge. In 1714, Gabriel Fahrenheit creates a reliable scale for measuring temperature with a mercury-type thermometer. In 1742, Anders Celsius, a Swedish astronomer, proposed the 'centigrade' temperature scale, the predecessor of the current Celsius scale. In 1783, the first hair hygrometer is demonstrated by Horace-Bénédict de Saussure. In 1802-1803, Luke Howard writes *On the Modification of Clouds* in which he assigns cloud types Latin names. In 1806, Francis Beaufort introduced his system for classifying wind speeds. Near the end of the 19th century the first cloud atlases were published, including the *International Cloud Atlas*, which has remained in print ever since. The April 1960 launch of the first successful weather satellite, TIROS-1, marked the beginning of the age where weather information became available globally.

Atmospheric Composition Research

In 1648, Blaise Pascal rediscovers that atmospheric pressure decreases with height, and deduces that there is a vacuum above the atmosphere. In 1738, Daniel Bernoulli publishes *Hydrodynamics*, initiating the kinetic theory of gases and established the basic laws for the theory of gases. In 1761, Joseph Black discovers that ice absorbs heat without changing its temperature when melting. In 1772, Black's student Daniel Rutherford discovers nitrogen, which he calls *phlogisticated air*, and together they developed the phlogiston theory. In 1777, Antoine Lavoisier discovers oxygen and develops an explanation for combustion. In 1783, in Lavoisier's book *Reflexions sur le phlogistique*, he deprecates the phlogiston theory and proposes a caloric theory. In 1804, Sir John Leslie observes that a matte black surface radiates heat more effectively than a polished surface, suggesting the importance of black body radiation. In 1808, John Dalton defends caloric theory in *A New System of Chemistry* and describes how it combines with matter, especially gases; he proposes that the heat capacity of gases varies inversely with atomic weight. In 1824, Sadi Carnot analyses the efficiency of steam engines using caloric theory; he develops the notion of a reversible process and, in postulating that no such thing exists in nature, lays the foundation for the second law of thermodynamics.

Research into Cyclones and Air Flow

In 1494, Christopher Columbus experiences a tropical cyclone, leads to the first written European account of a hurricane. In 1686, Edmund Halley presents a systematic study of the trade winds and monsoons and identifies solar heating as the cause of atmospheric motions. In 1735, an *ideal* explanation of global circulation through study of the Trade winds was written by George Hadley. In 1743, when Benjamin Franklin is prevented from seeing a lunar eclipse by a hurricane, he decides that cyclones move

in a contrary manner to the winds at their periphery. Understanding the kinematics of how exactly the rotation of the Earth affects airflow was partial at first. Gaspard-Gustave Coriolis published a paper in 1835 on the energy yield of machines with rotating parts, such as waterwheels. In 1856, William Ferrel proposed the existence of a circulation cell in the mid-latitudes with air being deflected by the Coriolis force to create the prevailing westerly winds. Late in the 19th century the full extent of the large scale interaction of pressure gradient force and deflecting force that in the end causes air masses to move *along* isobars was understood. By 1912, this deflecting force was named the Coriolis effect. Just after World War I, a group of meteorologists in Norway led by Vilhelm Bjerknes developed the Norwegian cyclone model that explains the generation, intensification and ultimate decay (the life cycle) of mid-latitude cyclones, introducing the idea of fronts, that is, sharply defined boundaries between air masses. The group included Carl-Gustaf Rossby (who was the first to explain the large scale atmospheric flow in terms of fluid dynamics), Tor Bergeron (who first determined the mechanism by which rain forms) and Jacob Bjerknes.

Observation Networks and Weather Forecasting

In 1654, Ferdinando II de Medici establishes the first *weather observing* network, that consisted of meteorological stations in Florence, Cutigliano, Vallombrosa, Bologna, Parma, Milan, Innsbruck, Osnabruck, Paris and Warsaw. Collected data was centrally sent to Florence at regular time intervals. In 1832, an electromagnetic telegraph was created by Baron Schilling. The arrival of the electrical telegraph in 1837 afforded, for the first time, a practical method for quickly gathering surface weather observations from a wide area. This data could be used to produce maps of the state of the atmosphere for a region near the Earth's surface and to study how these states evolved through time. To make frequent weather forecasts based on these data required a reliable network of observations, but it was not until 1849 that the Smithsonian Institution began to establish an observation network across the United States under the leadership of Joseph Henry. Similar observation networks were established in Europe at this time. In 1854, the United Kingdom government appointed Robert FitzRoy to the new office of *Meteorological Statist to the Board of Trade* with the role of gathering weather observations at sea. FitzRoy's office became the United Kingdom Meteorological Office in 1854, the first national meteorological service in the world. The first daily weather forecasts made by FitzRoy's Office were published in *The Times* newspaper in 1860. The following year a system was introduced of hoisting storm warning cones at principal ports when a gale was expected.

Over the next 50 years many countries established national meteorological services. The India Meteorological Department (1875) was established following tropical cyclone and monsoon related famines in the previous decades. The Finnish Meteorological Central Office (1881) was formed from part of Magnetic Observatory of Helsinki University. Japan's Tokyo Meteorological Observatory, the forerunner of the Japan

Meteorological Agency, began constructing surface weather maps in 1883. The United States Weather Bureau (1890) was established under the United States Department of Agriculture. The Australian Bureau of Meteorology (1906) was established by a Meteorology Act to unify existing state meteorological services.

Numerical Weather Prediction

In 1904, Norwegian scientist Vilhelm Bjerknes first argued in his paper *Weather Forecasting as a Problem in Mechanics and Physics* that it should be possible to forecast weather from calculations based upon natural laws.

It was not until later in the 20th century that advances in the understanding of atmospheric physics led to the foundation of modern numerical weather prediction. In 1922, Lewis Fry Richardson published "Weather Prediction By Numerical Process", after finding notes and derivations he worked on as an ambulance driver in World War I. He described therein how small terms in the prognostic fluid dynamics equations governing atmospheric flow could be neglected, and a finite differencing scheme in time and space could be devised, to allow numerical prediction solutions to be found. Richardson envisioned a large auditorium of thousands of people performing the calculations and passing them to others. However, the sheer number of calculations required was too large to be completed without the use of computers, and the size of the grid and time steps led to unrealistic results in deepening systems. It was later found, through numerical analysis, that this was due to numerical instability.

Starting in the 1950s, numerical forecasts with computers became feasible. The first weather forecasts derived this way used barotropic (that means, single-vertical-level) models, and could successfully predict the large-scale movement of midlatitude Rossby waves, that is, the pattern of atmospheric lows and highs.

In the 1960s, the chaotic nature of the atmosphere was first observed and mathematically described by Edward Lorenz, founding the field of chaos theory. These advances have led to the current use of ensemble forecasting in most major forecasting centres, to take into account uncertainty arising from the chaotic nature of the atmosphere. Climate models have been developed that feature a resolution comparable to older weather prediction models. These climate models are used to investigate long-term climate shifts, such as what effects might be caused by human emission of greenhouse gases.

Meteorologists

Meteorologists are scientists who study meteorology. Meteorologists work in government agencies, private consulting and research services, industrial enterprises, utilities, radio and television stations, and in education. In the United States, meteorologists held about 9,400 jobs in 2009. Meteorologists are best-known for forecasting the weather. Many radio and television weather forecasters are professional meteorologists, while others are merely reporters with no formal meteorological

training. The American Meteorological Society and National Weather Association issue “Seals of Approval” to weather broadcasters who meet certain requirements.

Equipment

Each science has its own unique sets of laboratory equipment. In the atmosphere, there are many things or qualities of the atmosphere that can be measured. Rain, which can be observed, or seen anywhere and anytime was one of the first ones to be measured historically. Also, two other accurately measured *qualities* are wind and humidity. Neither of these can be *seen* but can be felt. The devices to measure these three sprang up in the mid-15th century and were respectively the rain gauge, the anemometer, and the hygrometer. Sets of surface measurements are important data to meteorologists. They give a snapshot of a variety of weather conditions at one single location and are usually at a weather station, a ship or a weather buoy. The measurements taken at a weather station can include any number of atmospheric observables. Usually, temperature, pressure, wind measurements, and humidity are the variables that are measured by a thermometer, barometer, anemometer, and hygrometer, respectively. Upper air data are of crucial importance for weather forecasting. The most widely used technique is launches of radiosondes. Supplementing the radiosondes a network of aircraft collection is organized by the World Meteorological Organization.

Remote sensing, as used in meteorology, is the concept of collecting data from remote weather events and subsequently producing weather information. The common types of remote sensing are Radar, Lidar, and satellites (or photogrammetry). Each collects data about the atmosphere from a remote location and, usually, stores the data where the instrument is located. RADAR and LIDAR are not passive because both use EM radiation to illuminate a specific portion of the atmosphere. Weather satellites along with more general-purpose Earth-observing satellites circling the earth at various altitudes have become an indispensable tool for studying a wide range of phenomena from forest fires to El Nino.

Spatial Scales

In the study of the atmosphere, meteorology can be divided into distinct areas of emphasis depending on the temporal scope and spatial scope of interest. At one extreme of this scale is climatology. In the timescales of hours to days, meteorology separates into micro-, meso-, and synoptic scale meteorology. Respectively, the geospatial size of each of these three scales relates directly with the appropriate timescale. Other sub-classifications are available based on the need by or by the unique, local or broad effects that are studied within that sub-class.

Microscale

Microscale meteorology is the study of atmospheric phenomena of about 1 km or less. Individual thunderstorms, clouds, and local turbulence caused by buildings and other obstacles, such as individual hills fall within this category.

Mesoscale

Mesoscale meteorology is the study of atmospheric phenomena that has horizontal scales ranging from microscale limits to synoptic scale limits and a vertical scale that starts at the Earth's surface and includes the atmospheric boundary layer, troposphere, tropopause, and the lower section of the stratosphere. Mesoscale timescales last from less than a day to the lifetime of the event, which in some cases can be weeks. The events typically of interest are thunderstorms, squall lines, fronts, precipitation bands in tropical and extratropical cyclones, and topographically generated weather systems such as mountain waves and sea and land breezes.

Synoptic Scale

Synoptic scale meteorology is generally large area dynamics referred to in horizontal coordinates and with respect to time. The phenomena typically described by synoptic meteorology include events like extratropical cyclones, baroclinic troughs and ridges, frontal zones, and to some extent jet streams. All of these are typically given on weather maps for a specific time. The minimum horizontal scale of synoptic phenomena are limited to the spacing between surface observation stations.

Global Scale

Global scale meteorology is study of weather patterns related to the transport of heat from the tropics to the poles. Also, very large scale oscillations are of importance. These oscillations have time periods typically on the order of months, such as the Madden-Julian Oscillation, or years, such as the El Nino-Southern Oscillation and the Pacific decadal oscillation. Global scale pushes the thresholds of the perception of meteorology into climatology. The traditional definition of climate is pushed in to larger timescales with the further understanding of how the global oscillations cause both climate and weather disturbances in the synoptic and mesoscale timescales.

Numerical Weather Prediction is a main focus in understanding air-sea interaction, tropical meteorology, atmospheric predictability, and tropospheric/stratospheric processes. The Naval Research Laboratory in Monterey produces the atmospheric model called NOGAPS, a global scale atmospheric model, this model is run operationally at Fleet Numerical Meteorology and Oceanography Centre. Many other global atmospheric models are run by national meteorological agencies.

Some Meteorological Principles

Boundary Layer Meteorology

Boundary layer meteorology is the study of processes in the air layer directly above Earth's surface, known as the atmospheric boundary layer (ABL). The effects of the surface – heating, cooling, and friction – cause turbulent mixing within the air layer. Significant fluxes of heat, matter, or momentum on time scales of less than a

day are advected by turbulent motions. Boundary layer meteorology includes the study of all types of surface-atmosphere boundary, including ocean, lake, urban land and non-urban land.

Dynamic Meteorology

Dynamic meteorology generally focuses on the fluid dynamics of the atmosphere. The idea of air parcel is used to define the smallest element of the atmosphere, while ignoring the discrete molecular and chemical nature of the atmosphere. An air parcel is defined as a point in the fluid continuum of the atmosphere. The fundamental laws of fluid dynamics, thermodynamics, and motion are used to study the atmosphere. The physical quantities that characterize the state of the atmosphere are temperature, density, pressure, etc. These variables have unique values in the continuum.

Applications

Weather Forecasting

Weather forecasting is the application of science and technology to predict the state of the atmosphere for a future time and a given location. Human beings have attempted to predict the weather informally for millennia, and formally since at least the nineteenth century. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere and using scientific understanding of atmospheric processes to project how the atmosphere will evolve.

Once an all-human endeavour based mainly upon changes in barometric pressure, current weather conditions, and sky condition, forecast models are now used to determine future conditions. Human input is still required to pick the best possible forecast model to base the forecast upon, which involves pattern recognition skills, teleconnections, knowledge of model performance, and knowledge of model biases. The chaotic nature of the atmosphere, the massive computational power required to solve the equations that describe the atmosphere, error involved in measuring the initial conditions, and an incomplete understanding of atmospheric processes mean that forecasts become less accurate as the difference in current time and the time for which the forecast is being made (the *range* of the forecast) increases. The use of ensembles and model consensus help narrow the error and pick the most likely outcome.

There are a variety of end uses to weather forecasts. Weather warnings are important forecasts because they are used to protect life and property. Forecasts based on temperature and precipitation are important to agriculture, and therefore to commodity traders within stock markets. Temperature forecasts are used by utility companies to estimate demand over coming days. On an everyday basis, people use weather forecasts to determine what to wear on a given day. Since outdoor activities are severely curtailed by heavy rain, snow and the wind chill, forecasts can be used to plan activities around these events, and to plan ahead and survive them.

Aviation Meteorology

Aviation meteorology deals with the impact of weather on air traffic management. It is important for air crews to understand the implications of weather on their flight plan as well as their aircraft, as noted by the Aeronautical Information Manual:

The effects of ice on aircraft are cumulative-thrust is reduced, drag increases, lift lessens, and weight increases. The results are an increase in stall speed and a deterioration of aircraft performance. In extreme cases, 2 to 3 inches of ice can form on the leading edge of the airfoil in less than 5 minutes. It takes but 1/2 inch of ice to reduce the lifting power of some aircraft by 50 percent and increases the frictional drag by an equal percentage.

Agricultural Meteorology

Meteorologists, soil scientists, agricultural hydrologists, and agronomists are persons concerned with studying the effects of weather and climate on plant distribution, crop yield, water-use efficiency, phenology of plant and animal development, and the energy balance of managed and natural ecosystems. Conversely, they are interested in the role of vegetation on climate and weather.

Hydrometeorology

Hydrometeorology is the branch of meteorology that deals with the hydrologic cycle, the water budget, and the rainfall statistics of storms. A hydrometeorologist prepares and issues forecasts of accumulating (quantitative) precipitation, heavy rain, heavy snow, and highlights areas with the potential for flash flooding. Typically the range of knowledge that is required overlaps with climatology, mesoscale and synoptic meteorology, and other geosciences.

Nuclear Meteorology

Nuclear meteorology investigates the distribution of radioactive aerosols and gases in the atmosphere.

Maritime Meteorology

Maritime meteorology deals with air and wave forecasts for ships operating at sea. Organizations such as the Ocean Prediction Centre, Honolulu National Weather Service forecast office, United Kingdom Met Office, and JMA prepare high seas forecasts for the world's oceans.

Pedology (Soil Study)

Pedology is the study of soils in their natural environment. It is one of two main branches of soil science, the other being edaphology. Pedology deals with pedogenesis, soil morphology, and soil classification, while edaphology studies the way soils influence plants, fungi, and other living things.

Overview

Soil is not only a support for vegetation, but it is also the zone (the pedosphere) of numerous interactions between climate (water, air, temperature), soil life (micro-organisms, plants, animals) and its residues, the mineral material of the original and added rock, and its position in the landscape. During its formation and genesis, the soil profile slowly deepens and develops characteristic layers, called 'horizons', while a steady state balance is approached.

Soil users (such as agronomists) showed initially little concern in the dynamics of soil. They saw it as medium whose chemical, physical and biological properties were useful for the services of agronomic productivity.

On the other hand, pedologists and geologists did not initially focus on the agronomic applications of the soil characteristics (edaphic properties) but upon its relation to the nature and history of landscapes. Today, there's an integration of the two disciplinary approaches as part of landscape and environmental sciences.

Pedologists are now also interested in the practical applications of a good understanding of pedogenesis processes (the evolution and functioning of soils), like interpreting its environmental history and predicting consequences of changes in land use, while agronomists understand that the cultivated soil is a complex medium, often resulting from several thousands of years of evolution.

They understand that the current balance is fragile and that only a thorough knowledge of its history makes it possible to ensure its sustainable use.

Concepts

- Complexity in soil genesis is more common than simplicity.
- Soils lie at the interface of Earth's atmosphere, biosphere, hydrosphere and lithosphere. Therefore, a thorough understanding of soils requires some knowledge of meteorology, climatology, ecology, biology, hydrology, geomorphology, geology and many other earth sciences and natural sciences.
- Contemporary soils carry imprints of pedogenic processes that were active in the past, although in many cases these imprints are difficult to observe or quantify. Thus, knowledge of paleoecology, palaeogeography, glacial geology and paleoclimatology is important for the recognition and understanding of soil genesis and constitute a basis for predicting the future soil changes.
- Five major, external factors of soil formation (climate, organisms, relief, parent material and time), and several smaller, less identifiable ones, drive pedogenic processes and create soil patterns.
- Characteristics of soils and soil landscapes, e.g., the number, sizes, shapes and arrangements of soil bodies, each of which is characterized on the basis of soil horizons, degree of internal homogeneity, slope, aspect, landscape position,

age and other properties and relationships, can be observed and measured.

- Distinctive bioclimatic regimes or combinations of pedogenic processes produce distinctive soils. Thus, distinctive, observable morphological features, e.g., illuvial clay accumulation in B horizons, are produced by certain combinations of pedogenic processes operative over varying periods of time.
- Pedogenic (soil-forming) processes act to both create and destroy order (anisotropy) within soils; these processes can proceed simultaneously. The resulting soil profile reflects the balance of these processes, present and past.
- The geological Principle of Uniformitarianism applies to soils, i.e., pedogenic processes active in soils today have been operating for long periods of time, back to the time of appearance of organisms on the land surface. These processes do, however, have varying degrees of expression and intensity over space and time.
- A succession of different soils may have developed, eroded and / or regressed at any particular site, as soil genetic factors and site factors, e.g., vegetation, sedimentation, geomorphology, change.
- There are very few old soils (in a geological sense) because they can be destroyed or buried by geological events, or modified by shifts in climate by virtue of their vulnerable position at the surface of the earth. Little of the soil continuum dates back beyond the Tertiary period and most soils and land surfaces are no older than the Pleistocene Epoch. However, preserved / lithified soils (paleosols) are an almost ubiquitous feature in terrestrial (land-based) environments throughout most of geologic time. Since they record evidence of ancient climate change, they present immense utility in understanding climate evolution throughout geologic history.
- Knowledge and understanding of the genesis of a soil is important in its classification and mapping.
- Soil classification systems cannot be based entirely on perceptions of genesis, however, because genetic processes are seldom observed and because pedogenic processes change over time.
- Knowledge of soil genesis is imperative and basic to soil use and management. Human influence on, or adjustment to, the factors and processes of soil formation can be best controlled and planned using knowledge about soil genesis.
- Soils are natural clay factories (clay includes both clay mineral structures and particles less than 2 μm in diameter). Shales worldwide are, to a considerable extent, simply soil clays that have been formed in the pedosphere and eroded and deposited in the ocean basins, to become lithified at a later date.

Palaeogeography

Palaeogeography (also spelled paleogeography) is the study of what the geography was in times past. It is most often used about the physical landscape, although nothing excludes its use in reference to the human or cultural environment. If the topic is landforms it could also be called paleogeomorphology.

In petroleum geology the term paleogeographic analysis is used for the detailed study of sedimentary basins, since the ancient geomorphological environments of the Earth's surface are preserved in the stratigraphic record. Paleogeographers also study the sedimentary environment associated with fossils to aid in the understanding of evolutionary development of extinct species. The reconstruction of prehistoric continents and oceans depends partly on paleogeographic evidence. Thus paleogeography provides critical evidence for the development of continental drift and current plate tectonic theories. For example, knowledge of the shape and latitudinal location of supercontinents such as Pangaea and ancient oceans such as Panthalassa result from paleogeographic studies.

Not to be confused with palaeography (the study of ancient handwriting).

Coastal Geography

Coastal geography is the study of the dynamic interface between the ocean and the land, incorporating both the physical geography (i.e coastal geomorphology, geology and oceanography) and the human geography (sociology and history) of the coast. It involves an understanding of coastal weathering processes, particularly wave action, sediment movement and weather, and also the ways in which humans interact with the coast. Coastal geography is that branch of geography, incorporating physical and human geography, which deals with the study of the dynamic interface between ocean and land.

Wave Action and Longshore Drift

The waves of different strengths that constantly hit against the shoreline are the primary movers and shapers of the coastline. Despite the simplicity of this process, the differences between waves and the rocks they hit result in hugely varying shapes.

The effect that waves have depends on their strength. Strong, also called destructive waves occur on high energy beaches and are typical of Winter. They reduce the quantity of sediment present on the beach by carrying it out to bars under the sea. Constructive, weak waves are typical of low energy beaches and occur most during summer. They do the opposite to destructive waves and increase the size of the beach by piling sediment up onto the berm.

One of the most important transport mechanisms results from wave refraction. Since waves rarely break onto a shore at right angles, the upward movement of water

onto the beach (swash) occurs at an oblique angle. However, the return of water (backwash) is at right angles to the beach, resulting in the net movement of beach material laterally. This movement is known as beach drift. The endless cycle of swash and backwash and resulting beach drift can be observed on all beaches.

Probably the most important effect is longshore drift (LSD) (Also known as Littoral Drift), the process by which sediment is continuously moved along beaches by wave action. LSD occurs because waves hit the shore at an angle, pick up sediment (sand) on the shore and carry it down the beach at an angle (this is called swash). Due to gravity, the water then falls back perpendicular to the beach, dropping its sediment as it loses energy (this is called backwash). The sediment is then picked up by the next wave and pushed slightly further down the beach, resulting in a continual movement of sediment in one direction. This is the reason why long strips of coast are covered in sediment, not just the areas around river mouths, which are the main sources of beach sediment. LSD is reliant on a constant supply of sediment from rivers and if sediment supply is stopped or sediment falls into a submarine canals at any point along a beach, this can lead to bare beaches further along the shore.

LSD helps create many landforms including barriers, bay beaches and spits. In general LSD action serves to straighten the coast because the creation of barriers cuts off bays from the sea while sediment usually builds up in bays because the waves there are weaker (due to wave refraction), while sediment is carried away from the exposed headlands. The lack of sediment on headlands removes the protection of waves from them and makes them more vulnerable to weathering while the gathering of sediment in bays (where longshore drift is unable to remove it) protects the bays from further erosion and makes them pleasant recreational beaches.

Atmospheric Processes

- Onshore winds blowing “up” the beach, pick up sand and move it up the beach to form sand dunes.
- Rain hits the shore and erodes rocks, and carries weathered material to the shoreline to form beaches.
- Warm weather can encourage biological processes to occur more rapidly. In tropical areas some plants and animals protect stones from weathering, while other plants and animals actually eat away at the rocks.
- Temperatures that vary from below to above freezing point result in freeze-thaw weathering, while weather more than a few degrees below freezing point creates sea ice.

Biological Processes

In tropical regions in particular, plants and animals not only affect the weathering of rocks but are a source of sediment themselves. The shells and skeletons of many

organisms are of calcium carbonate and when this is broken down it forms sediment, limestone and clay.

Physical Processes

The main Physical Weathering process on beaches is salt-crystal growth. Wind carries salt spray onto rocks, where it is absorbed into small pores and cracks within the rocks. There the water evaporates and the salt crystallises, creating pressure and often breaking down the rock. In some beaches calcium carbonate is able to bind together other sediments to form beachrock and in warmer areas dunerock.

Because the sea level on earth regularly rises and falls due to climatic changes. During cold periods more of the Earth's water is stored as ice in glaciers while during warm periods it is released and sea levels rise to cover more land. Sea levels are currently quite high, while just 18,000 years ago during the Pleistocene ice age they were quite low. Global warming may result in further rises in the future, which presents a risk to coastal cities as most would be flooded by only small rises. As sea levels rise fjords and rias form. Fjords are flooded glacial valleys and rias are flooded river valleys. Fjords typically have steep rocky sides, while rias have dendritic drainage patterns typical of drainage zones. As tectonic plates move about the Earth they can rise and fall due to changing pressures and the presence of glaciers. If a beach is moving upwards relative to other plates this is known as isostatic change and raised beaches can be formed.

Land Level Changes (Isostatic Change)

This is found in the U.K. as above the line from the Wash to the Severn estuary, the land was covered in ice sheets during the last ice age. The weight of the ice caused northeast Scotland to sink, displacing the southeast and forcing it to rise. As the ice sheets receded the reverse process happened, as the land was released from the weight. At current estimates the southeast is sinking at a rate of about 2 mm per year, with northeast Scotland rising by the same amount.

Coastal Landforms

Spits

If the coast suddenly changes direction, especially around an estuary, spits are likely to form. Long shore drift pushes sediment along the beach but when it reaches a turn as in the diagram, the long shore drift does not always easily turn with it, especially near an estuary where the outward flow from a river may push sediment away from the coast. The area may be also be shielded from wave action, preventing much long shore drift. On the side of the headland receiving weaker waves, shingle and other large sediments will build up under the water where waves are not strong enough to move them along. This provides a good place for smaller sediments to build up to sea level. The sediment, after passing the headland will accumulate on the other

side and not continue down the beach, sheltered both by the headland and the shingle. Slowly over time sediment simply builds on this area, extending the spit outwards, forming a barrier of sand. Once in a while, the wind direction will change and come from the other direction. During this period the sediment will be pushed along in the other direction. The spit will start to grow backwards, forming a 'hook'. After this time the spit will grow again in the original direction. Eventually the spit will not be able to grow any further because it is no longer sufficiently sheltered from erosion by waves, or because the estuary current prevents sediment resting. Usually in the salty but calm waters behind the spit there will form a salt marshland. Spits often form around the breakwater of artificial harbours requiring dredging.

Occasionally, if there is no estuary then it is possible for the spit to grow across to the other side of the bay and form what is called a bar, or barrier. Barriers come in several varieties, but all form in a manner similar to spits. They usually enclose a bay to form a lagoon. They can join two headlands or join a headland to the mainland. When an island is joined to the mainland with a bar or barrier it is known as a tombolo. This usually occurs due to wave refraction, but can also be caused by isostatic change, a change in the level of the land (e.g. Chesil Beach). An example of this is along the Holderness coastline.

Oceanography

Oceanography, also called oceanology or marine science, is the branch of Earth science that studies the ocean. It covers a wide range of topics, including marine organisms and ecosystem dynamics; ocean currents, waves, and geophysical fluid dynamics; plate tectonics and the geology of the sea floor; and fluxes of various chemical substances and physical properties within the ocean and across its boundaries. These diverse topics reflect multiple disciplines that oceanographers blend to further knowledge of the world ocean and understanding of processes within it: biology, chemistry, geology, meteorology, and physics as well as geography.

History

Humans first acquired knowledge of the waves and currents of the seas and oceans in pre-historic times. Observations on tides are recorded by Aristotle and Strabo. Early modern exploration of the oceans was primarily for cartography and mainly limited to its surfaces and of the creatures that fishermen brought up in nets, though depth soundings by lead line were taken.

Although Juan Ponce de León in 1513 first identified the Gulf Stream, and the current was well-known to mariners, Benjamin Franklin made the first scientific study of it and gave it its name. Franklin measured water temperatures during several Atlantic crossings and correctly explained the Gulf Stream's cause. Franklin and Timothy Folger printed the first map of the Gulf Stream in 1769-1770.

When Louis Antoine de Bougainville, who voyaged between 1766 and 1769, and James Cook, who voyaged from 1768 to 1779, carried out their explorations in the South Pacific, information on the oceans themselves formed part of the reports. James Rennell wrote the first scientific textbooks about currents in the Atlantic and Indian oceans during the late 18th and at the beginning of 19th century. Sir James Clark Ross took the first modern sounding in deep sea in 1840, and Charles Darwin published a paper on reefs and the formation of atolls as a result of the second voyage of HMS *Beagle* in 1831-6. Robert FitzRoy published a report in four volumes of the three voyages of the *Beagle*. In 1841–1842 Edward Forbes undertook dredging in the Aegean Sea that founded marine ecology.

As first superintendent of the United States Naval Observatory (1842–1861) Matthew Fontaine Maury devoted his time to the study of marine meteorology, navigation, and charting prevailing winds and currents. His *Physical Geography of the Sea*, 1855 was the first textbook of oceanography. Many nations sent oceanographic observations to Maury at the Naval Observatory, where he and his colleagues evaluated the information and gave the results worldwide distribution.

The steep slope beyond the continental shelves was discovered in 1849. The first successful laying of transatlantic telegraph cable in August 1858 confirmed the presence of an underwater "telegraphic plateau" mid-ocean ridge. After the middle of the 19th century, scientific societies were processing a flood of new terrestrial botanical and zoological information.

In 1871, under the recommendations of the Royal Society of London, the British government sponsored an expedition to explore world's oceans and conduct scientific investigations. Under that sponsorship the Scots Charles Wyville Thompson and Sir John Murray launched the Challenger expedition (1872–1876). The results of this were published in 50 volumes covering biological, physical and geological aspects. 4417 new species were discovered.

Other European and American nations also sent out scientific expeditions (as did private individuals and institutions). The first purpose built oceanographic ship, the "Albatros" was built in 1882. The four-month 1910 North Atlantic expedition headed by Sir John Murray and Johan Hjort was at that time the most ambitious research oceanographic and marine zoological project ever, and led to the classic 1912 book *The Depths of the Ocean*.

Oceanographic institutes dedicated to the study of oceanography were founded. In the United States, these included the Scripps Institution of Oceanography in 1892, Woods Hole Oceanographic Institution in 1930, Virginia Institute of Marine Science in 1938, Lamont-Doherty Earth Observatory at Columbia University, and the School of Oceanography at University of Washington. In Britain, there is a major research institution: National Oceanography Centre, Southampton which is the successor to the Institute of Oceanography. In Australia, CSIRO Marine and Atmospheric Research,

known as CMAR, is a leading centre. In 1921 the International Hydrographic Bureau (IHB) was formed in Monaco.

In 1893, Fridtjof Nansen allowed his ship "Fram" to be frozen in the Arctic ice. As a result he was able to obtain oceanographic data as well as meteorological and astronomical data. The first international organization of oceanography was created in 1902 as the International Council for the Exploration of the Sea.

The first acoustic measurement of sea depth was made in 1914. Between 1925 and 1927 the "Meteor" expedition gathered 70,000 ocean depth measurements using an echo sounder, surveying the Mid atlantic ridge. The Great Global Rift, running along the Mid Atlantic Ridge, was discovered by Maurice Ewing and Bruce Heezen in 1953 while the mountain range under the Arctic was found in 1954 by the Arctic Institute of the USSR. The theory of seafloor spreading was developed in 1960 by Harry Hammond Hess. The Ocean Drilling Project started in 1966. Deep sea vents were discovered in 1977 by John Corlis and Robert Ballard in the submersible "Alvin".

In the 1950s, Auguste Piccard invented the bathyscaphe and used the "Trieste" to investigate the ocean's depths. The nuclear submarine Nautilus made the first journey under the ice to the North Pole in 1958. In 1962 there was the first deployment of FLIP (Floating Instrument Platform), a 355 foot spar buoy.

Then, in 1966, the U.S. Congress created a *National Council for Marine Resources and Engineering Development*. NOAA was put in charge of exploring and studying all aspects of Oceanography in the USA. It also enabled the National Science Foundation to award *Sea Grant College* funding to multi-disciplinary researchers in the field of oceanography.

From the 1970s, there has been much emphasis on the application of large scale computers to oceanography to allow numerical predictions of ocean conditions and as a part of overall environmental change prediction. An oceanographic buoy array was established in the Pacific to allow prediction of El Nino events.

1990 saw the start of the World Ocean Circulation Experiment (WOCE) which continued until 2002. Geosat seafloor mapping data became available in 1995.

In 1942, Sverdrup and Fleming published "The Ocean" which was a major landmark. "The Sea" (in three volumes covering physical oceanography, seawater and geology) edited by M.N. Hill was published in 1962 while the "Encyclopedia of Oceanography" by Rhodes Fairbridge was published in 1966.

Connection to the Atmosphere

The study of the oceans is linked to understanding global climate changes, potential global warming and related biosphere concerns. The atmosphere and ocean are linked because of evaporation and precipitation as well as thermal flux (and solar insolation). Wind stress is a major driver of ocean currents while the ocean is a sink

for atmospheric carbon dioxide. Our planet is invested with two great oceans; one visible, the other invisible; one underfoot, the other overhead; one entirely envelopes it, the other covers about two thirds of its surface. —Matthew F. Maury, *The Physical Geography of the Seas and Its Meteorology (1855).*

Branches

The study of oceanography is divided into branches:

- Biological oceanography, or marine biology, is the study of the plants, animals and microbes of the oceans and their ecological interaction with the ocean;
- Chemical oceanography, or marine chemistry, is the study of the chemistry of the ocean and its chemical interaction with the atmosphere;
- Geological oceanography, or marine geology, is the study of the geology of the ocean floor including plate tectonics and paleoceanography;
- Physical oceanography, or marine physics, studies the ocean's physical attributes including temperature-salinity structure, mixing, waves, internal waves, surface tides, internal tides, and currents.

These branches reflect the fact that many oceanographers are first trained in the exact sciences or mathematics and then focus on applying their interdisciplinary knowledge, skills and abilities to oceanography.

Data derived from the work of Oceanographers is used in marine engineering, in the design and building of oil platforms, ships, harbours, and other structures that allow us to use the ocean safely.

Oceanographic data management is the discipline ensuring that oceanographic data both past and present are available to researchers.

Quaternary Science

Quaternary science is an inter-disciplinary field of study focusing on the Quaternary period, which encompasses the last 2.6 million years. The field studies the last ice age and the recent interstadial the Holocene and uses proxy evidence to reconstruct the past environments during this period to infer the climatic and environmental changes that have occurred.

Landscape Ecology

Landscape ecology is the science of studying and improving relationships between urban development and ecological processes in the environment and particular ecosystems. This is done within a variety of landscape scales, development spatial patterns, and organizational levels of research and policy.

As a highly interdisciplinary science in systems ecology, landscape ecology integrates biophysical and analytical approaches with humanistic and holistic

perspectives across the natural sciences and social sciences. Landscapes are spatially heterogeneous geographic areas characterized by diverse interacting patches or ecosystems, ranging from relatively natural terrestrial and aquatic systems such as forests, grasslands and lakes to human-dominated environments including agricultural and urban settings. The most salient characteristics of landscape ecology are its emphasis on the relationship among pattern, process and scale and its focus on broad-scale ecological and environmental issues. These necessitate the coupling between biophysical and socioeconomic sciences. Key research topics in landscape ecology include ecological flows in landscape mosaics, land use and land cover change, scaling, relating landscape pattern analysis with ecological processes, and landscape conservation and sustainability (Wu & Hobbs 2002).

Terminology

The term *landscape ecology* was coined by Carl Troll, a German geographer, in 1939. He developed this terminology and many early concepts of landscape ecology as part of his early work, which consisted of applying aerial photograph interpretation to studies of interactions between environment and vegetation.

Explanation

Heterogeneity is the measure of how different parts of a landscape are from one another. Landscape ecology looks at how this spatial structure affects organism abundance at the landscape level, as well as the behaviour and functioning of the landscape as a whole. This includes studying the influence of pattern, or the internal order of a landscape, on process, or the continuous operation of functions of organisms. Landscape ecology also includes geomorphology as applied to the design and architecture of landscapes. Geomorphology is the study of how geological formations are responsible for the structure of a landscape.

History

Evolution of Theory

One central landscape ecology theory originated from MacArthur & Wilson's *The Theory of Island Biogeography*. This work considered the biodiversity on islands as the result of competing forces of colonization from a mainland stock and stochastic extinction. The concepts of island biogeography were generalized from physical islands to abstract patches of habitat by Levins' metapopulation model (which can be applied e.g. to forest islands in the agricultural landscape). This generalization spurred the growth of landscape ecology by providing conservation biologists a new tool to assess how habitat fragmentation affects population viability. Recent growth of landscape ecology owes much to the development of geographic information systems (GIS) and the availability of large-extent habitat data (e.g. remotely sensed datasets).

Development as a Discipline

Landscape ecology developed in Europe from historical planning on human-dominated landscapes. Concepts from general ecology theory were integrated in North America. While general ecology theory and its sub-disciplines focused on the study of more homogenous, discrete community units organized in a hierarchical structure (typically as ecosystems, populations, species, and communities), landscape ecology built upon heterogeneity in space and time. It frequently included human-caused landscape changes in theory and application of concepts.

By 1980, landscape ecology was a discrete, established, discipline. It was marked by the organization of the International Association for Landscape Ecology (IALE) in 1982. Landmark book publications defined the scope and goals of the discipline, including Naveh and Lieberman and Forman and Godron. Forman wrote that although study of "the ecology of spatial configuration at the human scale" was barely a decade old, there was strong potential for theory development and application of the conceptual framework. Today, theory and application of landscape ecology continues to develop through a need for innovative applications in a changing landscape and environment. Landscape ecology relies on advanced technologies such as remote sensing, GIS, and models. There has been associated development of powerful quantitative methods to examine the interactions of patterns and processes. An example would be determining the amount of carbon present in the soil based on landform over a landscape, derived from GIS maps, vegetation types, and rainfall data for a region.

Relationship to Ecological Theory

Landscape ecology theory may be slightly outside of the "classical and preferred domain of scientific disciplines" because of the large, heterogeneous areas of study. However, general ecology theory is central to landscape ecology theory in many aspects. Landscape ecology consists of four main principles: the development and dynamics of spatial heterogeneity, interactions and exchanges across heterogeneous landscapes, influences of spatial heterogeneity on biotic and abiotic processes, and the management of spatial heterogeneity. The main difference from traditional ecological studies, which frequently assume that systems are spatially homogenous, is the consideration of spatial patterns.

Important Terms in Landscape Ecology

Landscape ecology not only created new terms, but also incorporated existing ecological terms in new ways. Many of the terms used in landscape ecology are as interconnected and interrelated as the discipline itself. *Landscape* can be defined as an area containing two or more ecosystems in close proximity.

Scale and Heterogeneity (Incorporating Composition, Structure, and Function)

A main concept in landscape ecology is *scale*. Scale represents the real world as

translated onto a map, relating distance on a map image and the corresponding distance on earth. Scale is also the spatial or temporal measure of an object or a process, or amount of spatial resolution. Components of scale include composition, structure, and function, which are all important ecological concepts. Applied to landscape ecology, *composition* refers to the number of patch types represented on a landscape and their relative abundance. For example, the amount of forest or wetland, the length of forest edge, or the density of roads can be aspects of landscape composition. *Structure* is determined by the composition, the configuration, and the proportion of different patches across the landscape, while *function* refers to how each element in the landscape interacts based on its life cycle events. *Pattern* is the term for the contents and internal order of a heterogeneous area of land.

A landscape with structure and pattern implies that it has spatial *heterogeneity*, or the uneven distribution of objects across the landscape. Heterogeneity is a key element of landscape ecology that separates this discipline from other branches of ecology.

Patch and Mosaic

Patch, a term fundamental to landscape ecology, is defined as a relatively homogeneous area that differs from its surroundings. Patches are the basic unit of the landscape that change and fluctuate, a process called *patch dynamics*. Patches have a definite shape and spatial configuration, and can be described compositionally by internal variables such as number of trees, number of tree species, height of trees, or other similar measurements.

Matrix is the "background ecological system" of a landscape with a high degree of connectivity. *Connectivity* is the measure of how connected or spatially continuous a corridor, network, or matrix is. For example, a forested landscape (matrix) with fewer gaps in forest cover (open patches) will have higher connectivity. *Corridors* have important functions as strips of a particular type of landscape differing from adjacent land on both sides. A *network* is an interconnected system of corridors while *mosaic* describes the pattern of patches, corridors and matrix that form a landscape in its entirety.

Boundary and Edge

Landscape patches have a boundary between them which can be defined or fuzzy. The zone composed of the edges of adjacent ecosystems is the *boundary*. *Edge* means the portion of an ecosystem near its perimeter, where influences of the adjacent patches can cause an environmental difference between the interior of the patch and its edge. This edge effect includes a distinctive species composition or abundance. For example, when a landscape is a mosaic of perceptibly different types, such as a forest adjacent to a grassland, the edge is the location where the two types adjoin. In a continuous landscape, such as a forest giving way to open woodland, the exact

edge location is fuzzy and is sometimes determined by a local gradient exceeding a threshold, such as the point where the tree cover falls below thirty-five percent.

Ecotones, Ecoclines, and Ecotopes

A type of boundary is the *ecotone*, or the transitional zone between two communities. Ecotones can arise naturally, such as a lakeshore, or can be human-created, such as a cleared agricultural field from a forest. The ecotonal community retains characteristics of each bordering community and often contains species not found in the adjacent communities. Classic examples of ecotones include fencerows, forest to marshlands transitions, forest to grassland transitions, or land-water interfaces such as riparian zones in forests. Characteristics of ecotones include vegetational sharpness, physiognomic change, occurrence of a spatial community mosaic, many exotic species, ecotonal species, spatial mass effect, and species richness higher or lower than either side of the ecotone.

An *ecocline* is another type of landscape boundary, but it is a gradual and continuous change in environmental conditions of an ecosystem or community. Ecoclines help explain the distribution and diversity of organisms within a landscape because certain organisms survive better under certain conditions, which change along the ecocline. They contain heterogeneous communities which are considered more environmentally stable than those of ecotones.

An *ecotope* is a spatial term representing the smallest ecologically-distinct unit in mapping and classification of landscapes. Relatively homogeneous, they are spatially-explicit landscape units used to stratify landscapes into ecologically distinct features. They are useful for the measurement and mapping of landscape structure, function, and change over time, and to examine the effects of disturbance and fragmentation.

Disturbance and Fragmentation

Disturbance is an event that significantly alters the pattern of variation in the structure or function of a system. *Fragmentation* is the breaking up of a habitat, ecosystem, or land-use type into smaller parcels. Disturbance is generally considered a natural process. Fragmentation causes land transformation, an important process in landscapes as development occurs.

An important consequence of repeated, random clearing (whether by natural disturbance or human activity) is that contiguous cover can break down into isolated patches. This happens when the area cleared exceed a critical level, which means that landscapes exhibit two phases: connected and disconnected.

Landscape Ecology Theory

Landscape ecology theory stresses the role of human impacts on landscape structures and functions. It also proposes ways for restoring degraded landscapes. Landscape ecology explicitly includes humans as entities that cause functional changes

on the landscape. Landscape ecology theory includes the landscape stability principle, which emphasizes the importance of landscape structural heterogeneity in developing resistance to disturbances, recovery from disturbances, and promoting total system stability.

This principle is a major contribution to general ecological theories which highlight the importance of relationships among the various components of the landscape. Integrity of landscape components helps maintain resistance to external threats, including development and land transformation by human activity. Analysis of land use change has included a strongly geographical approach which has led to the acceptance of the idea of multifunctional properties of landscapes. There are still calls for a more unified theory of landscape ecology due to differences in professional opinion among ecologists and its interdisciplinary approach (Bastian 2001).

An important related theory is hierarchy theory, which refers to how systems of discrete functional elements operate when linked at two or more scales. For example, a forested landscape might be hierarchically composed of drainage basins, which in turn are composed of local ecosystems, which are in turn composed of individual trees and gaps. Recent theoretical developments in landscape ecology have emphasized the relationship between pattern and process, as well as the effect that changes in spatial scale has on the potential to extrapolate information across scales. Several studies suggest that the landscape has critical thresholds at which ecological processes will show dramatic changes, such as the complete transformation of a landscape by an invasive species with small changes in temperatures which favors the invasive's habitat requirements.

Landscape Ecology Application

Research Directions

Developments in landscape ecology illustrate the important relationships between spatial patterns and ecological processes. These developments incorporate quantitative methods that link spatial patterns and ecological processes at broad spatial and temporal scales. This linkage of time, space, and environmental change can assist managers in applying plans to solve environmental problems. The increased attention in recent years on spatial dynamics has highlighted the need for new quantitative methods that can analyse patterns, determine the importance of spatially explicit processes, and develop reliable models. Multivariate analysis techniques are frequently used to examine landscape level vegetation patterns. Studies use statistical techniques, such as cluster analysis, canonical correspondence analysis (CCA), or detrended correspondence analysis (DCA), for classifying vegetation. Gradient analysis is another way to determine the vegetation structure across a landscape or to help delineate critical wetland habitat for conservation or mitigation purposes (Choesin and Boerner 2002).

Climate change is another major component in structuring current research in landscape ecology. Ecotones, as a basic unit in landscape studies, may have significance for management under climate change scenarios, since change effects are likely to be seen at ecotones first because of the unstable nature of a fringe habitat. Research in northern regions has examined landscape ecological processes, such as the accumulation of snow, melting, freeze-thaw action, percolation, soil moisture variation, and temperature regimes through long-term measurements in Norway. The study analyses gradients across space and time between ecosystems of the central high mountains to determine relationships between distribution patterns of animals in their environment. Looking at where animals live, and how vegetation shifts over time, may provide insight into changes in snow and ice over long periods of time across the landscape as a whole.

Other landscape-scale studies maintain that human impact is likely the main determinant of landscape pattern over much of the globe. Landscapes may become substitutes for biodiversity measures because plant and animal composition differs between samples taken from sites within different landscape categories. Taxa, or different species, can "leak" from one habitat into another, which has implications for landscape ecology. As human land use practices expand and continue to increase the proportion of edges in landscapes, the effects of this leakage across edges on assemblage integrity may become more significant in conservation. This is because taxa may be conserved across landscape levels, if not at local levels.

Relationship to other Disciplines

Landscape ecology has been incorporated into a variety of ecological subdisciplines. For example, a recent development has been the more explicit consideration of spatial concepts and principles into the study of lakes, streams and wetlands in the field of landscape limnology. In addition, landscape ecology has important links to application-oriented disciplines such as agriculture and forestry. In agriculture, landscape ecology has introduced new options for the management of environmental threats brought about by the intensification of agricultural practices.

Agriculture has always been a strong human impact on ecosystems. In forestry, from structuring stands for fuelwood and timber to ordering stands across landscapes to enhance aesthetics, consumer needs have affected conservation and use of forested landscapes. Landscape forestry provides methods, concepts, and analytic procedures for landscape forestry. Finally, landscape ecology has been cited as a contributor to the development of fisheries biology as a distinct biological science discipline, and is frequently incorporated in study design for wetland delineation in hydrology.

Geomatics (also known as geospatial technology or geomatics engineering) is the discipline of gathering, storing, processing, and delivering geographic information, or spatially referenced information.

Overview and Etymology

Geomatics is relatively new as a scientific term. It was coined by *Dubuisson* in the year 1969 with the idea of combining the terms geodesy and geoinformatics. It includes the tools and techniques used in land surveying, remote sensing, cartography, geographic information systems (GIS), global navigation satellite systems (GPS, GLONASS, Galileo, Compass), photogrammetry, geography and related forms of earth mapping. It was originally used in Canada, because it is similar in origin to both French and English. The term geomatics has since been adopted by the International Organization for Standardization, the Royal Institution of Chartered Surveyors, and many other international authorities, although some (especially in the United States) have shown a preference for the term *geospatial technology*.

Similarly, the new related field hydrogeomatics covers the geomatics area associated with surveying work carried out on, above or below the surface of the sea or other areas of water. The older term of hydrographics was too specific to the preparation of marine charts and failed to include the broader concept of positioning or measurements in all marine environments. A geospatial network is a network of collaborating resources for sharing and coordinating geographical data, and data tied to geographical references. One example of such a network is the Open Geospatial Consortium's efforts to provide "ready global access to geographic information". A number of university departments which were once titled surveying, survey engineering or topographic science have re-titled themselves as geomatics or geomatic engineering.

The rapid progress, and increased visibility, of geomatics since 1990s has been made possible by advances in computer hardware, computer science, and software engineering, as well as airborne and space observation remote sensing technologies.

Science of deriving information about an object using sensor without physically contacting it is called remote sensing which is a part of geomatics.

Applications

Applications areas include:

- Air navigation services
- Archaeological excavation and survey for GIS applications
- Coastal zone management and mapping
- Criminology
- Disaster informatics for disaster risk reduction and response
- The environment
- Infrastructure management
- Land management and reform

- Natural resource monitoring and development
- Seismic Interpretation
- Sociology
- Subdivision planning
- Urban planning

Integrated Geography

Integrated geography is the branch of geography that describes the spatial aspects of interactions between humans and the natural world. It requires an understanding of the dynamics of geology, meteorology, hydrology, biogeography, ecology, and geomorphology, as well as the ways in which human societies conceptualize the environment.

The links between cultural and physical geography were once more readily apparent than they are today. As human experience of the world is increasingly mediated by technology, the relationships have often become obscured.

Integrated geography represents a critically important set of analytical tools for assessing the impact of human presence on the environment by measuring the result of human activity on natural landforms and cycles.

Integrated geography is the third branch of geography, as compared to physical and human geography. Integrated geography concentrates on the relationship between human and the surrounding world.

Journals and Literature

Physical geography and Earth Science journals communicate and document the results of research carried out in universities and various other research institutions. Most journals cover a specific field and publish the research within that field, however unlike human geographers, physical geographers tend to publish in inter-disciplinary journals rather than predominantly geography journal; the research is normally expressed in the form of a scientific paper. Additionally, textbooks, books, and magazines on geography communicate research to laypeople, although these tend to focus on environmental issues or cultural dilemmas.

Historical Evolution of the Discipline

From the birth of geography as a science during the Greek classical period and until the late nineteenth century with the birth of anthropogeography or Human Geography, Geography was almost exclusively a natural science: the study of location and descriptive gazetteer of all places of the known world. Several works among the best known during this long period could be cited as an example, from Strabo (Geography), Eratosthenes (Geography) or Dionisio Periegetes (Periegesis Oiceumene) in the Ancient Age to the Alexander von Humboldt (Cosmos) in the century XIX, in

which geography is regarded as a physical and natural science, of course, through the work Summa de Geografía of Martín Fernández de Enciso from the early sixteenth century, which is indicated for the first time the New World.

During the eighteenth and nineteenth centuries, a controversy exported from Geology, between supporters of James Hutton (uniformitarianism Thesis) and Georges Cuvier (catastrophism) strongly influenced the field of geography, because geography at this time was a natural science since Human Geography or Antropogeography had just developed as a discipline in the late nineteenth century.

Two historical events during the nineteenth century had a great effect in the further development of physical geography. The first was the European colonial expansion in Asia, Africa, Australia and even America in search of raw materials required by industries during the Industrial Revolution. This fostered the creation of geography departments in the universities of the colonial powers and the birth and development of national geographical societies, thus giving rise to the process identified by Horacio Capel as the institutionalization of geography.

One of the most prolific empires in this regard was the Russian. A mid-eighteenth century many geographers are sent by the Russian altamirazgo different opportunities to perform geographical surveys in the area of Arctic Siberia. Among these is who is considered the patriarch of Russian geography: Mikhail Lomonosov who in the mid-1750s began working in the Department of Geography, Academy of Sciences to conduct research in Siberia, their contributions are notable in this regard, shows the soil organic origin, develops a comprehensive law on the movement of the ice that still governs the basics, thereby founding a new branch of Geography: Glaciology. In 1755 his initiative was founded Moscow University where he promotes the study of geography and the training of geographers. In 1758 he was appointed director of the Department of Geography, Academy of Sciences, a post from which would develop a working methodology for geographical survey guided by the most important long expeditions and geographical studies in Russia. Thus followed the line of Lomonosov and the contributions of the Russian school became more frequent through his disciples, and in the nineteenth century we have great geographers as Vasily Dokuchaev who performed works of great importance as a "principle of comprehensive analysis of the territory" and "Russian Chernozem" latter being the most important where introduces the geographical concept of soil, as distinct from a simple geological strata, and thus founding a new geographic area of study: the Pedology. Climatology also receive a strong boost from the Russian school by Wladimir Koppen whose main contribution, climate classification, is still valid today. However, this great geographer also contributed to the Paleogeography through his work "The climates of the geological past" which is considered the father of Paleoclimatology. Russian geographers who made great contributions to the discipline in this period were: NM Sibirtsev, Pyotr Semyonov, K. D. Glinka, Neustrayev, among others.

The second important process is the theory of evolution by Darwin in mid-century (which decisively influenced the work of Ratzel, who had academic training as a zoologist and was a follower of Darwin's ideas) which meant an important impetus in the development of Biogeography.

Another major event in the late nineteenth and early twentieth century will give a major boost to development of geography and will take place in United States. It is the work of the famous geographer William Morris Davis who not only made important contributions to the establishment of discipline in his country, but revolutionized the field to develop geographical cycle theory which he proposed as a paradigm for Geography in general, although in actually served as a paradigm for Physical Geography. His theory explained that mountains and other landforms are shaped by the influence of a number of factors that are manifested in the geographical cycle. He explained that the cycle begins with the lifting of the relief by geological processes (faults, volcanism, tectonic upheaval, etc.).. Geographical factors such as rivers and runoff begins to create the V-shaped valleys between the mountains (the stage called "youth"). During this first stage, the terrain is steeper and more irregular. Over time, the currents can carve wider valleys ("maturity") and then start to wind, towering hills only ("senescence"). Finally, everything comes to what is a plain flat plain at the lowest elevation possible (called "baseline") This plain was called by Davis' "peneplain" meaning "almost plain" Then the rejuvenation occurs and there is another mountain lift and the cycle continues. Although Davis's theory is not entirely accurate, it was absolutely revolutionary and unique in its time and helped to modernize and create Geography subfield of Geomorphology. Its implications prompted a myriad of research in various branches of Physical Geography. In the case of the Paleogeography this theory provided a model for understanding the evolution of the landscape. For Hydrology, Glaciology and Climatology as a boost investigated as studying geographic factors shape the landscape and affect the cycle. The bulk of the work of William Morris Davis led to the development of a new branch of Physical Geography: Geomorphology whose contents until then did not differ from the rest of Geography. Shortly after this branch would present a major development. Some of his disciples made significant contributions to various branches of physical geography such as Curtis Marbut and his invaluable legacy for Pedology, Mark Jefferson, Isaiah Bowman, among others.

Notable Physical Geographers

- Eratosthenes (276 – 194 BC), who made the first known reliable estimation of the Earth's size. He is considered the father of geodesy.
- Ptolemy (c.90 – c.168), who compiled Greek and Roman knowledge to produce the book *Geographia*.
- Abû Rayhân Bîrûnî (973 – 1048 AD), considered the father of geodesy.

- Ibn Sina (Avicenna, 980–1037), who formulated the law of superposition and concept of uniformitarianism in *The Book of Healing*.
- Muhammad al-Idrisi (Dreses, 1100 – c.1165), who drew the *Tabula Rogeriana*, the most accurate world map in pre-modern times.
- Piri Reis (1465 – c.1554), whose Piri Reis map is the oldest surviving world map to include the Americas and possibly Antarctica.
- Gerardus Mercator (1512–1594), an innovative cartographer and originator of the Mercator projection.
- Bernhardus Varenius (1622–1650), Wrote his important work "General Geography" (1650), first overview of the geography, the foundation of modern geography.
- Mikhail Lomonosov (1711–1765), father of Russian geography and founded the study of glaciology.
- Alexander Von Humboldt (1769–1859), considered the father of modern geography. Published *Kosmos* and founded the study of biogeography.
- Arnold Henry Guyot (1807–1884), who noted the structure of glaciers and advanced the understanding of glacial motion, especially in fast ice flow.
- Louis Agassiz (1807–1873), the author of a glacial theory which disputed the notion of a steady-cooling Earth.
- Alfred Russel Wallace (1823–1913), founder of modern biogeography and the Wallace line.
- Vasily Dokuchaev (1840–1903), patriach of Russian geography and founder of pedology.
- Wladimir Peter Koppen (1846–1940), developer of most important climate classification and founder of Paleoclimatology.
- William Morris Davis (1850–1934), father of American geography, founder of Geomorphology and developer of the geographical cycle theory.
- Walther Penck (1888–1923), proponent of the cycle of erosion and the simultaneous occurrence of uplift and denudation.
- Sir Ernest Shackleton (1874–1922), Antarctic explorer during the Heroic Age of Antarctic Exploration.
- Robert E. Horton (1875–1945), founder of modern hydrology and concepts such as infiltration capacity and overland flow.
- J Harlen Bretz (1882–1981), pioneer of research into the shaping of landscapes by catastrophic floods, most notably the Bretz (Missoula) floods.
- Luis García Sáinz (1894–1965), was a pioneer of physical geography in Spain.
- Willi Dansgaard (1922–2011), palaeoclimatologist and quaternary scientist,

instrumental in the use of oxygen-isotope dating and co-identifier of Dansgaard-Oeschger events.

- Hans Oeschger (1927–1998), palaeoclimatologist and pioneer in ice core research, co-identifier of Dansgaard-Orschger events.
- Richard Chorley (1927–2002), a key contributor to the quantitative revolution and the use of systems theory in geography.
- Sir Nicholas Shackleton (1937–2006), who demonstrated that oscillations in climate over the past few million years could be correlated with variations in the orbital and positional relationship between the Earth and the Sun.
- Stefan Rahmstorf (born 1960), professor of abrupt climate changes and author on theories of thermohaline dynamics.

Human Geography

Human geography is one of the two major sub-fields of the discipline of geography. Human geography is the study of the world, its people, communities, and cultures. Human geography differs from physical geography mainly in that it has a greater focus on studying human activities and is more receptive to qualitative research methodologies. Broadly speaking, human geography is a social science discipline, while physical geography is an earth science. Human geography is concerned with the study of spatial patterns of interactions between human beings and their physical environment.

History

Geographical knowledge, both physical and social, has a long history. In the History of geography, geographers have often recorded and described features of the Earth that might now be considered the remit of human, rather than physical, geographers. For example Hecataeus of Miletus, a geographer and historian in ancient Greece, described inhabitants of the ancient world as well as physical features.

It was not until the 18th and 19th Centuries, however, that geography was recognised as a formal academic discipline. The Royal Geographical Society was founded in England in 1830, although the United Kingdom did not get its first full Chair of geography until 1917. The first real geographical intellect to emerge in United Kingdom geography was Halford John Mackinder, appointed reader at Oxford University in 1887. The National Geographic Society was founded in the USA in 1888 and began publication of the *National Geographic* magazine which became and continues to be a great popularizer of geographic information. The society has long supported geographic research and education.

One of the first examples of geographic methods being used for purposes other than to describe and theorise the physical properties of the earth is John Snow's map of the 1854 Broad Street cholera outbreak. Though a physician and a pioneer of

epidemiology, the map is probably one of the earliest examples of Health geography. The now fairly distinct differences between the subfields of physical and human geography developed at a later date.

This connection between both physical and human properties of geography is most apparent in the theory of Environmental determinism, made popular in the 19th Century by Carl Ritter and others, and with close links to evolutionary biology of the time. Environmental determinism is the theory that a people's physical, mental and moral habits are directly due to the influence of their natural environment. However, by the mid 19th Century, environmental determinism was under attack for lacking methodological rigour associated with modern science, and later as serving to justify racism and imperialism.

A similar concern with both human and physical aspects is apparent in the later Regional geography, during the later 19th and first half of the 20th Centuries. The goal of regional geography, through regionalization, was to delineate space into regions and then understand and describe the unique characteristics of each region, in both human and physical aspects. With links to possibilism and cultural ecology, some of the same notions of causal effect of the environment on society and culture, as with environmental determinism remained.

By the 1960s, however, the quantitative revolution lead to strong criticism of regional geography. Due to a perceived lack of scientific rigour in and overly descriptive nature of the discipline, and a continued separation of geography from geology and the two subfields of physical and human geography, geographers in the mid 20th Century began to apply statistical and mathematical model methods to solving spatial problems.

Much of the development during the quantitative revolution is now apparent in the use of Geographic information systems; the use of statistics, spatial modelling and positivist approaches is still important to many branches of human geography. Well-known geographers from this period are Fred K. Schaefer, Waldo Tobler, William Garrison, Peter Haggett, Richard J. Chorley, William Bunge, and Torsten Hägerstrand.

From the 1970s a number of critiques of the positivism now associated with geography emerged. Known under the term critical geography this signalled another turning point in the discipline. Behavioral geography emerged for some time as a means to understand how people made perceived spaces and places, and made locational decisions. More influentially, radical geography emerged in the 1970s and 1980s, drawing heavily on Marxist theory and techniques, and is associated with geographers such as David Harvey and Richard Peet. Seeking to say something 'meaningful' about the problems recognised through quantitative methods, to provide explanations rather than descriptions, to put forward alternatives and solutions and to be politically engaged, rather than the detachment associated with positivist methods. (The detachment and objectivity of the quantitative revolution was itself critiqued by

radical geographers as being a tool of capital). Radical geography and the links to Marxism and related theories remain an important part of contemporary human geography Critical geography also saw the introduction of humanistic geography, associated with the work of Yi-Fu Tuan, which, though similar to behavioural geography, pushed for a much more qualitative approach in methodology.

The changes under critical geography have led to contemporary approaches in the discipline such as Feminist geography, New cultural geography, and the engagement with postmodern and poststructural theories and philosophies.

Fields of Human Geography

The main fields of study in human geography focus around the core fields of:

Culture

Cultural geography is the study of cultural products and norms and their variation across and relations to spaces and places. It focuses on describing and analyzing the ways language, religion, economy, government, and other cultural phenomena vary or remain constant from one place to another and on explaining how humans function spatially.

- Subfields include: *Children's geographies, Animal geographies, Language geography, Sexuality and Space and Religion geography.*

Development

Development Geography is the study of the Earth's geography with reference to the Standard of living and the Quality of life of its human inhabitants, study of the location, distribution and spatial organization of economic activities, across the Earth. The subject matter investigated is strongly influenced by the researcher's methodological approach.

Economic

Economic geography examines relationships between human economic systems, states, and other factors, and the biophysical environment.

- Subfields include *Marketing geography.*

Health

Health geography is the application of geographical information, perspectives, and methods to the study of health, disease, and health care.

Historical

Historical Geography is the study of the human, physical, fictional, theoretical, and "real" geographies of the past. Historical geography studies a wide variety of issues and topics. A common theme is the study of the geographies of the past and

how a place or region changes through time. Many historical geographers study geographical patterns through time, including how people have interacted with their environment, and created the cultural landscape.

- Subfields include *Time geography.*

Political

Political geography is concerned with the study of both the spatially uneven outcomes of political processes and the ways in which political processes are themselves affected by spatial structures.

- Subfields include *Electoral geography*, *Geopolitics*, *Strategic geography* and *Military geography.*

Population

Population geography is the study of the ways in which spatial variations in the distribution, composition, migration, and growth of populations are related to the nature of places.

Tourism

Tourism geography is the study of travel and tourism as an industry, as a human activity, and especially as a place-based experience.

- Subfields include *Transportation geography.*

Urban

Urban geography is the study of urban areas with specific regards to spatial and relational aspects and theories. That is the study of areas which have a high concentration of buildings and infrastructure. These are areas where the majority of economic activities are in the secondary sector and tertiary sectors. They probably have a high population density.

Philosophical & Theoretical Approaches

Within each of the subfields, various philosophical approaches can be used in research; therefore, an urban geographer could be a Feminist or Marxist geographer, etc.

Such approaches are:

- Behavioral geography
- Critical geography
- Feminist geography
- Marxist geography
- Non-representational theory
- Postcolonialism

- Poststructuralist geography
- Psychoanalytic geography

Integrated Geography

Integrated geography is the branch of geography that describes the spatial aspects of interactions between humans and the natural world. It requires an understanding of the dynamics of geology, meteorology, hydrology, biogeography, ecology, and geomorphology, as well as the ways in which human societies conceptualize the environment.

The links between cultural and physical geography were once more readily apparent than they are today. As human experience of the world is increasingly mediated by technology, the relationships have often become obscured.

Integrated geography represents a critically important set of analytical tools for assessing the impact of human presence on the environment by measuring the result of human activity on natural landforms and cycles.

Integrated geography is the third branch of geography, as compared to physical and human geography. Integrated geography concentrates on the relationship between human and the surrounding world.

Regional Geography

Regional geography is the study of world regions. Attention is paid to unique characteristics of a particular region such as natural elements, human elements, and regionalization which covers the techniques of delineating space into regions.

Regional geography is also a certain approach to geographical study, comparable to quantitative geography or critical geography. This approach prevailed during the second half of the 19th century and the first half of the 20th century, a period when then regional geography paradigm was central within the geographical sciences. It was later criticised for its descriptiveness and the lack of theory. Strong criticism was leveled against it in particular during the 1950s and the quantitative revolution. Main critics were G. H. T. Kimble and Fred K. Schaefer.

The regional geography paradigm has had an impact on many other geographical sciences, including economic geography and geomorphology. Regional geography is still taught in some universities as a study of the major regions of the world, such as Northern and Latin America, Europe, and Asia and their countries. In addition, the notion of a city-regional approach to the study of geography gained some credence in the mid-1990s through the work of geographers such as Saskia Sassen, although it was also criticized, for example by Peter Storper.

Notable figures in regional geography were Alfred Hettner in Germany, with his concept of chorology; Paul Vidal de la Blache in France, with the possibilism approach

(possibilism being a softer notion than environmental determinism); and, in the United States, Richard Hartshorne with his concept of areal differentiation.

Some geographers have also attempted to reintroduce a certain amount of regionalism since the 1980s. This involves a complex definition of regions and their interactions with other scales.

Geomatics

Geomatics (also known as geospatial technology or geomatics engineering) is the discipline of gathering, storing, processing, and delivering geographic information, or spatially referenced information.

Overview and Etymology

Geomatics is relatively new as a scientific term. It was coined by *Dubuisson* in the year 1969 with the idea of combining the terms geodesy and geoinformatics. It includes the tools and techniques used in land surveying, remote sensing, cartography, geographic information systems (GIS), global navigation satellite systems (GPS, GLONASS, Galileo, Compass), photogrammetry, geography and related forms of earth mapping.

It was originally used in Canada, because it is similar in origin to both French and English. The term geomatics has since been adopted by the International Organization for Standardization, the Royal Institution of Chartered Surveyors, and many other international authorities, although some (especially in the United States) have shown a preference for the term *geospatial technology*.

Similarly, the new related field hydrogeomatics covers the geomatics area associated with surveying work carried out on, above or below the surface of the sea or other areas of water. The older term of hydrographics was too specific to the preparation of marine charts and failed to include the broader concept of positioning or measurements in all marine environments. A geospatial network is a network of collaborating resources for sharing and coordinating geographical data, and data tied to geographical references. One example of such a network is the Open Geospatial Consortium's efforts to provide "ready global access to geographic information". A number of university departments which were once titled surveying, survey engineering or topographic science have re-titled themselves as geomatics or geomatic engineering.

The rapid progress, and increased visibility, of geomatics since 1990s has been made possible by advances in computer hardware, computer science, and software engineering, as well as airborne and space observation remote sensing technologies.

Science of deriving information about an object using sensor without physically contacting it is called remote sensing which is a part of geomatics.

Applications

Applications areas include:

- Air navigation services
- Archaeological excavation and survey for GIS applications
- Coastal zone management and mapping
- Criminology
- Disaster informatics for disaster risk reduction and response
- The environment
- Infrastructure management
- Land management and reform
- Natural resource monitoring and development
- Seismic Interpretation
- Sociology
- Subdivision planning
- Urban planning

Related Fields

- Urban planning, regional planning and spatial planning: use the science of geography to assist in determining how to develop (or not develop) the land to meet particular criteria, such as safety, beauty, economic opportunities, the preservation of the built or natural heritage, and so on. The planning of towns, cities, and rural areas may be seen as applied geography.
- Regional science: In the 1950s the regional science movement led by Walter Isard arose, to provide a more quantitative and analytical base to geographical questions, in contrast to the descriptive tendencies of traditional geography programs. Regional science comprises the body of knowledge in which the spatial dimension plays a fundamental role, such as regional economics, resource management, location theory, urban and regional planning, transport and communication, human geography, population distribution, landscape ecology, and environmental quality.
- Interplanetary Sciences: While the discipline of geography is normally concerned with the Earth, the term can also be informally used to describe the study of other worlds, such as the planets of the Solar System and even beyond. The study of systems larger than the earth itself usually forms part of Astronomy or Cosmology. The study of other planets is usually called planetary science. Alternative terms such as Areology (the study of Mars) have been proposed, but are not widely used.

Techniques

As spatial interrelationships are key to this synoptic science, maps are a key tool. Classical cartography has been joined by a more modern approach to geographical analysis, computer-based geographic information systems (GIS).

In their study, geographers use four interrelated approaches:

- Systematic — Groups geographical knowledge into categories that can be explored globally.
- Regional — Examines systematic relationships between categories for a specific region or location on the planet.
- Descriptive — Simply specifies the locations of features and populations.
- Analytical — Asks *why* we find features and populations in a specific geographic area.

Cartography

Cartography studies the representation of the Earth's surface with abstract symbols (map making). Although other subdisciplines of geography rely on maps for presenting their analyses, the actual making of maps is abstract enough to be regarded separately. Cartography has grown from a collection of drafting techniques into an actual science.

Cartographers must learn cognitive psychology and ergonomics to understand which symbols convey information about the Earth most effectively, and behavioral psychology to induce the readers of their maps to act on the information. They must learn geodesy and fairly advanced mathematics to understand how the shape of the Earth affects the distortion of map symbols projected onto a flat surface for viewing. It can be said, without much controversy, that cartography is the seed from which the larger field of geography grew. Most geographers will cite a childhood fascination with maps as an early sign they would end up in the field.

Geographic Information Systems

Geographic information systems (GIS) deal with the storage of information about the Earth for automatic retrieval by a computer, in an accurate manner appropriate to the information's purpose. In addition to all of the other subdisciplines of geography, GIS specialists must understand computer science and database systems. GIS has revolutionized the field of cartography; nearly all mapmaking is now done with the assistance of some form of GIS software. GIS also refers to the science of using GIS software and GIS techniques to represent, analyse and predict spatial relationships. In this context, GIS stands for Geographic Information Science.

Remote Sensing

Remote sensing is the science of obtaining information about Earth features from measurements made at a distance. Remotely sensed data comes in many forms such

as satellite imagery, aerial photography and data obtained from hand-held sensors. Geographers increasingly use remotely sensed data to obtain information about the Earth's land surface, ocean and atmosphere because it: a) supplies objective information at a variety of spatial scales (local to global), b) provides a synoptic view of the area of interest, c) allows access to distant and / or inaccessible sites, d) provides spectral information outside the visible portion of the electromagnetic spectrum, and e) facilitates studies of how features / areas change over time. Remotely sensed data may be analysed either independently of, or in conjunction with, other digital data layers (e.g., in a Geographic Information System).

Quantitative Methods

Geostatistics deal with quantitative data analysis, specifically the application of statistical methodology to the exploration of geographic phenomena. Geostatistics is used extensively in a variety of fields including: hydrology, geology, petroleum exploration, weather analysis, urban planning, logistics, and epidemiology.

The mathematical basis for geostatistics derives from cluster analysis, linear discriminant analysis and non-parametric statistical tests, and a variety of other subjects. Applications of geostatistics rely heavily on geographic information systems, particularly for the interpolation (estimate) of unmeasured points. Geographers are making notable contributions to the method of quantitative techniques.

Qualitative Methods

Geographic qualitative methods, or ethnographical; research techniques, are used by human geographers. In cultural geography there is a tradition of employing qualitative research techniques also used in anthropology and sociology. Participant observation and in-depth interviews provide human geographers with qualitative data.

History

The oldest known world maps date back to ancient Babylon from the 9th century BC. The best known Babylonian world map, however, is the *Imago Mundi* of 600 BC. The map as reconstructed by Eckhard Unger shows Babylon on the Euphrates, surrounded by a circular landmass showing Assyria, Urartu and several cities, in turn surrounded by a "bitter river" (Oceanus), with seven islands arranged around it so as to form a seven-pointed star. The accompanying text mentions seven outer regions beyond the encircling ocean. The descriptions of five of them have survived. In contrast to the *Imago Mundi*, an earlier Babylonian world map dating back to the 9th century BC depicted Babylon as being further north from the centre of the world, though it is not certain what that centre was supposed to represent.

The ideas of Anaximander (c. 610 BC-c. 545 BC), considered by later Greek writers to be the true founder of geography, come to us through fragments quoted by his

successors. Anaximander is credited with the invention of the gnomon,the simple yet efficient Greek instrument that allowed the early measurement of latitude. Thales, Anaximander is also credited with the prediction of eclipses. The foundations of geography can be traced to the ancient cultures, such as the ancient, medieval, and early modern Chinese.

The Greeks, who were the first to explore geography as both art and science, achieved this through Cartography, Philosophy, and Literature, or through Mathematics. There is some debate about who was the first person to assert that the Earth is spherical in shape, with the credit going either to Parmenides or Pythagoras. Anaxagoras was able to demonstrate that the profile of the Earth was circular by explaining eclipses. However, he still believed that the Earth was a flat disk, as did many of his contemporaries. One of the first estimates of the radius of the Earth was made by Eratosthenes.

The first rigorous system of latitude and longitude lines is credited to Hipparchus. He employed a sexagesimal system that was derived from Babylonian mathematics. The parallels and meridians were sub-divided into 360°, with each degree further subdivided 602 (minutes). To measure the longitude at different location on Earth, he suggested using eclipses to determine the relative difference in time. The extensive mapping by the Romans as they explored new lands would later provide a high level of information for Ptolemy to construct detailed atlases. He extended the work of Hipparchus, using a grid system on his maps and adopting a length of 56.5 miles for a degree.

From the 3rd century onwards, Chinese methods of geographical study and writing of geographical literature became much more complex than what was found in Europe at the time (until the 13th century). Chinese geographers such as Liu An, Pei Xiu, Jia Dan, Shen Kuo, Fan Chengda, Zhou Daguan, and Xu Xiake wrote important treatises, yet by the 17th century, advanced ideas and methods of Western-style geography were adopted in China.

During the Middle Ages, the fall of the Roman empire led to a shift in the evolution of geography from Europe to the Islamic world. Muslim geographers such as Muhammad al-Idrisi produced detailed world maps (such as Tabula Rogeriana), while other geographers such as Yaqut al-Hamawi, Abu Rayhan Biruni, Ibn Battuta and Ibn Khaldun provided detailed accounts of their journeys and the geography of the regions they visited.

Turkish geographer, Mahmud al-Kashgari drew a world map on a linguistic basis, and later so did Piri Reis. Further, Islamic scholars translated and interpreted the earlier works of the Romans and Greeks and established the House of Wisdom in Baghdad for this purpose. Abû Zayd al-Balkhî, originally from Balkh, founded the "Balkhî school" of terrestrial mapping in Baghdad. Suhrâb, a late tenth century Muslim geographer, accompanied a book of geographical coordinates with instructions

for making a rectangular world map, with equirectangular projection or cylindrical equidistant projection.

Abu Rayhan Biruni (976-1048) first described a polar equi-azimuthal equidistant projection of the celestial sphere. He was regarded as the most skilled when it came to mapping cities and measuring the distances between them, which he did for many cities in the Middle East and Indian subcontinent. He often combined astronomical readings and mathematical equations, in order to develop methods of pin-pointing locations by recording degrees of latitude and longitude. He also developed similar techniques when it came to measuring the heights of mountains, depths of valleys, and expanse of the horizon.

He also discussed human geography and the planetary habitability of the Earth. He also calculated the latitude of Kath, Khwarezm, using the maximum altitude of the Sun, and solved a complex geodesic equation in order to accurately compute the Earth's circumference, which were close to modern values of the Earth's circumference. His estimate of 6,339.9 km for the Earth radius was only 16.8 km less than the modern value of 6,356.7 km. In contrast to his predecessors who measured the Earth's circumference by sighting the Sun simultaneously from two different locations, al-Biruni developed a new method of using trigonometric calculations based on the angle between a plain and mountain top which yielded more accurate measurements of the Earth's circumference and made it possible for it to be measured by a single person from a single location.

The European Age of Discovery during the 16th and 17th centuries, where many new lands were discovered and accounts by European explorers such as Christopher Columbus, Marco Polo and James Cook, revived a desire for both accurate geographic detail, and more solid theoretical foundations in Europe. The problem facing both explorers and geographers was finding the latitude and longitude of a geographic location. The problem of latitude was solved long ago but that of longitude remained; agreeing on what zero meridian should be was only part of the problem. It was left to John Harrison to solve it by inventing the chronometer H-4, in 1760, and later in 1884 for the International Meridian Conference to adopt by convention the Greenwich meridian as zero meridian.

The 18th and 19th centuries were the times when geography became recognized as a discrete academic discipline and became part of a typical university curriculum in Europe (especially Paris and Berlin). The development of many geographic societies also occurred during the 19th century with the foundations of the Société de Géographie in 1821, the Royal Geographical Society in 1830, Russian Geographical Society in 1845, American Geographical Society in 1851, and the National Geographic Society in 1888. The influence of Immanuel Kant, Alexander von Humboldt, Carl Ritter and Paul Vidal de la Blache can be seen as a major turning point in geography from a philosophy to an academic subject.

Over the past two centuries the advancements in technology such as computers, have led to the development of geomatics and new practices such as participant observation and geostatistics being incorporated into geography's portfolio of tools. In the West during the 20th century, the discipline of geography went through four major phases: environmental determinism, regional geography, the quantitative revolution, and critical geography.

The strong interdisciplinary links between geography and the sciences of geology and botany, as well as economics, sociology and demographics have also grown greatly especially as a result of Earth System Science that seeks to understand the world in a holistic view.

Notable Geographers

- Eratosthenes (276BC-194BC)-calculated the size of the Earth.
- Ptolemy (c.90–c.168)-compiled Greek and Roman knowledge into the book Geographia.
- Al Idrisi (1100–1165/66)-author of Nuzhatul Mushtaq.
- Gerardus Mercator (1512–1594)-innovative cartographer produced the mercator projection.
- Alexander von Humboldt (1769–1859)-Considered Father of modern geography, published the Kosmos and founder of the sub-field biogeography.
- Carl Ritter (1779–1859)-Considered Father of modern geography. Occupied the first chair of geography at Berlin University.
- Arnold Henry Guyot (1807–1884)-noted the structure of glaciers and advanced understanding in glacier motion, especially in fast ice flow.
- William Morris Davis (1850–1934)-father of American geography and developer of the cycle of erosion.
- Paul Vidal de la Blache (1845–1918)-founder of the French school of geopolitics and wrote the principles of human geography.
- Sir Halford John Mackinder (1861–1947)-Co-founder of the LSE, Geographical Association.
- Carl O. Sauer (1889–1975)-Prominent cultural geographer.
- Walter Christaller (1893–1969)-human geographer and inventor of Central place theory.
- Yi-Fu Tuan (1930-)-Chinese-American scholar credited with starting Humanistic Geography as a discipline.
- David Harvey (1935-)-Marxist geographer and author of theories on spatial and urban geography, winner of the Vautrin Lud Prize.
- Edward Soja (born 1941)-Noted for his work on regional development, planning and governance along with coining the terms Synekism and Postmetropolis.

- Michael Frank Goodchild (1944-)-prominent GIS scholar and winner of the RGS founder's medal in 2003.
- Doreen Massey (1944-)-Key scholar in the space and places of globalization and its pluralities, winner of the Vautrin Lud Prize.
- Nigel Thrift (1949-)-originator of non-representational theory.
- Ellen Churchill Semple (1863–1932)-She was America's first influential female geographer.

Institutions and societies

- Anton Melik Geographical Institute (Slovenia)
- National Geographic Society (U.S.)
- American Geographical Society (U.S.)
- National Geographic Bee (U.S.)
- Royal Canadian Geographical Society (Canada)
- Royal Geographical Society (UK)

2

Cartography

Cartography (from Greek *chartis* = map and *graphein* = write) is the study and practice of making maps. Combining science, aesthetics, and technique, cartography builds on the premise that reality can be modeled in ways that communicate spatial information effectively.

The fundamental problems of traditional cartography are to:

- Set the map's agenda and select traits of the object to be mapped. This is the concern of map editing. Traits may be physical, such as roads or land masses, or may be abstract, such as toponyms or political boundaries.
- Represent the terrain of the mapped object on flat media. This is the concern of map projections.
- Eliminate characteristics of the mapped object that are not relevant to the map's purpose. This is the concern of generalization.
- Reduce the complexity of the characteristics that will be mapped. This is also the concern of generalization.
- Orchestrate the elements of the map to best convey its message to its audience. This is the concern of map design.

Modern cartography is closely integrated with geographic information science (GIScience) and constitutes many theoretical and practical foundations of geographic information systems.

History

The earliest known map is a matter of some debate, both because the definition of "map" is not sharp and because some artifacts speculated to be maps might actually

be something else. A wall painting, which may depict the ancient Anatolian city of Catalhoyuk, has been dated to the late 7th millennium BCE. Other known maps of the ancient world include the Minoan "House of the Admiral" wall painting from c. 1600 BCE, showing a seaside community in an oblique perspective and an engraved map of the holy Babylonian city of Nippur, from the Kassite period (14th – 12th centuries BCE). The oldest surviving world maps are the Babylonian world maps from the 9th century BCE. One shows Babylon on the Euphrates, surrounded by a circular landmass showing Assyria, Urartu and several cities, in turn surrounded by a "bitter river" (Oceanus), with seven islands arranged around it. Another depicts Babylon as being further north from the centre of the world.

The ancient Greeks and Romans created maps, beginning at latest with Anaximander in the 6th century BC. In the 2nd century AD, Ptolemy produced his treatise on cartography, Geographia. This contained Ptolemy's world map-the world then known to Western society *(Ecumene)*. As early as the 8th century, Arab scholars were translating the works of the Greek geographers into Arabic.

In ancient China, geographical literature spans back to the 5th century BC. The oldest extant Chinese maps come from the State of Qin, dated back to the 4th century BC, during the Warring States Period. In the book of the *Xin Yi Xiang Fa Yao*, published in 1092 by the Chinese scientist Su Song, a star map on the equidistant cylindrical projection. Although this method of charting seems to have existed in China even prior to this publication and scientist, the greatest significance of the star maps by Su Song is that they represent the oldest existent star maps in printed form.

Early forms of cartography of India included the locations of the Pole star and other constellations of use. These charts may have been in use by the beginning of the Common Era for purposes of navigation.

Mappa mundi is the general term used to describe Medieval European maps of the world. Approximately 1,100 mappae mundi are known to have survived from the Middle Ages. Of these, some 900 are found illustrating manuscripts and the remainder exist as stand-alone documents.

The Arab geographer Muhammad al-Idrisi produced his medieval atlas *Tabula Rogeriana* in 1154. He incorporated the knowledge of Africa, the Indian Ocean and the Far East, gathered by Arab merchants and explorers with the information inherited from the classical geographers to create the most accurate map of the world up until his time. It remained the most accurate world map for the next three centuries.

In the Age of Exploration, from the 15th century to the 17th century, European cartographers both copied earlier maps (some of which had been passed down for

centuries) and drew their own based on explorers' observations and new surveying techniques. The invention of the magnetic compass, telescope and sextant enabled increasing accuracy. In 1492, Martin Behaim, a German cartographer, made the oldest extant globe of the Earth.

Johannes Werner refined and promoted the *Werner map projection*. In 1507, Martin Waldseemuller produced a globular world map and a large 12-panel world wall map (*Universalis Cosmographia*) bearing the first use of the name "America". Portuguese cartographer Diego Ribero was the author of the first known planisphere with a graduated Equator (1527). Italian cartographer Battista Agnese produced at least 71 manuscript atlases of sea charts.

Due to the sheer physical difficulties inherent in cartography, map-makers frequently lifted material from earlier works without giving credit to the original cartographer. For example, one of the most famous early maps of North America is unofficially known as the "Beaver Map", published in 1715 by Herman Moll. This map is an exact reproduction of a 1698 work by Nicolas de Fer. De Fer in turn had copied images that were first printed in books by Louis Hennepin, published in 1697, and François Du Creux, in 1664. By the 18th century, map-makers started to give credit to the original engraver by printing the phrase "After [the original cartographer]" on the work.

Technological Changes

In cartography, technology has continually changed in order to meet the demands of new generations of mapmakers and map users. The first maps were manually constructed with brushes and parchment; therefore, varied in quality and were limited in distribution. The advent of magnetic devices, such as the compass and much later, magnetic storage devices, allowed for the creation of far more accurate maps and the ability to store and manipulate them digitally.

Advances in mechanical devices such as the printing press, quadrant and vernier, allowed for the mass production of maps and the ability to make accurate reproductions from more accurate data. Optical technology, such as the telescope, sextant and other devices that use telescopes, allowed for accurate surveying of land and the ability of mapmakers and navigators to find their latitude by measuring angles to the North Star at night or the sun at noon.

Advances in photochemical technology, such as the lithographic and photochemical processes, have allowed for the creation of maps that have fine details, do not distort in shape and resist moisture and wear. This also eliminated the need for engraving, which further shortened the time it takes to make and reproduce maps.

Advances in electronic technology in the 20th century ushered in another revolution in cartography. Ready availability of computers and peripherals such as monitors, plotters, printers, scanners (remote and document) and analytic stereo

plotters, along with computer programs for visualization, image processing, spatial analysis, and database management, have democratized and greatly expanded the making of maps. The ability to superimpose spatially located variables onto existing maps created new uses for maps and new industries to explore and exploit these potentials.

These days most commercial-quality maps are made using software that falls into one of three main types: CAD, GIS and specialized illustration software. Spatial information can be stored in a database, from which it can be extracted on demand. These tools lead to increasingly dynamic, interactive maps that can be manipulated digitally.

With the field rugged computers, GPS and laser rangefinders, it is possible to perform mapping directly in the terrain. The construction of the map in real time improve productivity and quality of the result. Real time mapping is done for example with Field-map technology.

Map Types

General vs Thematic Cartography

In understanding basic maps, the field of cartography can be divided into two general categories: general cartography and thematic cartography. General cartography involves those maps that are constructed for a general audience and thus contain a variety of features. General maps exhibit many reference and location systems and often are produced in a series. For example, the 1:24,000 scale topographic maps of the United States Geological Survey (USGS) are a standard as compared to the 1:50,000 scale Canadian maps. The government of the UK produces the classic 1:50,000 (replacing the older 1 inch to 1 mile) "Ordnance Survey" maps of the entire UK and with a range of correlated larger-and smaller-scale maps of great detail.

Thematic cartography involves maps of specific geographic themes, oriented toward specific audiences. A couple of examples might be a dot map showing corn production in Indiana or a shaded area map of Ohio counties, divided into numerical choropleth classes. As the volume of geographic data has exploded over the last century, thematic cartography has become increasingly useful and necessary to interpret spatial, cultural and social data.

An orienteering map combines both general and thematic cartography, designed for a very specific user community. The most prominent thematic element is shading, that indicates degrees of difficulty of travel due to vegetation. The vegetation itself is not identified, merely classified by the difficulty ("fight") that it presents.

Topographic vs Topological

A topographic map is primarily concerned with the topographic description of a place, including (especially in the 20th and 21st centuries) the use of contour lines

showing elevation. Terrain or relief can be shown in a variety of ways. A topological map is a very general type of map, the kind you might sketch on a napkin. It often disregards scale and detail in the interest of clarity of communicating specific route or relational information. Beck's London Underground map is an iconic example. Though the most widely used map of "The Tube," it preserves little of reality: it varies scale constantly and abruptly, it straightens curved tracks, and it contorts directions. The only topography on it is the River Thames, letting the reader know whether a station is north or south of the river. That and the topology of station order and interchanges between train lines are all that is left of the geographic space. Yet those are all a typical passenger wishes to know, so the map fulfils its purpose.

Map Design

Map Purpose and Information Selection

Arthur H. Robinson, an American cartographer influential in thematic cartography, stated that a map not properly designed "will be a cartographic failure." He also claimed, when considering all aspects of cartography, that "map design is perhaps the most complex." Robinson codified the mapmaker's understanding that a map must be designed foremost with consideration to the audience and its needs.

From the very beginning of mapmaking, maps "have been made for some particular purpose or set of purposes". The intent of the map should be illustrated in a manner in which the percipient acknowledges its purpose in a timely fashion. The term *percipient* refers to the person receiving information and was coined by Robinson. The principle of figure-ground refers to this notion of engaging the user by presenting a clear presentation, leaving no confusion concerning the purpose of the map. This will enhance the user's experience and keep his attention. If the user is unable to identify what is being demonstrated in a reasonable fashion, the map may be regarded as useless.

Making a meaningful map is the ultimate goal. Alan MacEachren explains that a well designed map "is convincing because it implies authenticity" (1994). An interesting map will no doubt engage a reader. Information richness or a map that is multivariate shows relationships within the map. Showing several variables allows comparison, which adds to the meaningfulness of the map. This also generates hypothesis and stimulates ideas and perhaps further research. In order to convey the message of the map, the creator must design it in a manner which will aid the reader in the overall understanding of its purpose. The title of a map may provide the "needed link" necessary for communicating that message, but the overall design of the map fosters the manner in which the reader interprets it (Monmonier, 1993).

In the 21st century it is possible to find a map of virtually anything from the inner workings of the human body to the virtual worlds of cyberspace. Therefore there are now a huge variety of different styles and types of map-for example, one area which

has evolved a specific and recognisable variation are those used by public transport organisations to guide passengers, namely urban rail and metro maps, many of which are loosely based on 45 degree angles as originally perfected by Harry Beck and George Dow.

Naming Conventions

Most maps use text to label places and for such things as the map title, legend and other information. Although maps are often made in one specific language, place names often differ between languages. So a map made in English may use the name *Germany* for that country, while a German map would use *Deutschland* and a French map *Allemagne*. A word that describes a place using a non-native terminology or language is referred to as an exonym.

In some cases the correct name is not clear. For example, the nation of Burma officially changed its name to Myanmar, but many nations do not recognize the ruling junta and continue to use *Burma*. Sometimes an official name change is resisted in other languages and the older name may remain in common use. Examples include the use of *Saigon* for Ho Chi Minh City, *Bangkok* for Krung Thep and *Ivory Coast* for Côte d'Ivoire.

Difficulties arise when transliteration or transcription between writing systems is required. Some well-known places have well-established names in other languages and writing systems, but in other cases a system of transliteration or transcription is required. Even in the former case, the exclusive use of an exonym may be unhelpful for the map user. It will not be much use for an English user of a map of Italy to show Livorno *only* as "Leghorn" when road signs and railway timetables show it as "Livorno". In transliteration, the characters in one script are represented by characters in another. For example, the Cyrillic letter *Ð* is usually written as *R* in the Latin alphabet, although in many cases it is not as simple as a one-for-one equivalence. Systems exist for transliteration of Arabic, but the results may vary. For example, the Yemeni city of Mocha is written variously in English as Mocha, Al Mukha, al-Mukhâ, Mocca and Moka. Transliteration systems are based on relating written symbols to one another, while transcription is the attempt to spell in one language the phonetic sounds of another. Chinese writing is now usually converted to the Latin alphabet through the Pinyin phonetic transcription systems. Other systems were used in the past, such as Wade-Giles, resulting in the city being spelled *Beijing* on newer English maps and *Peking* on older ones.

Further difficulties arise when countries, especially former colonies, do not have a strong national geographic naming standard. In such cases, cartographers may have to choose between various phonetic spellings of local names versus older imposed, sometimes resented, colonial names. Some countries have multiple official languages, resulting in multiple official placenames. For example, the capital of Belgium is both *Brussel* and *Bruxelles*. In Canada, English and French are official languages and places

have names in both languages. British Columbia is also officially named *la Colombie-Britannique*. English maps rarely show the French names outside of Quebec, which itself is spelled *Québec* in French. The study of placenames is called toponymy, while that of the origin and historical usage of placenames as words is etymology.

In order to improve legibility or to aid the illiterate, some maps have been produced using pictograms to represent places. The iconic example of this practice is Lance Wyman's early plans for the Mexico City Metro, on which stations were shown simply as stylized logos. Wyman also prototyped such a map for the Washington Metro, though ultimately the idea was rejected. Other cities experimenting with such maps are Fukuoka, Guadalajara and Monterrey.

Map Symbology

The quality of a map's design affects its reader's ability to extract information and to learn from the map. Cartographic symbology has been developed in an effort to portray the world accurately and effectively convey information to the map reader. A legend explains the pictorial language of the map, known as its symbology. The title indicates the region the map portrays; the map image portrays the region and so on. Although every map element serves some purpose, convention only dictates inclusion of some elements, while others are considered optional. A menu of map elements includes the neatline (border), compass rose or north arrow, overview map, bar scale, projection and information about the map sources, accuracy and publication.

When examining a landscape, scale can be intuited from trees, houses and cars. Not so with a map. Even such a simple thing as a north arrow is crucial. It may seem obvious that the top of a map should point north, but this might not be the case.

Map colouring is also very important. How the cartographer displays the data in different hues can greatly affect the understanding or feel of the map. Different intensities of hue portray different objectives the cartographer is attempting to get across to the audience. Today, personal computers can display up to 16 million distinct colours at a time. This fact allows for a multitude of colour options for even for the most demanding maps. Moreover, computers can easily hatch patterns in colours to give even more options. This is very beneficial, when symbolizing data in categories such as quintile and equal interval classifications.

Quantitative symbols give a visual measure of the relative size / importance / number that a symbol represents and to symbolize this data on a map, there are two major classes of symbols used for portraying quantitative properties. Proportional symbols change their visual weight according to a quantitative property. These are appropriate for extensive statistics. Choropleth maps portray data collection areas, such as counties or census tracts, with colour. Using colour this way, the darkness and intensity (or value) of the colour is evaluated by the eye as a measure of intensity or concentration.

Map Generalization

A good map has to compromise between portraying the items of interest (or themes) in the right place on the map, and the need to show that item using text or a symbol, which take up space on the map and might displace some other item of information. The cartographer is thus constantly making judgements about what to include, what to leave out and what to show in a *slightly* incorrect place. This issue assumes more importance as the scale of the map gets smaller (i.e. the map shows a larger area) because the information shown on the map takes up more space *on the ground*. A good example from the late 1980s was the Ordnance Survey's first digital maps, where the *absolute* positions of major roads were sometimes a scale distance of hundreds of metres away from ground truth, when shown on digital maps at scales of 1:250,000 and 1:625,000, because of the overriding need to annotate the features.

Map Projections

The Earth being spherical, any flat representation generates distortions, where shapes, distances, and areas cannot all be conserved simultaneously. The mapmaker must choose a suitable map projection according to the space to be mapped and the purpose of the map.

Cartographic Errors

Some maps contain deliberate errors or distortions, either as propaganda or as a "watermark" to help the copyright owner identify infringement if the error appears in competitors' maps. The latter often come in the form of nonexistent, misnamed, or misspelled "trap streets". Other names and forms for this are paper townsites, fictitious entries, and copyright easter eggs.

Another motive for deliberate errors is cartographic "vandalism": a mapmaker wishing to leave his or her mark on the work. Mount Richard, for example, was a fictitious peak on the Rocky Mountains' continental divide that appeared on a Boulder County, Colorado map in the early 1970s. It is believed to be the work of draftsman Richard Ciacci. The fiction was not discovered until two years later.

Geographic Information System

A geographic information system, geographical information science, or geospatial information studies is a system designed to capture, store, manipulate, analyse, manage, and present all types of geographically referenced data. In the simplest terms, GIS is the merging of cartography, statistical analysis, and database technology.

A GIS can be thought of as a system—it digitally creates and "manipulates" spatial areas that may be jurisdictional, purpose or application-oriented for which a specific GIS is developed. Hence, a GIS developed for an application, jurisdiction, enterprise or purpose may not be necessarily interoperable or compatible with a GIS

that has been developed for some other application, jurisdiction, enterprise, or purpose. What goes beyond a GIS is a spatial data infrastructure (SDI), a concept that has no such restrictive boundaries.

Therefore, in a general sense, the term describes any information system that integrates, stores, edits, analyses, shares and displays geographic information for informing decision making. The term GIS-Centric, however, has been specifically defined as the use of the Esri ArcGIS geodatabase as the asset / feature data repository central to Computerized Maintenance Management System (CMMS) as a part of Enterprise Asset Management and analytical software systems. GIS-centric certification criteria has been specifically defined by NAGCS, the National Association of GIS-Centric Solutions. Geographic information science is the science underlying the geographic concepts, applications and systems.

Applications

GIS technology can be used for:

- earth surface-based scientific investigations;
- resource management;
- reference and projections of a geospatial nature, both artificial and natural;
- asset management and location planning;
- archaeology;
- environmental impact-assessment;
- infrastructure assessment and development;
- urban planning;
- cartography, for a thematic and / or time-based purpose;
- criminology;
- geospatial intelligence;
- GIS data development;
- geographic history;
- marketing;
- logistics;
- population and demographic studies;
- public health planning;
- prospectivity mapping;
- statistical analysis;
- GIS in environmental contamination;
- Disease surveillance;
- military planning.

Examples of use are:

- GIS may allow emergency planners to easily calculate emergency response times and the movement of response resources (for logistics) in the case of a natural disaster;
- GIS might be used to find wetlands that need protection strategies regarding pollution;
- GIS can be used by a company to site a new business location to take advantage of GIS data identified trends to respond to a previously under-served market; Most city and transportation systems planning offices have GIS sections;
- GIS can be used to track the spread of emerging infectious disease threats. This allows for informed pandemic planning and enhanced preparedness.

History of Development

In 1854, John Snow depicted a cholera outbreak in London using points to represent the locations of some individual cases, possibly the earliest use of the geographic method. His study of the distribution of cholera led to the source of the disease, a contaminated water pump (the Broad Street Pump, whose handle he had disconnected, thus terminating the outbreak) within the heart of the cholera outbreak. While the basic elements of topography and theme existed previously in cartography, the John Snow map was unique, using cartographic methods not only to depict but also to analyse clusters of geographically-dependent phenomena for the first time.

The early 20th century saw the development of photozincography, which allowed maps to be split into layers, for example one layer for vegetation and another for water. This was particularly used for printing contours – drawing these was a labour intensive task but having them on a separate layer meant they could be worked on without the other layers to confuse the draughtsman. This work was originally drawn on glass plates but later, plastic film was introduced, with the advantages of being lighter, using less storage space and being less brittle, among others. When all the layers were finished, they were combined into one image using a large process camera. Once colour printing came in, the layers idea was also used for creating separate printing plates for each colour. While the use of layers much later became one of the main typical features of a contemporary GIS, the photographic process just described is not considered to be a GIS in itself – as the maps were just images with no database to link them to.

Computer hardware development spurred by nuclear weapon research led to general-purpose computer 'mapping' applications by the early 1960s.

The year 1960 saw the development of the world's first true operational GIS in Ottawa, Ontario, Canada by the federal Department of Forestry and Rural Development. Developed by Dr. Roger Tomlinson, it was called the Canada Geographic Information System (CGIS) and was used to store, analyse, and manipulate data collected for the

Canada Land Inventory (CLI) – an effort to determine the land capability for rural Canada by mapping information about soils, agriculture, recreation, wildlife, waterfowl, forestry and land use at a scale of 1:50,000. A rating classification factor was also added to permit analysis.

CGIS was an improvement over 'computer mapping' applications as it provided capabilities for overlay, measurement and digitizing / scanning. It supported a national coordinate system that spanned the continent, coded lines as arcs having a true embedded topology and it stored the attribute and locational information in separate files. As a result of this, Tomlinson has become known as the 'father of GIS', particularly for his use of overlays in promoting the spatial analysis of convergent geographic data.

CGIS lasted into the 1990s and built a large digital land resource database in Canada. It was developed as a mainframe-based system in support of federal and provincial resource planning and management. Its strength was continent-wide analysis of complex datasets. The CGIS was never available in a commercial form.

In 1964, Howard T. Fisher formed the Laboratory for Computer Graphics and Spatial Analysis at the Harvard Graduate School of Design (LCGSA 1965–1991), where a number of important theoretical concepts in spatial data handling were developed, and which by the 1970s had distributed seminal software code and systems, such as 'SYMAP', 'GRID' and 'ODYSSEY' – that served as sources for subsequent commercial development—to universities, research centres and corporations worldwide.

By the early 1980s, M&S Computing (later Intergraph) along with Bentley Systems Incorporated for the CAD platform, Environmental Systems Research Institute (ESRI), CARIS (Computer Aided Resource Information System) and ERDAS emerged as commercial vendors of GIS software, successfully incorporating many of the CGIS features, combining the first generation approach to separation of spatial and attribute information with a second generation approach to organizing attribute data into database structures. In parallel, the development of two public domain systems began in the late 1970s and early 1980s.

The Map Overlay and Statistical System (MOSS) project started in 1977 in Fort Collins, Colorado under the auspices of the Western Energy and Land Use Team (WELUT) and the US Fish and Wildlife Service. GRASS GIS was introduced in 1982 by the US Army Corps of Engineering Research Laboratory (USA-CERL) in Champaign, Illinois, a branch of the US Army Corps of Engineers to meet the need of the US military for software for land management and environmental planning.

In the later 1980s and 1990s, industry growth was spurred on by the growing use of GIS on Unix workstations and the personal computer. By the end of the 20th century, the rapid growth in various systems had been consolidated and standardized

on relatively few platforms and users were beginning to explore the concept of viewing GIS data over the Internet, requiring data format and transfer standards. More recently, a growing number of free, open-source GIS packages run on a range of operating systems and can be customized to perform specific tasks. Increasingly geospatial data and mapping applications are being made available via the world wide web.

Several authoritative books on the history of GIS have been published.

GIS Techniques and Technology

Modern GIS technologies use digital information, for which various digitized data creation methods are used. The most common method of data creation is digitization, where a hard copy map or survey plan is transferred into a digital medium through the use of a computer-aided design (CAD) program, and geo-referencing capabilities. With the wide availability of ortho-rectified imagery (both from satellite and aerial sources), heads-up digitizing is becoming the main avenue through which geographic data is extracted. Heads-up digitizing involves the tracing of geographic data directly on top of the aerial imagery instead of by the traditional method of tracing the geographic form on a separate digitizing tablet (heads-down digitizing).

Relating Information from Different Sources

GIS uses spatio-temporal (space-time) location as the key index variable for all other information. Just as a relational database containing text or numbers can relate many different tables using common key index variables, GIS can relate otherwise unrelated information by using location as the key index variable. The key is the location and/or extent in space-time.

Any variable that can be located spatially, and increasingly also temporally, can be referenced using a GIS. Locations or extents in Earth space–time may be recorded as dates/times of occurrence, and x, y, and z coordinates representing, longitude, latitude, and elevation, respectively. These GIS coordinates may represent other quantified systems of temporo-spatial reference (for example, film frame number, stream gage station, highway mile-marker, surveyor benchmark, building address, street intersection, entrance gate, water depth sounding, POS or CAD drawing origin/units). Units applied to recorded temporal-spatial data can vary widely, but all Earth-based spatial–temporal location and extent references should, ideally, be relatable to one another and ultimately to a "real" physical location or extent in space–time.

Related by accurate spatial information, an incredible variety of real-world and projected past or future data can be analysed, interpreted and represented to facilitate education and decision making. This key characteristic of GIS has begun to open new avenues of scientific inquiry into behaviours and patterns of previously considered unrelated real-world information.

GIS Uncertainties

GIS accuracy depends upon source data, and how it is encoded to be data referenced. Land surveyors have been able to provide a high level of positional accuracy utilizing the GPS derived positions. [Retrieved from Federal Geographic Data Committee] the high-resolution digital terrain and aerial imagery, [Retrieved NJGIN] the powerful computers, Web technology, are changing the quality, utility, and expectations of GIS to serve society on a grand scale, but nevertheless there are other source data that has an impact on the overall GIS accuracy like: paper maps that are not found to be very suitable to achieve the desired accuracy since the aging of maps affects their dimensional stability.

In developing a digital topographic data base for a GIS, topographical maps are the main source of data. Aerial photography and satellite images are extra sources for collecting data and identifying attributes which can be mapped in layers over a location facsimile of scale. The scale of a map and geographical rendering area representation type are very important aspects since the information content depends mainly on the scale set and resulting locatability of the map's representations. In order to digitize a map, the map has to be checked within theoretical dimensions, then scanned into a raster format, and resulting raster data has to be given a theoretical dimension by a rubber sheeting/warping technology process.

Uncertainty is a significant problem in designing a GIS because spatial data tend to be used for purposes for which they were never intended. Some maps were made many decades ago, where at that time the computer industry was not even in its perspective establishments. This has led to historical reference maps without common norms. Map accuracy is a relative issue of minor importance in cartography. All maps are established for communication ends. Maps use a historically constrained technology of pen and paper to communicate a view of the world to their users. Cartographers feel little need to communicate information based on accuracy, for when the same map is digitized and input into a GIS, the mode of use often changes. The new uses extend well beyond a determined domain for which the original map was intended and designed.

A quantitative analysis of maps brings accuracy issues into focus. The electronic and other equipment used to make measurements for GIS is far more precise than the machines of conventional map analysis. [Retrieved USGS]. The truth is that all geographical data are inherently inaccurate, and these inaccuracies will propagate through GIS operations in ways that are difficult to predict, yet have goals of conveyance in mind for original design. Accuracy Standards for 1:24000 Scales Map: 1:24,000 ± 40.00 feet.

This means that when we see a point or attribute on a map, its "probable" location is within a +/-40 foot area of its rendered reference, according to area representations and scale.

A GIS can also convert existing digital information, which may not yet be in map form, into forms it can recognize, employ for its data analysis processes, and use in forming mapping output. For example, digital satellite images generated through remote sensing can be analysed to produce a map-like layer of digital information about vegetative covers on land locations. Another fairly recently developed resource for naming GIS location objects is the Getty Thesaurus of Geographic Names (GTGN), which is a structured vocabulary containing about 1,000,000 names and other information about places.

Likewise, researched census or hydrological tabular data can be displayed in map-like form, serving as layers of thematic information for forming a GIS map.

Data Representation

GIS data represents real objects (such as roads, land use, elevation, trees, waterways, etc.) with digital data determining the mix. Real objects can be divided into two abstractions: discrete objects (e.g., a house) and continuous fields (such as rainfall amount, or elevations). Traditionally, there are two broad methods used to store data in a GIS for both kinds of abstractions mapping references: raster images and vector. Points, lines, and polygons are the stuff of mapped location attribute references. A new hybrid method of storing data is that of identifying point clouds, which combine three-dimensional points with RGB information at each point, returning a “3D colour image”. GIS Thematic maps then are becoming more and more realistically visually descriptive of what they set out to show or determine.

Raster

A raster data type is, in essence, any type of digital image represented by reducible and enlargeable grids. Anyone who is familiar with digital photography will recognize the Raster graphics pixel as the smallest individual grid unit building block of an image, usually not readily identified as an artifact shape until an image is produced on a very large scale. A combination of the pixels making up an image colour formation scheme will compose details of an image, as is distinct from the commonly used points, lines, and polygon area location symbols of scalable vector graphics as the basis of the vector model of area attribute rendering. While a digital image is concerned with its output blending together its grid based details as an identifiable representation of reality, in a photograph or art image transferred into a computer, the raster data type will reflect a digitized abstraction of reality dealt with by grid populating tones or objects, quantities, cojoined or open boundaries, and map relief schemas. Aerial photos are one commonly used form of raster data, with one primary purpose in mind: to display a detailed image on a map area, or for the purposes of rendering its identifiable objects by digitization. Additional raster data sets used by a GIS will contain information regarding elevation, a digital elevation model, or reflectance of a particular wavelength of light, Landsat, or other electromagnetic

spectrum indicators. Raster data type consists of rows and columns of cells, with each cell storing a single value. Raster data can be images (raster images) with each pixel (or cell) containing a colour value. Additional values recorded for each cell may be a discrete value, such as land use, a continuous value, such as temperature, or a null value if no data is available. While a raster cell stores a single value, it can be extended by using raster bands to represent RGB (red, green, blue) colours, colormaps (a mapping between a thematic code and RGB value), or an extended attribute table with one row for each unique cell value. The resolution of the raster data set is its cell width in ground units.

Raster data is stored in various formats; from a standard file-based structure of TIF, JPEG, etc. to binary large object (BLOB) data stored directly in a relational database management system (RDBMS) similar to other vector-based feature classes. Database storage, when properly indexed, typically allows for quicker retrieval of the raster data but can require storage of millions of significantly sized records.

Vector

In a GIS, geographical features are often expressed as vectors, by considering those features as geometrical shapes. Different geographical features are expressed by different types of geometry:

- Points: Zero-dimensional points are used for geographical features that can best be expressed by a single point reference—in other words, by simple location. Examples include wells, peaks, features of interest, and trailheads. Points convey the least amount of information of these file types. Points can also be used to represent areas when displayed at a small scale. For example, cities on a map of the world might be represented by points rather than polygons. No measurements are possible with point features.
- Lines or polylines: One-dimensional lines or polylines are used for linear features such as rivers, roads, railroads, trails, and topographic lines. Again, as with point features, linear features displayed at a small scale will be represented as linear features rather than as a polygon. Line features can measure distance.
- Polygons: Two-dimensional polygons are used for geographical features that cover a particular area of the earth's surface. Such features may include lakes, park boundaries, buildings, city boundaries, or land uses. Polygons convey the most amount of information of the file types. Polygon features can measure perimeter and area.

Each of these geometries are linked to a row in a database that describes their attributes. For example, a database that describes lakes may contain a lake's depth, water quality, pollution level. This information can be used to make a map to describe a particular attribute of the dataset. For example, lakes could be coloured depending

on level of pollution. Different geometries can also be compared. For example, the GIS could be used to identify all wells (point geometry) that are within one kilometre of a lake (polygon geometry) that has a high level of pollution.

Vector features can be made to respect spatial integrity through the application of topology rules such as 'polygons must not overlap'. Vector data can also be used to represent continuously varying phenomena.

Contour lines and triangulated irregular networks (TIN) are used to represent elevation or other continuously changing values. TINs record values at point locations, which are connected by lines to form an irregular mesh of triangles. The face of the triangles represent the terrain surface.

Advantages and Disadvantages

There are some important advantages and disadvantages to using a raster or vector data model to represent reality:

- Raster datasets record a value for all points in the area covered which may require more storage space than representing data in a vector format that can store data only where needed
- Raster data is computationally less expensive to render than vector graphics
- There are transparency and aliasing problems when overlaying multiple stacked pieces of raster images
- Vector data allows for visually smooth and easy implementation of overlay operations, especially in terms of graphics and shape-driven information like maps, routes and custom fonts, which are more difficult with raster data
- Vector data can be displayed as vector graphics used on traditional maps, whereas raster data will appear as an image that may have a blocky appearance for object boundaries (depending on the resolution of the raster file)
- Vector data can be easier to register, scale, and re-project, which can simplify combining vector layers from different sources
- Vector data is more compatible with relational database environments, where they can be part of a relational table as a normal column and processed using a multitude of operators
- Vector file sizes are usually smaller than raster data, which can be tens, hundreds or more times larger than vector data (depending on resolution)
- Vector data is simpler to update and maintain, whereas a raster image will have to be completely reproduced (Example: a new road is added)
- Vector data allows much more analysis capability, especially for "networks" such as roads, power, rail, telecommunications, etc. (Examples: Best route, largest port, airfields connected to two-lane highways). Raster data will not have all the characteristics of the features it displays

Non-spatial Data

Additional non-spatial data can also be stored along with the spatial data represented by the coordinates of a vector geometry or the position of a raster cell. In vector data, the additional data contains attributes of the feature. For example, a forest inventory polygon may also have an identifier value and information about tree species. In raster data the cell value can store attribute information, but it can also be used as an identifier that can relate to records in another table.

Software is currently being developed to support spatial and non-spatial decision-making, with the solutions to spatial problems being integrated with solutions to non-spatial problems. The end result with these flexible spatial decision-making support systems (FSDSSs) is expected to be that non-experts will be able to use GIS, along with spatial criteria, and simply integrate their non-spatial criteria to view solutions to multi-criteria problems. This system is intended to assist decision-making.

Data Capture

Data capture—entering information into the system—consumes much of the time of GIS practitioners. There are a variety of methods used to enter data into a GIS where it is stored in a digital format.

Existing data printed on paper or PET film maps can be digitized or scanned to produce digital data. A digitizer produces vector data as an operator traces points, lines, and polygon boundaries from a map. Scanning a map results in raster data that could be further processed to produce vector data.

Survey data can be directly entered into a GIS from digital data collection systems on survey instruments using a technique called coordinate geometry (COGO). Positions from a global navigation satellite system (GNSS) like Global Positioning System (GPS), another survey tool, can also be directly entered into a GIS. Current trend is data collection and field mapping carried out directly with field computers (position from GPS and/or laser rangefinder). New technologies allow to create maps as well as analysis directly in the field, projects are more efficient and mapping is more accurate.

Remotely sensed data also plays an important role in data collection and consist of sensors attached to a platform. Sensors include cameras, digital scanners and LIDAR, while platforms usually consist of aircraft and satellites. Recently with the development of Miniature UAVs, aerial data collection is becoming possible at much lower costs, and on a more frequent basis. For example, the Aeryon Scout was used to map a 50 acre area with a Ground sample distance of 1 inch in only 12 minutes.

The majority of digital data currently comes from photo interpretation of aerial photographs. Soft-copy workstations are used to digitize features directly from stereo pairs of digital photographs. These systems allow data to be captured in two and three dimensions, with elevations measured directly from a stereo pair using principles of

photogrammetry. Currently, analog aerial photos are scanned before being entered into a soft-copy system, but as high quality digital cameras become cheaper this step will be skipped.

Satellite remote sensing provides another important source of spatial data. Here satellites use different sensor packages to passively measure the reflectance from parts of the electromagnetic spectrum or radio waves that were sent out from an active sensor such as radar. Remote sensing collects raster data that can be further processed using different bands to identify objects and classes of interest, such as land cover.

When data is captured, the user should consider if the data should be captured with either a relative accuracy or absolute accuracy, since this could not only influence how information will be interpreted but also the cost of data capture.

In addition to collecting and entering spatial data, attribute data is also entered into a GIS. For vector data, this includes additional information about the objects represented in the system.

After entering data into a GIS, the data usually requires editing, to remove errors, or further processing. For vector data it must be made "topologically correct" before it can be used for some advanced analysis. For example, in a road network, lines must connect with nodes at an intersection. Errors such as undershoots and overshoots must also be removed. For scanned maps, blemishes on the source map may need to be removed from the resulting raster. For example, a fleck of dirt might connect two lines that should not be connected.

Raster-to-vector Translation

Data restructuring can be performed by a GIS to convert data into different formats. For example, a GIS may be used to convert a satellite image map to a vector structure by generating lines around all cells with the same classification, while determining the cell spatial relationships, such as adjacency or inclusion.

More advanced data processing can occur with image processing, a technique developed in the late 1960s by NASA and the private sector to provide contrast enhancement, false colour rendering and a variety of other techniques including use of two dimensional Fourier transforms.

Since digital data is collected and stored in various ways, the two data sources may not be entirely compatible. So a GIS must be able to convert geographic data from one structure to another.

Projections, Coordinate Systems and Registration

A property ownership map and a soils map might show data at different scales. Map information in a GIS must be manipulated so that it registers, or fits, with information gathered from other maps. Before the digital data can be analysed, they may have to undergo other manipulations—projection and coordinate conversions,

for example—that integrate them into a GIS. The earth can be represented by various models, each of which may provide a different set of coordinates (e.g., latitude, longitude, elevation) for any given point on the Earth's surface. The simplest model is to assume the earth is a perfect sphere. As more measurements of the earth have accumulated, the models of the earth have become more sophisticated and more accurate. In fact, there are models that apply to different areas of the earth to provide increased accuracy (e.g., North American Datum, 1927 – NAD27 – works well in North America, but not in Europe).

Projection is a fundamental component of map making. A projection is a mathematical means of transferring information from a model of the Earth, which represents a three-dimensional curved surface, to a two-dimensional medium—paper or a computer screen. Different projections are used for different types of maps because each projection particularly suits specific uses. For example, a projection that accurately represents the shapes of the continents will distort their relative sizes.

Since much of the information in a GIS comes from existing maps, a GIS uses the processing power of the computer to transform digital information, gathered from sources with different projections and/or different coordinate systems, to a common projection and coordinate system. For images, this process is called rectification.

Data Modelling

It is difficult to relate wetlands maps to rainfall amounts recorded at different points such as airports, television stations, and high schools. A GIS, however, can be used to depict two-and three-dimensional characteristics of the Earth's surface, subsurface, and atmosphere from information points. For example, a GIS can quickly generate a map with isopleth or contour lines that indicate differing amounts of rainfall. Such a map can be thought of as a rainfall contour map. Many sophisticated methods can estimate the characteristics of surfaces from a limited number of point measurements. A two-dimensional contour map created from the surface modelling of rainfall point measurements may be overlaid and analysed with any other map in a GIS covering the same area.

Additionally, from a series of three-dimensional points, or digital elevation model, isopleth lines representing elevation contours can be generated, along with slope analysis, shaded relief, and other elevation products. Watersheds can be easily defined for any given reach, by computing all of the areas contiguous and uphill from any given point of interest. Similarly, an expected thalweg of where surface water would want to travel in intermittent and permanent streams can be computed from elevation data in the GIS.

Topological Modelling

A GIS can recognize and analyse the spatial relationships that exist within

digitally stored spatial data. These topological relationships allow complex spatial modelling and analysis to be performed. Topological relationships between geometric entities traditionally include adjacency (what adjoins what), containment (what encloses what), and proximity (how close something is to something else).

Networks

Geometric networks are linear networks of objects that can be used to represent interconnected features, and to perform special spatial analysis on them. A geometric network is composed of edges, which are connected at junction points, similar to graphs in mathematics and computer science. Just like graphs, networks can have weight and flow assigned to its edges, which can be used to represent various interconnected features more accurately. Geometric networks are often used to model road networks and public utility networks, such as electric, gas, and water networks. Network modelling is also commonly employed in transportation planning, hydrology modelling, and infrastructure modelling.

Hydrological Modelling

GIS hydrological models can provide a spatial element that other hydrological models lack, with the analysis of variables such as slope, aspect and watershed or catchment area. Terrain analysis is fundamental to hydrology, since water always flows down a slope. As basic terrain analysis of a DEM involves calculation of slope and aspect, DEMs are very useful for hydrological analysis. Slope and aspect can then be used to determine direction of surface runoff, and hence flow accumulation for the formation of streams, rivers and lakes. Areas of divergent flow can also give a clear indication of the boundaries of a catchment. Once a flow direction and accumulation matrix has been created, queries can be performed that show contributing or dispersal areas at a certain point. More detail can be added to the model, such as terrain roughness, vegetation types and soil types, which can influence infiltration and evapotranspiration rates, and hence influencing surface flow. These extra layers of detail ensures a more accurate model. Also, check out GIS in Water Contamination and GIS in Environmental Contamination.

Cartographic Modelling

The term “cartographic modelling” was (probably) coined by Dana Tomlin in his PhD dissertation and later in his book which has the term in the title. Cartographic modelling refers to a process where several thematic layers of the same area are produced, processed, and analysed. Tomlin used raster layers, but the overlay method can be used more generally. Operations on map layers can be combined into algorithms, and eventually into simulation or optimization models.

Map Overlay

The combination of several spatial datasets (points, lines or polygons) creates a

new output vector dataset, visually similar to stacking several maps of the same region. These overlays are similar to mathematical Venn diagram overlays. A union overlay combines the geographic features and attribute tables of both inputs into a single new output. An intersect overlay defines the area where both inputs overlap and retains a set of attribute fields for each. A symmetric difference overlay defines an output area that includes the total area of both inputs except for the overlapping area.

Data extraction is a GIS process similar to vector overlay, though it can be used in either vector or raster data analysis. Rather than combining the properties and features of both datasets, data extraction involves using a "clip" or "mask" to extract the features of one data set that fall within the spatial extent of another dataset.

In raster data analysis, the overlay of datasets is accomplished through a process known as "local operation on multiple rasters" or "map algebra," through a function that combines the values of each raster's matrix. This function may weigh some inputs more than others through use of an "index model" that reflects the influence of various factors upon a geographic phenomenon.

Automated Cartography

Digital cartography and GIS both encode spatial relationships in structured formal representations. GIS is used in digital cartography modelling as a (semi) automated process of making maps, so called Automated Cartography. In practice, it can be a subset of a GIS, within which it is equivalent to the stage of visualization, since in most cases not all of the GIS functionality is used. Cartographic products can be either in a digital or in a hardcopy format. Powerful analysis techniques with different data representation can produce high-quality maps within a short time period. The main problem in Automated Cartography is to use a single set of data to produce multiple products at a variety of scales, a technique known as cartographic generalization.

Geostatistics

Geostatistics is a point-pattern analysis that produces field predictions from data points. It is a way of looking at the statistical properties of those special data. It is different from general applications of statistics because it employs the use of graph theory and matrix algebra to reduce the number of parameters in the data. Only the second-order properties of the GIS data are analysed.

When phenomena are measured, the observation methods dictate the accuracy of any subsequent analysis. Due to the nature of the data (e.g. traffic patterns in an urban environment; weather patterns over the Pacific Ocean), a constant or dynamic degree of precision is always lost in the measurement. This loss of precision is determined from the scale and distribution of the data collection.

To determine the statistical relevance of the analysis, an average is determined so that points (gradients) outside of any immediate measurement can be included to determine their predicted behaviour. This is due to the limitations of the applied statistic and data collection methods, and interpolation is required to predict the behaviour of particles, points, and locations that are not directly measurable.

Interpolation is the process by which a surface is created, usually a raster dataset, through the input of data collected at a number of sample points. There are several forms of interpolation, each which treats the data differently, depending on the properties of the data set. In comparing interpolation methods, the first consideration should be whether or not the source data will change (exact or approximate). Next is whether the method is subjective, a human interpretation, or objective. Then there is the nature of transitions between points: are they abrupt or gradual. Finally, there is whether a method is global (it uses the entire data set to form the model), or local where an algorithm is repeated for a small section of terrain.

Interpolation is a justified measurement because of a spatial autocorrelation principle that recognizes that data collected at any position will have a great similarity to, or influence of those locations within its immediate vicinity.

Digital elevation models (DEM), triangulated irregular networks (TIN), edge finding algorithms, Thiessen polygons, Fourier analysis, (weighted) moving averages, inverse distance weighting, kriging, spline, and trend surface analysis are all mathematical methods to produce interpolative data.

Address Geocoding

Geocoding is interpolating spatial locations (X,Y coordinates) from street addresses or any other spatially referenced data such as ZIP Codes, parcel lots and address locations. A reference theme is required to geocode individual addresses, such as a road centerline file with address ranges. The individual address locations have historically been interpolated, or estimated, by examining address ranges along a road segment. These are usually provided in the form of a table or database. The GIS will then place a dot approximately where that address belongs along the segment of centerline. For example, an address point of 500 will be at the midpoint of a line segment that starts with address 1 and ends with address 1000. Geocoding can also be applied against actual parcel data, typically from municipal tax maps. In this case, the result of the geocoding will be an actually positioned space as opposed to an interpolated point. This approach is being increasingly used to provide more precise location information.

There are several potentially dangerous caveats that are often overlooked when using interpolation.

Various algorithms are used to help with address matching when the spellings of addresses differ. Address information that a particular entity or organization has

data on, such as the post office, may not entirely match the reference theme. There could be variations in street name spelling, community name, etc. Consequently, the user generally has the ability to make matching criteria more stringent, or to relax those parameters so that more addresses will be mapped. Care must be taken to review the results so as not to map addresses incorrectly due to overzealous matching parameters.

Reverse Geocoding

Reverse geocoding is the process of returning an estimated street address number as it relates to a given coordinate. For example, a user can click on a road centerline theme (thus providing a coordinate) and have information returned that reflects the estimated house number. This house number is interpolated from a range assigned to that road segment. If the user clicks at the midpoint of a segment that starts with address 1 and ends with 100, the returned value will be somewhere near 50. Note that reverse geocoding does not return actual addresses, only estimates of what should be there based on the predetermined range.

3

Data Output and Cartography

Cartography is the design and production of maps, or visual representations of spatial data. The vast majority of modern cartography is done with the help of computers, usually using a GIS but production quality cartography is also achieved by importing layers into a design program to refine it. Most GIS software gives the user substantial control over the appearance of the data.

Cartographic work serves two major functions: First, it produces graphics on the screen or on paper that convey the results of analysis to the people who make decisions about resources. Wall maps and other graphics can be generated, allowing the viewer to visualize and thereby understand the results of analyses or simulations of potential events. Web Map Servers facilitate distribution of generated maps through web browsers using various implementations of web-based application programming interfaces (AJAX, Java, Flash, etc.). Second, other database information can be generated for further analysis or use. An example would be a list of all addresses within one mile (1.6 km) of a toxic spill.

Graphic Display Techniques

Traditional maps are abstractions of the real world, a sampling of important elements portrayed on a sheet of paper with symbols to represent physical objects. People who use maps must interpret these symbols. Topographic maps show the shape of land surface with contour lines or with shaded relief.

Today, graphic display techniques such as shading based on altitude in a GIS can make relationships among map elements visible, heightening one's ability to extract and analyse information. For example, two types of data were combined in a GIS to produce a perspective view of a portion of San Mateo County, California.

- The digital elevation model, consisting of surface elevations recorded on a 30-meter horizontal grid, shows high elevations as white and low elevation as black.

- The accompanying Landsat Thematic Mapper image shows a false-colour infrared image looking down at the same area in 30-meter pixels, or picture elements, for the same coordinate points, pixel by pixel, as the elevation information.

A GIS was used to register and combine the two images to render the three-dimensional perspective view looking down the San Andreas Fault, using the Thematic Mapper image pixels, but shaded using the elevation of the landforms. The GIS display depends on the viewing point of the observer and time of day of the display, to properly render the shadows created by the sun's rays at that latitude, longitude, and time of day.

An archeochrome is a new way of displaying spatial data. It is a thematic on a 3D map that is applied to a specific building or a part of a building. It is suited to the visual display of heat loss data.

Spatial ETL

Spatial ETL tools provide the data processing functionality of traditional Extract, Transform, Load (ETL) software, but with a primary focus on the ability to manage spatial data. They provide GIS users with the ability to translate data between different standards and proprietary formats, whilst geometrically transforming the data en-route.

GIS Development

Many disciplines can benefit from GIS technology. An active GIS market has resulted in lower costs and continual improvements in the hardware and software components of GIS. These developments will, in turn, result in a much wider use of the technology throughout science, government, business, and industry, with applications including real estate, public health, crime mapping, national defense, sustainable development, natural resources, landscape architecture, archaeology, regional and community planning, transportation and logistics. GIS is also diverging into location-based services (LBS). LBS allows GPS enabled mobile devices to display their location in relation to fixed assets (nearest restaurant, gas station, fire hydrant), mobile assets (friends, children, police car) or to relay their position back to a central server for display or other processing. These services continue to develop with the increased integration of GPS functionality with increasingly powerful mobile electronics (cell phones, PDAs, laptops).

OGC Standards

The Open Geospatial Consortium (OGC) is an international industry consortium of 384 companies, government agencies, universities and individuals participating in a consensus process to develop publicly available geoprocessing specifications. Open interfaces and protocols defined by OpenGIS Specifications support interoperable

solutions that "geo-enable" the Web, wireless and location-based services, and mainstream IT, and empower technology developers to make complex spatial information and services accessible and useful with all kinds of applications. Open Geospatial Consortium (OGC) protocols include Web Map Service (WMS) and Web Feature Service (WFS).

GIS products are broken down by the OGC into two categories, based on how completely and accurately the software follows the OGC specifications.

Compliant Products are software products that comply to OGC's OpenGIS Specifications. When a product has been tested and certified as compliant through the OGC Testing Program, the product is automatically registered as "compliant" on this site.

Implementing Products are software products that implement OpenGIS Specifications but have not yet passed a compliance test. Compliance tests are not available for all specifications. Developers can register their products as implementing draft or approved specifications, though OGC reserves the right to review and verify each entry.

Web Mapping

In recent years there has been an explosion of mapping applications on the web such as Google Maps and Bing Maps. These websites give the public access to huge amounts of geographic data.

Some of them, like Google Maps and OpenLayers, expose an API that enable users to create custom applications. These toolkits commonly offer street maps, aerial / satellite imagery, geocoding, searches, and routing functionality.

Other applications for publishing geographic information on the web include GeoBase (Telogis GIS software), Smallworld's SIAS or GSS, MapInfo's MapXtreme or PlanAcess or Stratus Connect, Cadcorp's GeognoSIS, Intergraph's GeoMedia WebMap (TM), ESRI's ArcIMS, ArcGIS Server, Autodesk's Mapguide, Bentley's Geo Web Publisher, SeaTrails' AtlasAlive, ObjectFX's Web Mapping Tools, ERDAS APOLLO Suite, Google Earth, Google Fusion Tables, and the open source MapServer or GeoServer.

In recent years web mapping services have begun to adopt features more common in GIS. Services such as Google Maps and Bing Maps allow users to access and annotate maps and share the maps with others.

Global Change, Climate History Program and Prediction of its Impact

Maps have traditionally been used to explore the Earth and to exploit its resources. GIS technology, as an expansion of cartographic science, has enhanced the efficiency and analytic power of traditional mapping. Now, as the scientific community recognizes the environmental consequences of anthropogenic activities influencing climate change,

GIS technology is becoming an essential tool to understand the impacts of this change over time. GIS enables the combination of various sources of data with existing maps and up-to-date information from earth observation satellites along with the outputs of climate change models. This can help in understanding the effects of climate change on complex natural systems. One of the classic examples of this is the study of Arctic Ice Melting.

The outputs from a GIS in the form of maps combined with satellite imagery allow researchers to view their subjects in ways that literally never have been seen before. The images are also invaluable for conveying the effects of climate change to non-scientists.

Adding the Dimension of Time

The condition of the Earth's surface, atmosphere, and subsurface can be examined by feeding satellite data into a GIS. GIS technology gives researchers the ability to examine the variations in Earth processes over days, months, and years.

As an example, the changes in vegetation vigor through a growing season can be animated to determine when drought was most extensive in a particular region. The resulting graphic, known as a normalized vegetation index, represents a rough measure of plant health. Working with two variables over time would then allow researchers to detect regional differences in the lag between a decline in rainfall and its effect on vegetation.

GIS technology and the availability of digital data on regional and global scales enable such analyses. The satellite sensor output used to generate a vegetation graphic is produced for example by the Advanced Very High Resolution Radiometer (AVHRR). This sensor system detects the amounts of energy reflected from the Earth's surface across various bands of the spectrum for surface areas of about 1 square kilometer. The satellite sensor produces images of a particular location on the Earth twice a day. AVHRR and more recently the Moderate-Resolution Imaging Spectroradiometer (MODIS) are only two of many sensor systems used for Earth surface analysis. More sensors will follow, generating ever greater amounts of data.

GIS and related technology will help greatly in the management and analysis of these large volumes of data, allowing for better understanding of terrestrial processes and better management of human activities to maintain world economic vitality and environmental quality.

In addition to the integration of time in environmental studies, GIS is also being explored for its ability to track and model the progress of humans throughout their daily routines. A concrete example of progress in this area is the recent release of time-specific population data by the US Census. In this data set, the populations of cities are shown for daytime and evening hours highlighting the pattern of concentration and dispersion generated by North American commuting patterns. The manipulation

and generation of data required to produce this data would not have been possible without GIS.

Using models to project the data held by a GIS forward in time have enabled planners to test policy decisions. These systems are known as Spatial Decision Support Systems.

Semantics

Tools and technologies emerging from the W3C's Semantic Web Activity are proving useful for data integration problems in information systems. Correspondingly, such technologies have been proposed as a means to facilitate interoperability and data reuse among GIS applications and also to enable new analysis mechanisms.

Ontologies are a key component of this semantic approach as they allow a formal, machine-readable specification of the concepts and relationships in a given domain. This in turn allows a GIS to focus on the intended meaning of data rather than its syntax or structure. For example, reasoning that a land cover type classified as *deciduous needleleaf trees* in one dataset is a specialization or subset of land cover type *forest* in another more roughly classified dataset can help a GIS automatically merge the two datasets under the more general land cover classification. Tentative ontologies have been developed in areas related to GIS applications, for example the hydrology ontology developed by the Ordnance Survey in the United Kingdom and the SWEET ontologies developed by NASA's Jet Propulsion Laboratory. Also, simpler ontologies and semantic metadata standards are being proposed by the W3C Geo Incubator Group to represent geospatial data on the web.

Recent research results in this area can be seen in the International Conference on Geospatial Semantics and the Terra Cognita – Directions to the Geospatial Semantic Web workshop at the International Semantic Web Conference.

Society

With the popularization of GIS in decision making, scholars have begun to scrutinize the social implications of GIS. It has been argued that the production, distribution, utilization, and representation of geographic information are largely related with the social context. Other related topics include discussion on copyright, privacy, and censorship. A more optimistic social approach to GIS adoption is to use it as a tool for public participation.

Neogeography

Neogeography literally means "new geography" (aka Volunteered Geographic Information), and is commonly applied to the usage of geographical techniques and tools used for personal and community activities or for utilization by a non-expert group of users. Application domains of neogeography are typically not formal or analytical.

History

The term neogeography has been used since at least 1922. In the early 1950s in the U.S. it was a term used in the sociology of production & work. The French philosopher Francois Dagognet used it in the title of his 1977 book *Une Epistemologie de l'espace concret: Neo-geographie*. The word was first used in relation to the study of online communities in the 1990s by Kenneth Dowling, the Librarian of the City and County of San Francisco. Immediate precursor terms in the industry press were: "the geospatial Web" and "the geoaware Web" (both 2005); "Where 2.0" (2005); "a dissident cartographic aesthetic" and "mapping and counter-mapping" (2006). These terms arose with the concept of Web 2.0, around the increased public appeal of mapping and geospatial technologies that occurred with the release of such tools as "slippy maps" such as Google Maps, Google Earth, and also with the decreased cost of geolocated mobile devices such as GPS units. Subsequently, the use of geospatial technologies began to see increased integration with non-geographically focused applications.

The term neogeography was first defined in its contemporary sense by Randall Szott in 2006. He argued for a broad scope, to include artists, psychogeography, and more. The technically-oriented aspects of the field, far more tightly defined than in Szott's definition, were outlined by Andrew Turner in his *Introduction to Neogeography* (O'Reilly, 2006). The contemporary use of the term, and the field in general, owes much of its inspiration to the locative media movement that sought to expand the use of location-based technologies to encompass personal expression and society.

Traditional Geographic Information Systems historically have developed tools and techniques targeted towards formal applications that require precision and accuracy. By contrast, neogeography tends to apply to the areas of approachable, colloquial applications. The two realms can have overlap as the same problems are presented to different sets of users: experts and non-experts.

User-generated Geographic Content

Neogeography has also been connected with the increase in user-generated geographic content, closely related to Volunteered Geographic Information. This can be active collection of data such as OpenStreetMap, or passive collection of user-data such as Flickr tags for folksonomic toponyms.

Discussion about the Definition

There is currently much debate about the scope and application of Neogeography in the web mapping, geography, and GIS fields. Some of this discussion considers neogeography to be the ease of use of geographic tools and interfaces while other points focus on the domains of application. Neogeography is not limited to a specific technology and is not strictly web-based, so is not synonymous with web mapping though it is commonly conceived as such.

A number of geographers and geoinformatics scientists (such as Mike Goodchild) have expressed strong reservations about the term "neogeography". They say that geography is an established scientific discipline; uses such as mashups and tags in Google Earth are not scientific works, but are better described as Volunteered Geographic Information.

There are also a great many artists and inter-disciplinary practitioners involved in an engagement with new forms of mapping and locative art. It is thus far wider than simply web mapping.

Geostatistics

Geostatistics is a branch of statistics focusing on spatial or spatiotemporal datasets. Developed originally to predict probability distributions of ore grades for mining operations, it is currently applied in diverse disciplines including petroleum geology, hydrogeology, hydrology, meteorology, oceanography, geochemistry, geometallurgy, geography, forestry, environmental control, landscape ecology, soil science, and agriculture (esp. in precision farming). Geostatistics is applied in varied branches of geography, particularly those involving the spread of diseases (epidemiology), the practice of commerce and military planning (logistics), and the development of efficient spatial networks. Geostatistical algorithms are incorporated in many places, including geographic information systems (GIS) and the R statistical environment.

Background

Geostatistics is intimately related to interpolation methods, but extends far beyond simple interpolation problems. It consists of a collection of numerical and mathematical techniques dealing with the characterization of spatial phenomena. Geostatistical techniques rely on statistical model that is based on random function (or random variable) theory to model the uncertainty associated with spatial estimation and simulation.

A number of simpler interpolation methods / algorithms, such as inverse distance weighting, bilinear interpolation and nearest-neighbor interpolation, were already well known before geostatistics. Geostatistics goes beyond the interpolation problem by considering the studied phenomenon at unknown locations as a set of correlated random variables.

Let $Z(x)$ be the value of the variable of interest at a certain location x. This value is unknown (e.g. temperature, rainfall, piezometric level, geological facies, etc.). Although there exists a value at location x that could be measured, geostatistics considers this value as random since it was not measured, or has not been measured yet. However, the randomness of $Z(x)$ is not complete, but defined by a cumulative distribution function (cdf) that depends on certain information that is known about the value $Z(x)$:

$$F(z, \mathbf{x}) = \text{Prob}\{Z(\mathbf{x}) \leqslant z \mid \text{information}\}.$$

Typically, if the value of *Z* is known at locations close to x (or in the neighborhood of x) one can constrain the pdf of *Z*(x) by this neighbourhood: if a high spatial continuity is assumed, *Z*(x) can only have values similar to the ones found in the neighbourhood. Conversely, in the absence of spatial continuity *Z*(x) can take any value. The spatial continuity of the random variables is described by a model of spatial continuity that can be either a parametric function in the case of variogram-based geostatistics, or have a non-parametric form when using other methods such as multiple-point simulation or pseudo-genetic techniques.

By applying a single spatial model on an entire domain, one makes the assumption that *Z* is a stationary process. It means that the same statistical properties are applicable on the entire domain. Several geostatistical methods provide ways of relaxing this stationarity assumption.

In this framework, one can distinguish two modelling goals:

- 1) Estimating the value for *Z*(x), typically by the expectation, the median or the mode of the pdf *f*(*z*,x). This is usually denoted as an estimation problem.
- 2) Sampling from the entire probability density function *f*(*z*,x) by actually considering each possible outcome of it at each location. This is generally done by creating several alternative maps of *Z*, called realizations. Consider a domain discretized in *N* grid nodes (or pixels). Each realization is a sample of the complete *N*-dimensional joint distribution function.

$$F(\mathbf{z}, \mathbf{x}) = \text{Prob}\{Z(\mathbf{x}_1) \leqslant z_1, Z(\mathbf{x}_2) \leqslant z_2, ..., Z(\mathbf{x}_N) \leqslant z_N\}.$$

In this approach, the presence of multiple solutions to the interpolation problem is acknowledged. Each realization is considered as a possible scenario of what the real variable could be. All associated workflows are then considering ensemble of realizations, and consequently ensemble of predictions that allow for probabilistic forecasting. Therefore, geostatistics is often used to generate or update spatial models when solving inverse problems.

A number of methods exist for both geostatistical estimation and multiple realizations approaches. Several reference books provide a comprehensive overview of the discipline.

Definitions and Tools

- Regionalized variable theory
- Covariance function
- Semi-variance
- Variogram
- Kriging

- Range (geostatistics)
- Sill (geostatistics)
- Nugget effect
- Training image

Main Scientific Journals Related to Geostatistics

- Water Resources Research
- Advances in Water Resources
- Ground Water
- Mathematical Geosciences
- Computers & Geosciences
- Computational Geosciences
- J. Soil Science Society of America
- Environmetrics
- Remote Sensing of the Environment
- Stochastic Environmental Research and Risk Assessment

Scientific Orgainisations Related to Geostatistics

- European Forum for GeoStatistics

Related Software

- gslib is a set of Fortran 77 routines (open source) implementing most of the classical geostatistics estimation and simulation algorithms.
- sgems is a cross-platform (Windows, Unix), open-source software that implements most of the classical geostatistics algorithms (kriging, Gaussian and indicator simulation, etc.) as well as new developments (multiple-points geostatistics). It also provides an interactive 3D visualization and offers the scripting capabilities of Python.
- mgstat is a free MATLAB toolbox that allows calling sgems from MATLAB and transparent import/export of objects.
- gstat is an open source computer code for multivariable geostatistical modelling, prediction and simulation. It is also available as R package.
- R has around 20 other packages dedicated to geostatistics, and around 30 dedicated to other areas of spatial statistics.

Ethnography

Ethnography is a qualitative method aimed to learn and understand cultural phenomena which reflect the knowledge and system of meanings guiding the life of

a cultural group. It is a popular method in various fields of social sciences—particularly in sociology, communication studies, history. —that studies people, ethnic groups and other ethnic formations, their ethnogenesis, composition, resettlement, social welfare characteristics, as well as their material and spiritual culture. It is often employed for gathering empirical data on human societies and cultures. Data collection is often done through participant observation, interviews, questionnaires, etc. Ethnography aims to describe the nature of those who are studied (i.e. to describe a people, an *ethnos*) through writing. In the biological sciences, this type of study might be called a "field study" or a "case report," both of which are used as common synonyms for "ethnography".

Data Collection Methods

One of the most common methods for collecting data in an ethnographic study is direct, first-hand observation of daily participation. This can include participant observation. Another common method is interviewing, which may include conversation with different levels of form and can involve small talk to long interviews. A particular approach to transcribing interview data might be genealogical method. This is a set of procedures by which ethnographers discover and record connections of kinship, descent and marriage using diagrams and symbols. Questionnaires can be used to aid the discovery of local beliefs and perceptions and in the case of longitudinal research, where there is continuous long-term study of an area or site, they can act as valid instrument for measuring changes in the individuals or groups studied. Traditionally, the ethnographer focuses attention on a community, selecting knowledgeable informants who know well the activities of the community. These informants are typically asked to identify other informants who represent the community, often using chain sampling. This process is often effective in revealing common cultural common denominators connected to the topic being studied. Ethnography relies greatly on up-close, personal experience. Participation, rather than just observation, is one of the keys to this process. Ethnography is very useful in social research.

Differences Across Disciplines

The ethnographic method is used across a range of different disciplines, primarily by anthropologists but also frequently by sociologists. Cultural studies, economics, social work, education, ethnomusicology, folklore, religious studies geography, history, linguistics, communication studies, performance studies, advertising, psychology, usability and criminology are other fields which have made use of ethnography.

Cultural and Social Anthropology

Cultural anthropology and social anthropology were developed around ethnographic research and their canonical texts which are mostly ethnographies: e.g. *Argonauts of the Western Pacific* (1922) by Bronis³aw Malinowski, *Ethnologische Excursion*

in Johore by famous Russian ethnographer and naturalist ("The moon man") (1875) Nicholas Miklouho-Maclay, *Coming of Age in Samoa* (1928) by Margaret Mead, *The Nuer* (1940) by E. E. Evans-Pritchard, *Naven* (1936, 1958) by Gregory Bateson or "The Lele of the Kasai" (1963) by Mary Douglas. Cultural and social anthropologists today place such a high value on actually doing ethnographic research that ethnology—the comparative synthesis of ethnographic information—is rarely the foundation for a career.

The typical ethnography is a document written about a particular people, almost always based at least in part on emic views of where the culture begins and ends. Using language or community boundaries to bound the ethnography is common. Ethnographies are also sometimes called "case studies." Ethnographers study and interpret culture, its universalities and its variations through ethnographic study based on fieldwork. An ethnography is a specific kind of written observational science which provides an account of a particular culture, society, or community. The fieldwork usually involves spending a year or more in another society, living with the local people and learning about their ways of life. Ethnographers are participant observers. They take part in events they study because it helps with understanding local behaviour and thought. Classic examples are Carol Stack's *All Our Kin*, Jean Briggs' "Never in Anger", Richard Lee's "Kalahari Hunter-Gatherers," Victor Turner's "Forest of Symbols," David Maybry-Lewis' "Akew-Shavante Society," E.E. Evans-Pritchard's "The Nuer" and Claude Lévi-Strauss' "Tristes Tropiques". Iterations of ethnographic representations in the classic, modernist camp include Bartholomew Dean's recent (2009) contribution, *Urarina Society, Cosmology, and History in Peruvian Amazonia*.

A typical ethnography attempts to be holistic and typically follows an outline to include a brief history of the culture in question, an analysis of the physical geography or terrain inhabited by the people under study, including climate, and often including what biological anthropologists call habitat. Folk notions of botany and zoology are presented as ethnobotany and ethnozoology alongside references from the formal sciences. Material culture, technology and means of subsistence are usually treated next, as they are typically bound up in physical geography and include descriptions of infrastructure. Kinship and social structure (including age grading, peer groups, gender, voluntary associations, clans, moieties, and so forth, if they exist) are typically included. Languages spoken, dialects and the history of language change are another group of standard topics. Practices of childrearing, acculturation and emic views on personality and values usually follow after sections on social structure. Rites, rituals, and other evidence of religion have long been an interest and are sometimes central to ethnographies, especially when conducted in public where visiting anthropologists can see them.

As ethnography developed, anthropologists grew more interested in less tangible aspects of culture, such as values, worldview and what Clifford Geertz termed the

"ethos" of the culture. Clifford Geertz's own fieldwork used elements of a phenomenological approach to fieldwork, tracing not just the doings of people, but the cultural elements themselves. For example, if within a group of people, winking was a communicative gesture, he sought to first determine what kinds of things a wink might mean (it might mean several things). Then, he sought to determine in what contexts winks were used, and whether, as one moved about a region, winks remained meaningful in the same way. In this way, cultural boundaries of communication could be explored, as opposed to using linguistic boundaries or notions about residence. Geertz, while still following something of a traditional ethnographic outline, moved outside that outline to talk about "webs" instead of "outlines" of culture.

Within cultural anthropology, there are several sub-genres of ethnography. Beginning in the 1950s and early 1960s, anthropologists began writing "bio-confessional" ethnographies that intentionally exposed the nature of ethnographic research. Famous examples include *Tristes Tropiques* (1955) by Claude Lévi-Strauss, *The High Valley* by Kenneth Read, and *The Savage and the Innocent* by David Maybury-Lewis, as well as the mildly fictionalized *Return to Laughter* by Elenore Smith Bowen (Laura Bohannan).

Later "reflexive" ethnographies refined the technique to translate cultural differences by representing their effects on the ethnographer. Famous examples include "Deep Play: Notes on a Balinese Cockfight" by Clifford Geertz, *Reflections on Fieldwork in Morocco* by Paul Rabinow, *The Headman and I* by Jean-Paul Dumont, and *Tuhami* by Vincent Crapanzano. In the 1980s, the rhetoric of ethnography was subjected to intense scrutiny within the discipline, under the general influence of literary theory and post-colonial / post-structuralist thought. "Experimental" ethnographies that reveal the ferment of the discipline include *Shamanism, Colonialism, and the Wild Man* by Michael Taussig, *Debating Muslims* by Michael F. J. Fischer and Mehdi Abedi, *A Space on the Side of the Road* by Kathleen Stewart, and *Advocacy after Bhopal* by Kim Fortun.

Sociology

Sociology is another field which prominently features ethnographies. Urban sociology and the Chicago School in particular are associated with ethnographic research, with some well-known early examples being *Street Corner Society* by William Foote Whyte and *Black Metropolis* by St. Clair Drake and Horace R. Cayton, Jr.. Some of the influence for this can be traced to the anthropologist Lloyd Warner who was on the Chicago sociology faculty, and to Robert Park's experience as a journalist. Symbolic interactionism developed from the same tradition and yielded several excellent sociological ethnographies, including *Shared Fantasy* by Gary Alan Fine, which documents the early history of fantasy role-playing games. Other important ethnographies in the discipline of sociology include Pierre Bourdieu's work on Algeria and France, Paul Willis's *Learning To Labour* on working class youth, and the work

of Elijah Anderson, Mitchell Duneier, Loic Wacquant on black America and Glimpses of Madrasa From Africa, 2010 Lai Olurode. But even though many sub-fields and theoretical perspectives within sociology use ethnographic methods, ethnography is not the *sine qua non* of the discipline, as it is in cultural anthropology.

Communication Studies

Beginning in the 1960s and 1970s, ethnographic research methods began to be widely employed by communication scholars. Studies such as Gerry Philipsen's analysis of cultural communication strategies in a blue-collar, working class neighbourhood on the south side of Chicago, *Speaking 'Like a Man' in Teamsterville,* paved the way for the expansion of ethnographic research in the study of communication. Scholars of communication studies use ethnographic research methods to analyse communication behaviours, seeking to answer the "why" and "how come" questions of human communication. Often this type of research results in a case study or field study such as an analysis of speech patterns at a protest rally or the way firemen communicate during "down time" at a fire station. Like anthropology scholars, communication scholars often immerse themselves, participate in and/or directly observe the particular social group being studied.

Other Fields

The American anthropologist George Spindler was a pioneer in applying ethnographic methodology to the classroom.

Anthropologists like Daniel Miller and Mary Douglas have used ethnographic data to answer academic questions about consumers and consumption. In this sense, Tony Salvador, Genevieve Bell, and Ken Anderson describe design ethnography as being "a way of understanding the particulars of daily life in such a way as to increase the success probability of a new product or service or, more appropriately, to reduce the probability of failure specifically due to a lack of understanding of the basic behaviours and frameworks of consumers."

Businesses, too, have found ethnographers helpful for understanding how people use products and services, as indicated in the increasing use of ethnographic methods to understand consumers and consumption, or for new product development (such as video ethnography). Naked Eye Research is a UK based company that specialises in video ethnography, which involves participating, observing and describing how people from particular cultural groups respond to the situations they find themselves in. The recent Ethnographic Praxis in Industry (EPIC) conference is evidence of this. Ethnographers' systematic and holistic approach to real-life experience is valued by product developers, who use the method to understand unstated desires or cultural practices that surround products. Where focus groups fail to inform marketers about what people really do, ethnography links what people say to what they actually do—avoiding the pitfalls that come from relying only on self-reported, focus-group data.

Evaluating Ethnography

Ethnographic methodology is not usually evaluated in terms of philosophical standpoint (such as positivism and emotionalism). Ethnographic studies nonetheless need to be evaluated in some manner. While there is no consensus on evaluation standards, Richardson (2000) provides 5 criteria that ethnographers might find helpful.

1. *Substantive Contribution*: "Does the piece contribute to our understanding of social-life?"
2. *Aesthetic Merit*: "Does this piece succeed aesthetically?"
3. *Reflexivity*: "How did the author come to write this text…Is there adequate self-awareness and self-exposure for the reader to make judgments about the point of view?"
4. *Impact*: "Does this affect me? Emotionally? Intellectually?" Does it move me?
5. *Expresses a Reality*: "Does it seem 'true'—a credible account of a cultural, social, individual, or communal sense of the 'real'?"

Ethics

Gary Alan Fine argues that the nature of ethnographic inquiry demands that researchers deviate from formal and idealistic rules or ethics that have come to be widely accepted in qualitative and quantitative approaches in research. Many of these ethical assumptions are rooted in positivist and post-positivist epistemologies that have adapted over time, but nonetheless are apparent and must be accounted for in all research paradigms. These ethical dilemmas are evident throughout the entire process of conducting ethnographies, including the design, implementation, and reporting of an ethnographic study. Essentially, Fine maintains that researchers are typically not as ethical as they claim or assume to be — and that "each job includes ways of doing things that would be inappropriate for others to know".

Fine is not necessarily casting blame or pointing his finger at ethnographic researchers, but rather is attempting to show that researchers often make idealized ethical claims and standards which in actuality are inherently based on partial truths and self-deceptions. Fine also acknowledges that many of these partial truths and self-deceptions are unavoidable. He maintains that "illusions" are essential to maintain an occupational reputation and avoid potentially more caustic consequences. He claims, "Ethnographers cannot help but lie, but in lying, we reveal truths that escape those who are not so bold". Based on these assertions, Fine establishes three conceptual clusters in which ethnographic ethical dilemmas can be situated: "Classic Virtues," "Technical Skills," and "Ethnographic Self."

Much debate surrounding the issue of ethics arose after the ethnographer Napoleon Chagnon conducted his ethnographic fieldwork with the Yanomamo people of South America.

Classic Virtues

- "The kindly ethnographer" – Most ethnographers present themselves as being more sympathetic than they actually are, which aids in the research process, but is also deceptive. The identity that we present to subjects is different from who we are in other circumstances.
- "The friendly ethnographer" – Ethnographers operate under the assumption that they should not dislike anyone. In actuality, when hated individuals are found within research, ethnographers often crop them out of the findings.
- "The honest ethnographer" – If research participants know the research goals, their responses will likely be skewed. Therefore, ethnographers often conceal what they know in order to increase the likelihood of acceptance.

Technical Skills

- "The Precise Ethnographer" – Ethnographers often create the illusion that field notes are data and reflect what "really" happened. They engage in the opposite of plagiarism, giving credit to those undeserving by not using precise words but rather loose interpretations and paraphrasing. Researchers take near-fictions and turn them into claims of fact. The closest ethnographers can ever really get to reality is an approximate truth.
- "The Observant Ethnographer" – Readers of ethnography are often led to assume the report of a scene is complete – that little of importance was missed. In reality, an ethnographer will always miss some aspect because they are not omniscient. Everything is open to multiple interpretations and misunderstandings. The ability of the ethnographer to take notes and observe varies, and therefore, what is depicted in ethnography is not the whole picture.
- "The Unobtrusive Ethnographer" – As a "participant" in the scene, the researcher will always have an effect on the communication that occurs within the research site. The degree to which one is an "active member" affects the extent to which sympathetic understanding is possible.

The Ethnographic Self

The following appellations are commonly misconceived conceptions of ethnographers:

- "The Candid Ethnographer" – Where the researcher situates themselves within the ethnography is ethically problematic. There is an illusion that everything reported has actually happened because the researcher has been directly exposed to it.
- "The Chaste Ethnographer" – When ethnographers participate within the field, they invariably develop relationships with research subjects / participants. These relationships are sometimes not accounted for within the reporting of

the ethnography despite the fact that they seemingly would influence the research findings.

- "The Fair Ethnographer" – Fine claims that objectivity is an illusion and that everything in ethnography is known from a perspective. Therefore, it is unethical for a researcher to report fairness in their findings.
- "The Literary Ethnographer" – Representation is a balancing act of determining what to "show" through poetic/prosaic language and style versus what to "tell" via straightforward, 'factual' reporting. The idiosyncratic skill of the ethnographer influences the face-value of the research.

Seven principles should be considered for observing, recording and sampling data according to Denzin:

1. The groups should combine symbolic meanings with patterns of interaction.
2. Observe the world from the point of view of the subject, while maintaining the distinction between everyday and scientific perceptions of reality.
3. Link the group's symbols and their meanings with the social relationships.
4. Record all behaviour.
5. Methodology should highlight phases of process, change and stability.
6. The act should be a type of symbolic interactionism.
7. Use concepts that would avoid casual explanations.

Demographic Transition

The demographic transition model (DT) is the transition from high birth and death rates to low birth and death rates as a country develops from a pre-industrial to an industrialized economic system. The theory is based on an interpretation of demographic history developed in 1929 by the American demographer Warren Thompson (1887–1973). Thompson observed changes, or transitions, in birth and death rates in industrialized societies over the previous 200 years. Most developed countries are in stage 3 or 4 of the model; the majority of developing countries have reached stage 2 or stage 3. The major (relative) exceptions are some poor countries, mainly in sub-Saharan Africa and some Middle Eastern countries, which are poor or affected by government policy or civil strife, notably Pakistan, Palestinian Territories, Yemen and Afghanistan. Although this model predicts ever decreasing fertility rates, recent data show that beyond a certain level of development fertility rates increase again.

It is important to note the key differences between developed and less developed countries in understanding the dynamics of demographic transition. The traditional demographic transition began in developed countries in the 18th century and continues in our current era. In less developed countries, this demographic transition started later and is still at an earlier stage in the transition.

Summary of the Theory

The transition involves four stages, or possibly five:

- In stage one, pre-industrial society, death rates and birth rates are high and roughly in balance. All human populations are believed to have had this balance until the late 18th century when this balance ended in Western Europe. In fact, growth rates were less than 0.05% at least since the Agricultural Revolution over 10,000 years ago. Birth and death rates both tend to be very high in this stage. Because both rates are approximately in balance, population growth is typically very slow in stage one.
- In stage two, that of a developing country, the death rates drop rapidly due to improvements in food supply and sanitation, which increase life spans and reduce disease. The improvements specific to food supply typically include selective breeding and crop rotation and farming techniques. Other improvements generally include access to technology, basic healthcare, and education. For example, numerous improvements in public health reduce mortality, especially childhood mortality. Prior to the mid-20th century, these improvements in public health were primarily in the areas of food handling, water supply, sewage, and personal hygiene. Interestingly, one of the variables often cited is the increase in female literacy combined with public health education programs which emerged in the late 19th and early 20th centuries. In Europe, the death rate decline started in the late 18th century in northwestern Europe and spread to the south and east over approximately the next 100 years. Without a corresponding fall in birth rates this produces an imbalance, and the countries in this stage experience a large increase in population.
- In stage three, birth rates fall due to access to contraception, increases in wages, urbanization, a reduction in subsistence agriculture, an increase in the status and education of women, a reduction in the value of children's work, an increase in parental investment in the education of children and other social changes. Population growth begins to level off. The birth rate decline in developed countries started in the late 19th century in northern Europe. While improvements in contraception do play a role in birth rate decline, it should be noted that contraceptives were not generally available nor widely used in the 19th century and as a result likely did not play a significant role in the decline then. It is important to note that birth rate decline is caused also by a transition in values; not just because of the availability of contraceptives.
- During stage four there are both low birth rates and low death rates. Birth rates may drop to well below replacement level as has happened in countries like Germany, Italy, and Japan, leading to a shrinking population, a threat to many industries that rely on population growth. As the large group born during stage two ages, it creates an economic burden on the shrinking working

population. Death rates may remain consistently low or increase slightly due to increases in lifestyle diseases due to low exercise levels and high obesity and an aging population in developed countries. By the late 20th century, birth rates and death rates in developed countries leveled off at lower rates.

As with all models, this is an idealized picture of population change in these countries. The model is a generalization that applies to these countries as a group and may not accurately describe all individual cases. The extent to which it applies to less-developed societies today remains to be seen. Many countries such as China, Brazil and Thailand have passed through the Demographic Transition Model (DTM) very quickly due to fast social and economic change. Some countries, particularly African countries, appear to be stalled in the second stage due to stagnant development and the effect of AIDS.

Stage One

In pre-industrial society, death rates and birth rates were both high and fluctuated rapidly according to natural events, such as drought and disease, to produce a relatively constant and young population. Family planning and contraception were virtually nonexistent; therefore, birth rates were essentially only limited by the ability of women to bear children. Emigration depressed death rates in some special cases (for example, Europe and particularly the Eastern United States during the 19th century), but, overall, death rates tended to match birth rates, often exceeding 40 per 1000 per year.

Children contributed to the economy of the household from an early age by carrying water, firewood, and messages, caring for younger siblings, sweeping, washing dishes, preparing food, and working in the fields. Raising a child cost little more than feeding him; there were no education or entertainment expenses and, in equatorial Africa, there were no clothing expenses either. Thus, the total cost of raising children barely exceeded their contribution to the household. In addition, as they became adults they become a major input to the family business, mainly farming, and were the primary form of insurance for adults in old age. In India, an adult son was all that prevented a widow from falling into destitution. While death rates remained high there was no question as to the need for children, even if the means to prevent them had existed.

During this stage, the society evolves in accordance with Malthusian paradigm, with population essentially determined by the food supply. Any fluctuations in food supply (either positive, for example, due to technology improvements, or negative, due to droughts and pest invasions) tend to translate directly into population fluctuations. Famines resulting in significant mortality are frequent. Overall, the population dynamics during stage one is highly reminiscent of that commonly observed in animals.

Stage Two

This stage leads to a fall in death rates and an increase in population. The changes leading to this stage in Europe were initiated in the Agricultural Revolution of the 18th century and were initially quite slow. In the 20th century, the falls in death rates in developing countries tended to be substantially faster. Countries in this stage include Yemen, Afghanistan, the Palestinian territories, Bhutan and Laos and much of Sub-Saharan Africa (but do not include South Africa, Zimbabwe, Botswana, Swaziland, Lesotho, Namibia, Kenya and Ghana, which have begun to move into stage 3).

The decline in the death rate is due initially to two factors:

- First, improvements in the food supply brought about by higher yields in agricultural practices and better transportation prevent death due to starvation and lack of water. Agricultural improvements included crop rotation, selective breeding, and seed drill technology.
- Second, significant improvements in public health reduce mortality, particularly in childhood. These are not so much medical breakthroughs (Europe passed through stage two before the advances of the mid-20th century, although there was significant medical progress in the 19th century, such as the development of vaccination) as they are improvements in water supply, sewerage, food handling, and general personal hygiene following from growing scientific knowledge of the causes of disease and the improved education and social status of mothers.

A consequence of the decline in mortality in Stage Two is an increasingly rapid rise in population growth (a "population explosion") as the gap between deaths and births grows wider. Note that this growth is not due to an increase in fertility (or birth rates) but to a decline in deaths. This change in population occurred in northwestern Europe during the 19th century due to the Industrial Revolution. During the second half of the 20th century less-developed countries entered Stage Two, creating the worldwide population explosion that has demographers concerned today.

Another characteristic of Stage Two of the demographic transition is a change in the age structure of the population. In Stage One, the majority of deaths are concentrated in the first 5–10 years of life. Therefore, more than anything else, the decline in death rates in Stage Two entails the increasing survival of children and a growing population.

Hence, the age structure of the population becomes increasingly youthful and more of these children enter the reproductive cycle of their lives while maintaining the high fertility rates of their parents. The bottom of the "age pyramid" widens first, accelerating population growth. The age structure of such a population is illustrated by using an example from the Third World today.

Stage Three

Stage Three moves the population towards stability through a decline in the birth rate. Several factors contribute to this eventual decline, although some of them remain speculative:

- In rural areas continued decline in childhood death means that at some point parents realize they need not require so many children to be born to ensure a comfortable old age. As childhood death continues to fall and incomes increase parents can become increasingly confident that fewer children will suffice to help in family business and care for them in old age.
- Increasing urbanization changes the traditional values placed upon fertility and the value of children in rural society. Urban living also raises the cost of dependent children to a family. A recent theory suggests that urbanization also contributes to reducing the birth rate because it disrupts optimal mating patterns. A 2008 study in Iceland found that the most fecund marriages are between distant cousins. Genetic incompatibilities inherent in more distant outbreeding makes reproduction harder.
- In both rural and urban areas, the cost of children to parents is exacerbated by the introduction of compulsory education acts and the increased need to educate children so they can take up a respected position in society. Children are increasingly prohibited under law from working outside the household and make an increasingly limited contribution to the household, as school children are increasingly exempted from the expectation of making a significant contribution to domestic work. Even in equatorial Africa, children now need to be clothed, and may even require school uniforms. Parents begin to consider it a duty to buy children books and toys. Partly due to education and access to family planning, people begin to reassess their need for children and their ability to raise them.
- Increasing female literacy and employment lower the uncritical acceptance of childbearing and motherhood as measures of the status of women. Working women have less time to raise children; this is particularly an issue where fathers traditionally make little or no contribution to child-raising, such as southern Europe or Japan. Valuation of women beyond childbearing and motherhood becomes important.
- Improvements in contraceptive technology are now a major factor. Fertility decline is caused as much by changes in values about children and sex as by the availability of contraceptives and knowledge of how to use them.

The resulting changes in the age structure of the population include a reduction in the youth dependency ratio and eventually population aging. The population structure becomes less triangular and more like an elongated balloon. During the

period between the decline in youth dependency and rise in old age dependency there is a demographic window of opportunity that can potentially produce economic growth through an increase in the ratio of working age to dependent population; the demographic dividend.

However, unless factors such as those listed above are allowed to work, a society's birth rates may not drop to a low level in due time, which means that the society cannot proceed to Stage Four and is locked in what is called a demographic trap.

Countries that have experienced a fertility decline of over 40% from their pre-transition levels include: Costa Rica, El Salvador, Panama, Jamaica, Mexico, Colombia, Ecuador, Guyana, Surinam, Philippines, Indonesia, Malaysia, Sri Lanka, Turkey, Azerbaijan, Turkmenistan, Uzbekistan, Egypt, Tunisia, Algeria, Morocco, Lebanon, South Africa, India, Saudi Arabia, and many Pacific islands.

Countries that have experienced a fertility decline of 25-40% include: Honduras, Guatemala, Nicaragua, Paraguay, Bolivia, Vietnam, Myanmar, Bangladesh, Tajikistan, Jordan, Qatar, Albania, United Arab Emirates, Zimbabwe, and Botswana.

Countries that have experienced a fertility decline of 10-25% include: Haiti, Papua New Guinea, Nepal, Pakistan, Syria, Iraq, Libya, Sudan, Kenya, Ghana and Senegal.

Stage Four

This occurs where birth and death rates are both low. Therefore the total population is high and stable. Some theorists consider there are only 4 stages and that the population of a country will remain at this level. The DTM is only a suggestion about the future population levels of a country. It is not a prediction.

Countries that are at this stage (Total Fertility Rate of less than 2.5 in 1997) include: United States, Canada, Argentina, Australia, New Zealand, most of Europe, Bahamas, Puerto Rico, Trinidad and Tobago, Brazil, Sri Lanka, South Korea, Singapore, Iran, China, Turkey, Thailand and Mauritius.

Stage Five

The original Demographic Transition model has just four stages; however, some theorists consider that a fifth stage is needed to represent countries that have sub-replacement fertility (that is, below 2.1 children per woman). Most European and many East Asian countries now have higher death rates than birth rates. In this stage, population aging and population decline will eventually occur to some extent, presuming that sustained mass immigration does not occur.

Possible Stage Six

There may even be a further stage of demographic development. In an article in the August 2009 issue of *Nature*, Myrskyla, Kohler and Billari show that previously

negative relationship between national wealth (as measured by the human development index (HDI)) and birth rates has become J-shaped. Development promotes fertility decline at low and medium HDI levels, but advanced HDI promotes a rebound in fertility. In many countries with very high levels of development (around 0.95) fertility rates are now approaching two children per woman-although there are exceptions, notably Germany and Japan.

Effects on Age Structure

The decline in death rate and birth rate that occurs during the demographic transition leads to a radical transformation of the age structure. When the death rate declines during the second stage of the transition, the result is primarily an increase in the child population. The reason is that when the death rate is high (stage one), the infant mortality rate is very high, often above 200 deaths per 1000 children born. When the death rate falls or improves, this, in general, results in a significantly lower infant mortality rate and, hence, increased child survival. Over time, as cohorts increased by higher survival rates get older, there will also be an increase in the number of older children, teenagers, and young adults. This implies that there is an increase in the fertile population which, with constant fertility rates, will lead to an increase in the number of children born. This will further increase the growth of the child population. The second stage of the demographic transition, therefore, implies a rise in child dependency.

Historical Studies

Britain

Between 1750 and 1975 England experienced the transition from high levels of both mortality and fertility, to low levels. A major factor was the sharp decline in the death rate for infectious diseases, which has fallen from about 11 per 1,000 to less than 1 per 1,000. By contrast, the death rate from other causes was 12 per 1,000 in 1850 and has not declined markedly. The agricultural revolution and the development of transport, initiated by the construction of canals, led to greater availability of food and coal, and enabled the Industrial Revolution to improve the standard of living. Scientific discoveries and medical breakthroughs did not, in general, contribute importantly to the early major decline in infectious disease mortality, and the decline in fertility occurred before efficient contraception became available.

Ireland

In the 1980s and early 1990s, the Irish demographic status converged to the European norm. Mortality rose above the European Community average, and in 1991 Irish fertility fell to replacement level. The peculiarities of Ireland's past demography and its recent rapid changes challenge established theory. The recent changes have mirrored inward changes in Irish society, with respect to family planning, women

in the work force, the sharply declining power of the Catholic Church, and the emigration factor.

France

France displays real divergences from the standard model of Western demographic evolution. The uniqueness of the French case arises from its specific demographic history, its historic cultural values, and its internal regional dynamics. France's demographic profile is similar to its European neighbours and to developed countries in general, yet it seems to be staving off the population decline of Western countries. With 62.9 million inhabitants in 2006, it is the second most populous country in the European Union, and it displays a certain demographic dynamism, with a growth rate of 2.4% between 2000 and 2005, above the European average. More than two-thirds of that growth can be ascribed to a natural increase resulting from high fertility and birthrates. In contrast, France is one of the developed nations whose migratory balance is rather weak, which is an original feature at the European level. Several interrelated reasons account for such singularities, in particular the impact of pro-family policies accompanied by greater unmarried households and out-of-wedlock births.

These general demographic trends parallel equally important changes in regional demographics. Since 1982 the same significant tendencies have occurred throughout mainland France: demographic stagnation in the least-populated rural regions and industrial regions in the northwest, with strong growth in the southwest and along the Atlantic coast, plus dynamism in metropolitan areas. Shifts in population between regions account for most of the differences in growth. The varying demographic evolution regions can be analysed though the filter of several parameters, including residential facilities, economic growth, and urban dynamism, which yield several distinct regional profiles. The distribution of the French population therefore seems increasingly defined not only by interregional mobility but also by the residential preferences of individual households. These challenges, linked to configurations of population and the dynamics of distribution, inevitably raise the issue of town and country planning. The most recent census figures show that an outpouring of the urban population means that fewer rural areas are continuing to register a negative migratory flow-two-thirds of rural communities have shown some since 2000. The spatial demographic expansion of large cities amplifies the process of peri-urbanization yet is also accompanied by movement of selective residential flow, social selection, and sociospatial segregation based on income.

Asia

McNicoll (2006) examines the common features behind the striking changes in health and fertility in East and Southeast Asia in the 1960s-1990s, focusing on seven countries: Taiwan and South Korea ("tiger" economies), Thailand, Malaysia, and

Indonesia ("second wave" countries), and China and Vietnam ("market-Leninist" economies). Demographic change can be seen as a byproduct of social and economic development together with, in some cases, strong governmental pressures. The transition sequence entailed the establishment of an effective, typically authoritarian, system of local administration, providing a framework for promotion and service delivery in health, education, and family planning. Subsequent economic liberalization offered new opportunities for upward mobility-and greater risks of backsliding-but these opportunities were accompanied by the erosion of social capital and the breakdown or privatization of service programs.

Korea

Cha (2007) analyses a panel dataset to explore how industrial revolution, demographic transition, and human capital accumulation interacted in Korea from 1916-38. Income growth and public investment in health caused mortality to fall, which suppressed fertility and promoted education. Industrialization, skill premium, and closing gender wage gap further induced parents to opt for child quality. Expanding demand for education was accommodated by an active public school building program. The interwar agricultural depression aggravated traditional income inequality, raising fertility and impeding the spread of mass schooling. Landlordism collapsed in the wake of de-colonization, and the consequent reduction in inequality accelerated human and physical capital accumulation, hence growth in South Korea.

Africa

Campbell has studied the demography of 19th-century Madagascar in the light of demographic transition theory. Both supporters and critics of the theory hold to an intrinsic opposition between human and "natural" factors, such as climate, famine, and disease, influencing demography. They also suppose a sharp chronological divide between the precolonial and colonial eras, arguing that whereas "natural" demographic influences were of greater importance in the former period, human factors predominated thereafter. Campbell argues that in 19th-century Madagascar the human factor, in the form of the Merina state, was the predominant demographic influence.

However, the impact of the state was felt through natural forces, and it varied over time. In the late 18th and early 19th centuries Merina state policies stimulated agricultural production, which helped to create a larger and healthier population and laid the foundation for Merina military and economic expansion within Madagascar. From 1820, the cost of such expansionism led the state to increase its exploitation of forced labour at the expense of agricultural production and thus transformed it into a negative demographic force. Infertility and infant mortality, which were probably more significant influences on overall population levels than the adult mortality rate, increased from 1820 due to disease, malnutrition, and stress, all of which stemmed from state forced labour policies. Available estimates indicate little if any population

growth for Madagascar between 1820 and 1895. The demographic "crisis" in Africa, ascribed by critics of the demographic transition theory to the colonial era, stemmed in Madagascar from the policies of the imperial Merina regime, which in this sense formed a link to the French regime of the colonial era. Campbell thus questions the underlying assumptions governing the debate about historical demography in Africa and suggests that the demographic impact of political forces be reevaluated in terms of their changing interaction with "natural" demographic influences.

Russia

Russia has been undergoing a unique demographic transition since the 1980s; observers call it a "demographic catastrophe": the number of deaths exceeds the number of births, life expectancy is drastically decreasing and the number of suicides has increased.

United States

Greenwood and Seshadri (2002) show that from 1800 to 1940 there was a demographic shift from a mostly rural US population with high fertility, with an average of seven children born per white woman, to a minority (43%) rural population with low fertility, with an average of two births per white woman. This shift resulted from technological progress. A sixfold increase in real wages made children more expensive in terms of forgone opportunities to work and increases in agricultural productivity reduced rural demand for labour, a substantial portion of which traditionally had been performed by children in farm families.

A simplification of the DTM theory proposes an initial decline in mortality followed by a later drop in fertility. The changing demographics of the U.S. in the last two centuries did not parallel this model. Beginning around 1800, there was a sharp fertility decline; at this time, an average woman usually produced seven births per lifetime, but by 1900 this number had dropped to nearly four. A mortality decline was not observed in the U.S. until almost 1900—a hundred years following the drop in fertility.

However, this late decline occurred from a very low initial level. It's been estimated that the crude death rate in 17th century rural New England was already as low as 20 deaths per 1000 residents per year (levels of up to 40 per 1000 being typical during stages one and two). The phenomenon is explained by the pattern of colonization of the United States: high death rates unavoidably had to match high birth rates in densely populated Europe, whereas in the United States, westward expansion of the frontier into sparsely populated interior allowed ample room to absorb all the excess people, resulting in exponential population growth (from less than 4 million people in 1790, to 23 million in 1850, to 76 million in 1900.) Today, the U.S. is recognized as having both low fertility and mortality rates. Specifically, birth rates stand at 14 per 1000 per year and death rates at 8 per 1000 per year.

Critical Evaluation

It has to be remembered that the DTM is only a model and cannot necessarily predict the future. It does however give an indication of what the future birth and death rates may be for an underdeveloped country, together with the total population size. Most particularly, of course, the DTM makes no comment on change in population due to migration. It is not applicable for high levels of development, as it has been shown that after a HDI of 0.9 the fertility increases again.

Non-applicability to Less-developed Countries

DTM has a questionable applicability to less economically developed countries (LEDCs), where wealth and information access are limited. For example, the DTM has been validated primarily in Europe, Japan and North America where demographic data exists over centuries, whereas high quality demographic data for most LDCs did not become widely available until the mid-20th century. DTM does not account for recent phenomena such as AIDS; in these areas HIV has become the leading source of mortality. Some trends in waterborne bacterial infant mortality are also disturbing in countries like Malawi, Sudan and Nigeria; for example, progress in the DTM clearly arrested and reversed between 1975 and 2005.

Economic Development not Sufficient cause to Effect Demographic Change

DTM assumes that population changes are induced by industrial changes and increased wealth, without taking into account the role of social change in determining birth rates, e.g., the education of women. In recent decades more work has been done on developing the social mechanisms behind it.

DTM assumes that the birth rate is independent of the death rate. Nevertheless, demographers maintain that there is no historical evidence for society-wide fertility rates rising significantly after high mortality events. Notably, some historic populations have taken many years to replace lives after events such as the Black Death.

Some have claimed that DTM does not explain the early fertility declines in much of Asia in the second half of the 20th century or the delays in fertility decline in parts of the Middle East. Nevertheless, the demographer John C Caldwell has suggested that the reason for the rapid decline in fertility in some developing countries compared to Western Europe, the United States, Canada, Australia and New Zealand is mainly due to government programs and a massive investment in education both by governments and parents.

Malthusian Catastrophe

A Malthusian catastrophe (also phrased Malthusian check, Malthusian crisis, Malthusian disaster, Malthusian fallacy, Malthusian nightmare, or Malthusian theory of population) was originally foreseen to be a forced return to subsistence-level

conditions once population growth had outpaced agricultural production. Later formulations consider economic growth limits as well. The term is also commonly used in discussions of oil depletion.

Based on the work of political economist Thomas Malthus (1766–1834), theories of Malthusian catastrophe are very similar to the Iron Law of Wages. The main difference is that the Malthusian theories predict what will happen over several generations or centuries, whereas the Iron Law of Wages predicts what will happen in a matter of years and decades.

An August 2007 science review in *The New York Times* raised the claim that the Industrial Revolution had enabled the modern world to break out of the Malthusian growth model, while a front page *Wall Street Journal* article in March 2008 pointed out various limited resources which may soon limit human population growth because of a widespread belief in the importance of prosperity for every individual and the rising consumption trends of large developing nations such as China and India.

Work by Thomas Malthus

In 1798, Thomas Malthus published *An Essay on the Principle of Population*, in which he wrote:

"The power of population is so superior to the power of the earth to produce subsistence for man, that premature death must in some shape or other visit the human race. The vices of mankind are active and able ministers of depopulation. They are the precursors in the great army of destruction, and often finish the dreadful work themselves.

But should they fail in this war of extermination, sickly seasons, epidemics, pestilence, and plague advance in terrific array, and sweep off their thousands and tens of thousands. Should success be still incomplete, gigantic inevitable famine stalks in the rear, and with one mighty blow levels the population with the food of the world". —Malthus T.R. 1798. *An essay on the principle of population*.

Notwithstanding the apocalyptic image conveyed by this particular paragraph, Malthus himself did not subscribe to the notion that mankind was fated for a "catastrophe" due to population overshoot. Rather, he believed that population growth was generally restricted by available resources:

"The passion between the sexes has appeared in every age to be so nearly the same that it may always be considered, in algebraic language, as a given quantity. The great law of necessity which prevents population from increasing in any country beyond the food which it can either produce or acquire, is a law so open to our view...that we cannot for a moment doubt it. The different modes which nature takes to prevent or repress a redundant population do not appear, indeed, to us so certain and regular, but though we cannot always predict the mode we may with certainty predict the fact. —Malthus, 1798.

Neo-Malthusian Theory

Neo-Malthusian theory argues that unless at or below subsistence level, a population's fertility will tend to move upwards. Assume for example that a country has 10 breeding groups. Over time this country's fertility will approach that of its fastest growing group in the same way that:

$f(t) = a \times 1.01^t + b \times 1.02^t$, where $a > 0$ and $b > 0$.

Will eventually come to resemble:

$g(t) = b \times 1.02^t$

Regardless of how large the constant *a* is or how small the constant *b* is. Under subsistence conditions the fastest growing group is likely to be that group progressing most rapidly in agricultural technology. However, in above-subsistence conditions the fastest growing group is likely to be the one with the highest fertility. Therefore the fertility of the country will approach that of its most fertile group. This, however, is only part of the problem.

In any group some individuals will be more pro-fertility in their beliefs and practices than others. According to neo-Malthusian theory, these pro-fertility individuals will not only have more children, but also pass their pro-fertility on to their children, meaning a constant selection for pro-fertility similar to the constant natural selection for fertility genes (except much faster because of greater diversity). According to neo-Malthusians, this increase in fertility will lead to hyperexponential population growth that will eventually outstrip growth in economic production. This appears to make any sort of voluntary fertility control futile, in the long run. Neo-Malthusians argue that although adult immigrants (who, at the very least, arrive with human capital) contribute to economic production, there is little or no increase in economic production from increased natural growth and fertility. Neo-Malthusians argue that hyperexponential population growth has begun or will begin soon in developed countries. To this can be added that farmland deteriorates with use. Some areas where there was intensive agriculture in classic times (i.e., the feudal era) had already declined in population because their farmland was worn out, long before he wrote.

At the time Malthus wrote, and for 150 years thereafter, most societies had populations at or beyond their agricultural limits. After World War II, mechanized agriculture produced a dramatic increase in productivity of agriculture and the so-called Green Revolution greatly increased crop yields, expanding the world's food supply while lowering food prices. In response, the growth rate of the world's population accelerated rapidly, resulting in predictions by Paul R. Ehrlich, Simon Hopkins, and many others of an imminent Malthusian catastrophe. However, populations of most developed countries grew slowly enough to be outpaced by gains

in productivity. By 1990, agricultural production appeared to begin peaking in several world regions.

By the early 21st century, many technologically developed countries had passed through the demographic transition, a complex social development encompassing a drop in total fertility rates in response to lower infant mortality, more education of women, increased urbanization, and a wider availability of effective birth control, causing the demographic-economic paradox. Developed and developing countries follow two distinct paths. Most developed countries have sufficient food supply, low fertility rates, and stable (in some cases even declining) populations. In some cases, population growth occurs due to increasing life expectancies, even though fertility rates are below replacement. For example, As of 2008, Spain has approximately 4.6 km^2 of arable land or permanent crops per 1,000 residents, and its average fertility rate is well below replacement level (1.3 children / woman). Its population has grown less than 50% in the last 40 years. The corresponding ratio for Nigeria is only 2.1 km^2 of arable land or crops per 1,000 residents; Nigerian total fertility rate is 5.0 children / woman, and its population has more than tripled during the same 40 years. The Malthusian catastrophe appears to have been temporarily averted in Spain, although the Neo-Malthusian theory argues that the situation is only temporary; on the other hand, a significant part of population of Nigeria lives near subsistence levels.

David Pimentel and Ron Nielsen, working independently, determined that the human population as a whole has passed the numerical point where all can live in comfort, and that we have entered a stage where many of the world's citizens and future generations are trapped in misery. There is evidence that a catastrophe is underway as of at least the 1990s; for example, by the year 2000, children in developing countries were dying at the rate of approximately 11,000,000 per annum from strictly preventable diseases. These data suggest that, by the standard of misery, the catastrophe is underway. The term 'misery' can generally be construed as: high infant mortality, low standards of sanitation, malnutrition, inadequate drinking water, widespread diseases, war, and political unrest.

Regarding famines, data demonstrate the world's food production has peaked in some of the very regions where food is needed the most. For example, in South Asia, approximately half of the land has been degraded such that it no longer has the capacity for food production. In China there has been a 27% irreversible loss of land for agriculture, and it continues to lose arable land at the rate of 2,500 square kilometres per year. In Madagascar, at least 30% of the land previously regarded as arable is irreversibly barren. On the other hand, recent data show the number of overweight people in the world now outnumbers that of malnourished, and the rising tide of obesity continues to expand in both rich and poor countries.

On the assumption that the demographic transition is now spreading from the developed countries to less developed countries, the United Nations Population Fund

estimates that human population may peak in the late 21st century rather than continue to grow until it has exhausted available resources. Some have expressed doubt about the validity of the UN projections, claiming that they are below the projections by others. The most important point is that none of the projections show the population growth beginning to decline before 2050. Indeed, the UN "high" estimate does not decline at all, even out to 2300, indicating the potential for a Malthusian catastrophe. The actual growth curve of the human population is another issue. In the latter part of the 20th century, many argued that it followed exponential growth; however, a more recent view is that the growth in the last millennium has been faster, at a superexponential (possibly hyperbolic, double-exponential, or hyper-exponential) rate. Alternatively, the apparently exponential portion of the human population growth curve may actually fit the lower limb of a logistic curve, or a section of a Lotka–Volterra cycle.

Historians have estimated the total human population on earth back to 10,000 BC. The figure on the right shows the trend of total population from 1800 to 2005, and from there in three projections out to 2100 (low, medium, and high). If population growth were exactly exponential, then the growth rate would be a flat line. The fact that it was increasing from 1920 to 1960 indicates faster-than-exponential growth over this period. However, the growth rate has been decreasing since then, and is projected to continue decreasing. The United Nations population projections out to 2100 (the red, orange, and green lines) show a possible peak in the world's population occurring as early as 2040 in the most optimistic scenario, and by 2075 in the "medium" scenario.

The graph of annual growth rates (below) does not appear exactly as one would expect for long-term exponential growth. For exponential growth it should be a straight line at constant height, whereas in fact the graph from 1800 to 2005 is dominated by an enormous hump that began about 1920, peaked in the mid-1960s, and has been steadily eroding away for the last 40 years. The sharp fluctuation between 1959 and 1960 was due to the combined effects of the Great Leap Forward and a natural disaster in China. Also visible on this graph are the effects of the Great Depression, the two world wars, and possibly also the 1918 flu pandemic.

Though short-term trends, even on the scale of decades or centuries, cannot prove or disprove the existence of mechanisms promoting a Malthusian catastrophe over longer periods, the prosperity of a small fraction of the human population at the beginning of the 21st century, and the debatability of ecological collapse made by Paul R. Ehrlich in the 1960s and 1970s, has led some people, such as economist Julian L. Simon, to question its inevitability.

A 2004 study by a group of prominent economists and ecologists, including Kenneth Arrow and Paul Ehrlich suggests that the central concerns regarding sustainability have shifted from population growth to the consumption / savings ratio, due to shifts in population growth rates since the 1970s. Empirical estimates show

that public policy (taxes or the establishment of more complete property rights) can promote more efficient consumption and investment that are sustainable in an ecological sense; that is, given the current (relatively low) population growth rate, the Malthusian catastrophe can be avoided by either a shift in consumer preferences or public policy that induces a similar shift. However, some contend that the Malthusian catastrophe is not imminent. A 2002 study by the UN Food and Agriculture Organization predicts that world food production will be in excess of the needs of the human population by the year 2030; however, that source also states that hundreds of millions will remain hungry (presumably due to economic realities and political issues).

Application to Energy/Resource Consumption

Another way of applying the Malthusian theory is to substitute other resources, such as sources of energy for food, and energy consumption for population. (Since modern food production and logistics is energy and resource intensive, this is not a big jump. Most of the criteria for applying the theory are still satisfied.) Since energy consumption is increasing much faster than population, and most energy comes from non-renewable sources, the catastrophe appears more imminent, though perhaps not as certain, than when considering food and population continue to behave in a manner contradicting Malthus's assumptions.

Retired physics professor Albert Bartlett, a modern-day Malthusian, has lectured on "Arithmetic, Population and Energy" over 1,500 times. He published an article entitled *Thoughts on Long-Term Energy Supplies: Scientists and the Silent Lie* in Physics Today (July 2004). For a response to Bartlett's argument, see two articles on energy and population in Physics Today, November 2004, and following letters to the editor. A further way of analyzing resource limitation is the dwindling area for storage of soil contaminants and water pollution. The high rate of increase in toxic chemicals in the environment (especially persistent organic chemicals and endocrine altering chemicals) is creating a circumstance of resource limitation (e.g. safe potable water and safe arable land).

Criticism

The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure, that population levels determine agricultural methods, rather than agricultural methods determining population (via food supply). A major point of her book is that "necessity is the mother of invention". Julian Simon was one of many economists who challenged the Malthusian catastrophe, citing (1) the existence of new knowledge, and educated people to take advantage of it, and (2) "economic freedom", that is, the ability of the world to increase production when there is a profitable opportunity to do so. The economist Henry George argued that Malthus didn't provide any evidence of a natural tendency for a population to

overwhelm its ability to provide for itself. George wrote that even the main body of Malthus' work refuted this theory; that examples given show social causes for misery, such as "ignorance and greed... bad government, unjust laws, or war," rather than insufficient food production.

Karl Marx's Theory of Population

Karl Marx (1818-1883) is regarded as the Father of Communism. He did not separately propose any theory of population, but his surplus population theory has been deduced from his theory of communism. Marx opposed and criticized the Malthusian theory of population.

According to Marx, population increase must be interpreted in the context of the capitalistic economic system. A capitalist gives to labour as wage a small share of labour's productivity, and the capitalist himself takes the lion's share. The capitalist introduces more and more machinery and thus increases the surplus value of labour's productivity, which is pocketed by the capitalist. The surplus is the difference between labour's productivity and the wage level. A worker is paid less than the value of his productivity. When machinery is introduced, unemployment increases and, consequently, a reserve army of labour is created. Under these situations, the wage level goes down further, the poor parents cannot properly rear their children and a large part of the population becomes virtually surplus. Poverty, hunger and other social ills are the result of socially unjust practices associated with capitalism.

Population growth, according to Marx, is therefore not related to the alleged ignorance or moral inferiority of the poor, but is a consequence of the capitalist economic system. Marx points out that landlordism, unfavorable and high man-land ratio, uncertainty regarding land tenure system and the like are responsible for low food production in a country. Only in places where the production of food is not adequate does population growth become a problem.

Paul Ehrlich: Neo-Malthusian

As global populations rose spectacularly in the 20th century, theoretical debates over the extent and causes of the population problem expanded. Thomas Malthus and Karl Marx had set the initial stage for the world population debate, but other population theorists-including Paul Ehrlich, Julian Simon, Garrett Hardin, and Barry Commoner-would carry the ongoing discussion in the second half of the 20th century.

In 1968, as world population hovered above 3 billion, Paul Ehrlich authored the book *The Population Bomb,* a widely read publication that sold several million copies in the United States alone. Ehrlich, a biologist, maintained that the rate of population growth was outstripping agricultural growth and the capacity for renewal of Earth's resources. Given current rates of natural increase, Ehrlich predicted "certain" demographic disaster in response to eventual food shortages and disease. In the

opening to his book, he wrote: "The battle to feed all of humanity is over" and later stated that, "In the 1970s and 1980s hundreds of millions of people will starve to death in spite of any crash programs" (Ehrlich 1968). Ehrlich argued that industrialized regions such North America and Europe would be required to undertake "mild" food rationing as starvation spread across the developing worlds of Asia, Latin America, and Africa. In a worst case scenario, he predicted that the lack of food security in the developing world would set into motion several geopolitical crises that could result in thermonuclear war. At its core, Ehrlich's population theory contained three major elements: a rapid rate of change, a limit of some sort, and delays in perceiving the limit.

While some criticized Ehrlich's work as simply a repetition of Malthus's 19th century argument, Ehrlich's most vocal opponent, economist Julian Simon, was skeptical of the more central tenets of the population bomb, particularly the definition of limits. In the 1970s, Julian Simon published two central pieces that served to galvanize the population debate: *The Economics of Population Growth* (1977) followed by *The Ultimate Resource* (1981).

Simon argued that the relationship between population growth and economic growth was not as simple as Ehrlich believed, and that the extent to which population pressure impacted resources was overstated. The crux of Simon's argument centered on his belief that Ehrlich's limit on the availability of resources was misdirected. Simon instead argued that it was not possible to have too many people, for the only limit in determining the scarcity of resources was human imagination. People, the economist suggested, were the ultimate resource. According to Simon, ingenious, resourceful humans had the capacity to invent crops with higher yields, or to construct inexpensive, safe housing for growing populations. Simon's other contention was that current views on population and resource issues failed to take the long view, and that frequently too short a time frame was considered when examining demographic problems.

The Simon-Ehrlich Wager

In 1980, Julian Simon and Paul Ehrlich engaged in a very public debate that underscored their disparate standpoints on population and resource scarcity. Known as the Simon-Ehrlich wager, Simon invited Ehrlich and his colleagues to select and purchase five non-government controlled resources worth a total of $1000 whose value would be measured over time. Agreeing to the wager, Ehrlich's team selected chromium, copper, nickel, tin and tungsten as the commodities and then chose 1990 as the payoff date. If the price of the resource bundle rose, this implied that the resource had become scarcer and Simon therefore would be forced to pay the difference. If the price of the bundle had dropped, this would signify greater abundance, and Simon would receive the monetary difference.

Between 1980 and 1990, the world's population grew by more than 800 million, the largest increase in one decade, causing many to believe that the value of the bundle would rise due to population pressure and corresponding resource scarcity. Yet in September 1990, the inflation adjusted price of all five metals had fallen, forcing Ehrlich to mail Simon a check for $576 to settle the wager. Wired Magazine eventually dubbed Simon a "doomslayer" for his stance against those who argued that an ecological Armageddon was around the corner.

In contrast, while Ehrlich was often criticized as a "doomsdayer" theorist, he is credited for developing a simple equation that examines population's relationship to environmental impact. Known as the IPAT equation, Ehrlich argued that environmental impacts (I) are the result of three variables: population (P); affluence (A); and technology (T), as follows:

$$I = P \times A \times T$$

Not surprisingly, Ehrlich implicated population size as the main driver behind environmental problems, disagreeing with environmentalists such as Barry Commoner, who believed inappropriate technologies and consumption to be the prime causes of degradation. Nevertheless, in developing IPAT, Ehrlich put in place a new framework for population debates that looked beyond numbers to include human impact. Measuring the variables, however, can be challenging, particularly the technology variable.

Garrett Hardin and Lifeboat Ethics

Ehrlich and Simon were not the only theorists of the 1970s to debate the extent and causes of the population problem, nor were they the last to discuss the merits of possible solutions. Biologist Garrett Hardin, known primarily for his research on common property resources, published "Life Boat Ethics" in 1974, a manuscript in which he outlined the case for and against aiding poor, populous nations. Using a lifeboat as a metaphor for the position of rich, industrialized countries, Hardin questioned the ethics of whether "swimmers" surrounding the lifeboat should be taken aboard (or given aid) in light of the vessel's limited carrying capacity.

To explain the metaphor, Hardin pointed to proposals to create a World Food Bank, an international cache of food reserves to which "nations would contribute according to their abilities and from which they would draw according to their needs" (Hardin 1974). Hardin questioned whether we should appeal to our humanitarian impulses and provide aid or whether we'd be better served caring for those individuals already positioned in the boat.

Hardin concluded that the World Food Bank is essentially a commons in disguise where the less "provident" will be able to "multiply" and tax the planet's resources at the expense of other nations that had planned for potential famine and disease through appropriate policies (Hardin 1974: 39). Hardin argued that ultimately, this

disparity would bring eventual ruin upon all those who share in the commons. In the short run, Hardin concluded, a World Food Bank would diminish the need for food but in the long run would increase it without limit given rapid rates of population growth in developing nations.

While some have criticized the lifeboat ethics stance as harsh or callous, Hardin actually supported those humanitarian projects that stressed technology and advice rather than those that supplied food or cash. In drafting his solutions to the population problem, Hardin invoked the Chinese Proverb: "Give a man a fish and he will eat for a day; teach him how to fish and he will eat for the rest of his days". While Hardin criticized foreign aid that "frequently inspires mistrust rather than gratitude on the part of the recipient nation", he supported Rockefeller and Ford Foundation agricultural development projects that funded local, community-based solutions to poverty (Hardin 1974: 40).

Barry Commoner and Poverty

In 1980, biologist Barry Commoner entered the population debate with his chapter entitled "Poverty Breeds Overpopulation". A strong critic of Hardin's lifeboat ethics, Commoner questioned how passengers in the lifeboat and swimmers in the ocean assumed their relative positions in the first place. Tracing the roots of the problem to the colonial period, Commoner argues that initially, colonialism served to improve conditions and develop resources within colonies through the construction of roads, communication, and medical services. However, over time the resultant wealth in the developing world was siphoned away to developed nations in what Commoner calls a process of "demographic parasitism" (Commoner 1980: 4). More simply, the gap between the rich and poor nations grew as the rich fed the poor with their own resources. Commoner suggests that this process of international exploitation had the added effect of rapid population growth in former colonies. In other words, without financial resources available to improve living conditions, people in developing countries relied more heavily upon increased birth rates as a form of social security. Commoner summarized: "The poor countries have high birthrates because they are extremely poor, and they are extremely poor because other countries are extremely rich" (Commoner 1980: 4).

Commoner therefore concluded that the birth rate is not only affected by biological factors such as fertility and contraception but by social factors, such as quality of life. If the standard of living continues to increase, Commoner argued, population rates eventually level off in a self-regulating process. Commoner's solution to the population problem was to increase GDP per capita as a way to motivate voluntary reduction of fertility. He argued that the developed world has a duty to restore the imbalance in wealth between the developed and developing worlds by returning wealth to impoverished nations and abolishing poverty.

Demographic Trap

During "stage 2" of the demographic transition, quality of health care improves and death rates fall, but birth rates still remain high, resulting in a period of high population growth. The term "demographic trap" is used by some demographers to describe a situation where stage 2 persists because "falling living standards reinforce the prevailing high fertility, which in turn reinforces the decline in living standards." This results in more poverty, where people rely on more children to provide them with economic security. Social scientist John Avery explains that this results because the high birth rates and low death rates "lead to population growth so rapid that the development that could have slowed population is impossible."

Results

One of the significant outcomes of the "demographic trap" is explosive population growth. This is currently seen throughout Asia, Africa and Latin America, where death rates have dropped during the last half of the 20th century due to advanced health care. However, in subsequent decades most of those countries were unable to keep improving economic development to match their population's growth: by filling the education needs for more school age children; creating more jobs for the expanding workforce; and providing basic infrastructure and services, such as sewage, roads, bridges, water supplies, electricity, and stable food supplies.

A possible result of a country remaining trapped in stage 2 is its government may reach a state of "demographic fatigue," writes Donald Kaufman. In this condition, the government will lack financial resources to stabilize its population's growth and becomes unable to deal effectively with threats from natural disasters, such as hurricanes, floods, landslides, drought, and disease.

According to Kaufman, many countries suffering from "demographic fatigue" will slip back into stage 1, resulting in both high fertility and high mortality rates. "If they do," he states, "these countries may soon reach zero population growth, but at a terrible price." He gives the example of Zimbabwe, where 26 percent of the adult population has AIDS and the average person has a life expectancy of only 40 years.

Environmentalist Lester Brown notes that 16 of the 20 countries designated as "failed states" in 2010 were caught in this demographic trap, and would most likely be unable to break out of it on their own. Brown describes Sudan as a "classic case" of a country caught in the demographic trap:

"It has developed far enough economically and socially to reduce mortality, but not far enough to quickly reduce fertility. As a result, women on average have four children, double the two needed for replacement, and the population of 41 million is growing by over 2,000 per day. Under this pressure, Sudan—like scores of other countries—is breaking down."

Examples of developing nations that successfully went from stage 2 to stage 3 are South Korea and Taiwan, which were able to move toward smaller families, and thereby improved living standards. This resulted in further reduction in fertility rates.

Other Viewpoints

The existence of the "trap" is controversial. Some demographers see it as only a temporary problem, which can be eliminated with better education and better "family planning." While others consider the "trap" more of a longer-term symptom of the failure to educate children and provide safety nets against poverty, resulting in more families seeing children as a form of "securing incomes" for the future. Nonetheless, many social scientists agree that family planning should be an important part of public health and economic development.

Central Places Theory

The central places theory was conceived, primarily by W. Christaller and A. Losch, in order to explain size and number of cities and their spacing in a territory. It relies on a definition of city that considers it essentially as a distribution centre of goods and services to a scattered population, and on optimisation principles (which take transport costs into account). This theory stands on the limit between geography and spatial economy, and may be claimed by both disciplines. The theory is basically formalised in a static way, several derived models are proposed which represent equilibrium states, but its authors have suggested tracks that should allow making it evolve.

The theory is based on a distinction between centres, which are the seats of a supply of goods and services, and peripheries (regions complementing the centre) where demand, i.e. population using them, resides. The notion of centrality justifies clustering in a same place production of services of same level and of same range intended at the population which is scattered in the complementary region (or influence area), whose customers are polarised by the centre. The centres are indeed hierarchised, due to the existence of several levels of services defined by their spatial ranges (distance that the consumer is willing to travel in order to acquire the service, defined by the additional transport cost which can be afforded when buying the product) and by emergence thresholds (fixed by the volume of customers needed for the service supply to be profitable). Frequently used and cheap services are offered in numerous small centres located close to consumers, while those less frequently used are located in cities that are larger, but also more distant. According to versions of the theory, influence areas of centres fit inside each other (for Christaller), because centres of upper level generally provide all services of lower level, or more or less apart from each other (for Losch). The hypothesis of rational behaviour of consumers, which visit the closest centre, and competition between centres that share the customers have as consequence that cities are regularly spaced, and hierarchy of services levels is

translated into a smaller number and a wider spacing of cities when moving upward in urban hierarchy.

Quite numerous observations carried on in various areas of the world have shown how useful the theory is to understand spatial organisation of most services to resident population. The theory gives well enough account of differentiation of urban networks at middle levels scales, in relatively homogeneous regions. The hierarchy of urban centres fits in large part with a hierarchy of levels of services they concentrate, organised by frequency of use, amplitude of their spatial range and size of their thresholds of emergence. The theory has been used by spatial planning, notably to implement settlement of the polders in the Netherlands, or also to justify policy of "métropoles d'équilibre" in France. It is also used as reference model by archaeologists studying ancient settlement systems. On the other side, the theory hardly allows to predict distribution of retail and services in declining rural areas, where local factors encouraging a persistent location play a more prominent role than effects of additional cost of distance, or also in an urban environment where time accessibility takes a much stronger importance than physical distance and generates configurations that are much more complex than Christaller models.

Several critics are formulated against the central places theory. Some of them question theory hypotheses :

The closest centre choice is not systematically practiced by the consumer. It has been demonstrated that, in a rather densely populated area (in the Netherlands and probably more generally in an urban environment), around 40 pc of the purchase power is spent at the occasion of « travels with multiple purposes, i.e. in places where the consumer makes at the same time provision of goods and services of lower level in an upper level centre, thus compensating a longer average distance by benefit of a more diversified supply of services. This way of doing tends to by-pass smaller centres and to increase size of the larger ones, and thus generates a stronger hierarchising of centres than is predicted by the theory.

Regular hexagonal spatial models as proposed by Christaller are invalidated from the start, as their configuration relies on hypothesis of a uniform distribution of population to cover, an hypothesis in contradiction with the existence of centres, which necessarily induce strong centre-periphery gradients in terms of population density for example. Configurations that would take this fact into account have been simulated, but geometric or analytical models could not be demonstrated yet.

The central places theory is an incomplete explanation of the hierarchised organisation of urban systems. It is based on a form of spatial organisation of production of goods and services that is strongly conditioned by the requirement of proximity between the producer and its customers, be it because of a marked sensitiveness to transport costs (craft bakery, post office) either because product is perishable (slaughterhouses, market gardening and milk production belts around cities before

diffusion of frigorific transport), either because of nature of the provided service (hairdresser, doctor). Industrial production replacing craft has loosened these links, and location and size of cities born during the Industrial revolution (mine cities, steel cities, textile cities or chemical plants) do not follow the logic of central places (even if the latter plays a role in a later stage, because presence of a population to serve has for effect that services are established in proportion: for example a University had to be created in Valenciennes). Other urban functions such as defence or harbour functions also escape the frame of central places theory. One can imagine that an economic transformation of production and distribution that would completely suppress the link of proximity between producer and customer would make the central places theory totally obsolete, and would turn it into a merely temporary explanation of organisation of urban systems, linked to a moment in their development history where distance played a fundamental role in the spatial organisation of urban activities.

The theory nonetheless maintains its strength, as numerous activities (for example business services, high technology) locate in function of the presence of urban services, and strengthen correlation between level of those services and rank of cities in urban hierarchy, defined according to city population or to production weight. The explanation then changes its focus and calls for integration of the central places theory in a more general theory that could be an evolutionary theory of urban systems.

Hierarchy

The notion of hierarchy is used with two distinct meanings.

1. It is a social, political or administrative organisation in levels where each element belonging to a level is strictly subordinated to an element of upper level. The higher it is gone in the order of power or domination, the fewer elements are comprised in each level: hierarchy implies a pyramidal organisation. The benefit of such form of organisation is to allow making information circulate or imposing decisions while reducing transmission delays, its disadvantage is some rigidity in adaptation to change. In this sense, hierarchical organisations of work in firms or in the way that groups function are sometimes opposed to “network” organisations (despite the fact that a hierarchy is a particular form of network), where connections are more numerous and which are more versatile. In this precise meaning, territories rarely show a hierarchical organisation, except for what concerns organisation of some administrative grids with their levels fitting inside each other. By extension, is called hierarchy a system organised according to a tree-like relationship (i.e. which is represented by a tree as meant in graphs theory), which defines embedded sub-systems, which however are not necessarily subordinated (in an ascending hierarchical classification for example).

2. Hierarchy also describes a form of organisation of a system into sub-systems in such a way that the number of sub-systems varies according to an inverse geometrical progression of the number of elements (size) of each sub-system. Statistical models of these hierarchies are distribution of Pareto, or the lognormal distribution (so-called of Galton-Gibrat). It is the most frequent form of organisation in nature. Some examples in geography are the size of firms in a sector of activity, the area of farms in a region, the dimension of territories in the world, the size of cities in a State... Urban hierarchy, as described through the rank-size rule, or the central places theory, is a particularly stable and universal form of organisation of settlement and activities in a territory.

Centre/Periphery

The geometrical metaphor of centre and periphery is often used to describe opposition between the two basic types of places in a spatial system : the one which is commanding and benefitting it, the centre, and those which are subjected to it, in a peripheral position. This conceptual pair goes back at least to the time of Werner Sombart (*Der moderne Kapitalismus*, 1902), if not of Marx (city / countryside relations), and was used by the imperialism theoreticians (Rosa Luxemburg, Boukharine) but its contemporaneous form was given by economists specialised in development inequalities (Samir Amin, *Le développement inégal*, 1973). Alain Reynaud developed the notion in the field of geography (*Société, espace et justice*, 1981).

It is thus rigorous to avoid using these terms in their common sense, which is used in particular about urban realities, in order to distinguish what is in the middle from what is outside. There should be particular caution about representations in form of rings fitted inside each other, which may point as well to simple discretisations (rings of increasing or decreasing densities), to successive stages (rings of urban growth, for example), to distribution of phenomena of various nature in function of a distance parameter (rings of Thunen), as to an opposition between dominating and dominated places.

The concept may be used at all levels of the geographic scale (centre and periphery of a village farmland, of a city, of a region, etc.). But it has registered a particular success at the global level, as an equivalent of the "developed world / underdeveloped world" or "North / South" pairs. Centre / Periphery allows description of the opposition of places but moreover proposes an explanatory model of this differentiation : periphery is subjected because centre is dominating-and reciprocally. This concept has thus been mostly used in the context of third-worldist thinking, more or less as an instrument for making inhabitants of Western countries feel guilty. This is an excessively restrictive use of a much more efficient notion. To think in terms of centre(s) and periphery(ies) allows to reflect on interactions between places in the World : links of reciprocal dependency where inequalities are the rule, but which are not working one-way.

For the pair to make sense, there must be relationships between the two types of places, thus flows (of persons, goods, money, information, decisions,...) and these relationships must be dissymmetrical, (unbalance of flows, hierarchy of power relationships,...). The centre is central precisely because it benefits from this inequality and, in turn, the periphery(ies) is(are) characterised by a deficit which maintains its(their) dominated position. The hereby described system is auto-regulated : the centre reproduces conditions for its centrality and the periphery does the reverse. It is thus a pleonasm to talk about a "dominated periphery". However, precisely because it is based on a logic of (inequal) exchange, the system is dynamic. Whereas some peripheries may become "dead ends" (they are then said « abandoned »), others may benefit from their situation (advantage on the long-run because of a greater size, of a location in contact with the outside of the spatial system,...); this can generate either polarity reversals in a logic that remains globally identical or system changes.

The centre / periphery model has thus a robust heuristic potential, provided it is not overused. Its use should be reserved to formalisation of any system based on inequality relationships, and it should not be used as a simple descriptor of spatial gradient or differentiation.

4

Essential Terminologies

Sequent occupance: The notion that successive societies leave their cultural imprints on a place, each contributing to the cumulative cultural landscape. This is an important concept in geography because it symbolizes how humans interact with their surroundings.

Cultural landscape: Fashioning of a natural landscape by a cultural group. This is the essence of how humans interact with nature.

Arithmetic density: The total number of people divided by the total land area. This is what most people think of as density; how many people per area of land.

Physiological density: The number of people per unit of area of arable land, which is land suitable for agriculture. This is important because it relates to how much land is being used by how many people.

Hearth: The region from which innovative ideas originate. This relates to the important concept of the spreading of ideas from one area to another (diffusion).

Diffusion: The process of spread of a feature or trend from one place to another over time.

Relocation diffusion: The spread of an idea through physical movement of people from one place to another. Ex: spread of AIDS from New York, California, & Florida.

Expansion diffusion: The spread of a feature from one place to another in a snowballing process. This can happen in 3 ways:

Hierarchical diffusion: The spread of an idea from persons or nodes of authority or power to other persons or places (Ex: hip-hop/rap music).

Contagious diffusion: The rapid, widespread diffusion of a characteristic throughout the population. (Ex: ideas placed on the internet).

Stimulus diffusion: the spread of an underlying principle, even though a characteristic itself apparently fails to diffuse. (Ex: PC & Apple competition).

Absolute distance: Exact measurement of the physical space between two places.

Relative distance: Approximate measurement of the physical space between two places.

Distribution: The arrangement of something across Earth's surface.

Environmental determinism: A 19th-and early 20th-century approach to the study of geography that argued that the general laws sought by human geographers could be found in the physical sciences. Geography was therefore the study of how the physical environment caused human activities.

Absolute location: Position on Earth's surface using the coordinate system of longitude (that runs from North to South Pole) and latitude (that runs parallel to the equator).

Relative location: Position on Earth's surface relative to other features. (Ex: My house is west of 394).

Site: The physical character of place; what is found at the location and why it is significant.

Situation: The location of a place relative to other places.

Space Time Compression-The reduction in the time it takes to diffuse something to a distant place, as a result of improved communications and transportation system.

Friction of Distance-is based on the notion that distance usually requires some amount of effort, money, and/or energy to overcome. Because of this "friction," spatial interactions will tend to take place more often over shorter distances; quantity of interaction will decline with distance.

Distance Decay-The diminishing in importance and eventual disappearance of a phenomenon with increasing distance from its origin. Typically, the farther away one group is from another, the less likely the two groups are to interact. (Electronic devices such as the internet and e-mail have aided in eliminating barriers to interaction between people who are far from each other.

Networks-defined by Manuel Castells as a set of interconnected nodes without a centre.

Connectivity-The relationships among people and objects across the barrier of space. Geographers are concerned with the various means by which connections occur.

Accessibility-The degree of ease with which it is possible to reach certain location from other locations. Accessibility varies from place to place and can be measured.

Space- Refers to the physical gap or interval between two objects.

Spatial Distribution-Physical location of geographic phenomena across SPACE

Size-Is the estimation or determination of extent.

Scale-Representation of a real-world phenomenon at a certain level of reduction or generalization. In cartography, the ratio of map distance to ground distance, indicated on a map as a bar graph, representative fraction, and/or verbal statement.

Formal Region-(uniform) or homogenous region is an area within which everyone shares in common one or mare distinctive characteristics. The shared feature could be a cultural value such as a common language, or an environmental climate.

Functional Region- (nodal region) Area organized around a node or focal point. The characteristic chosen to define a functional region dominates at a central focus or node and diminishes in importance outward. This region is tied to the central point by transportation or communication systems or by economic or functional associations.

Vernacular Region-(Perceptual Region) is a place that people believe exists as a part of their cultural identity. Such regions emerge from peoples informal sense of place rather than from scientific models developed through geographic thought. (Often identified using a mental map-which is an internal representation of a portion of Earths surface).

Possibilism- The physical environment may limit some human actions, but people have the ability to adjust to their environment.

Natural Landscape- (xxx).

Pattern-A common property of distribution, which is the geometric arrangement of objects in space. Some features are organized in a geometric pattern, whereas others are distributed irregularly. Geographers observe that many objects form a linear distribution, such as the arrangement of houses along a street or stations along a subway line.

Place Name-Often referred to as a places toponym (the name given to a place on Earth.

Population – Migration & Dispersion

Age Distribution: (Population pyramid) is two back-to-back bar graphs, one showing the number of males and one showing females in a particular population in five-year age groups. This is important because you can tell from the age distribution important characteristic of a country, whether high guest worker population, they just had a war or a deadly disease and more.

Carry capacity: This is the population level that can be supported, given the quantity of food, habitat, water and other life infrastructure present. This is important because it tells how many people an area will be able to support. **Cohort**: Population of various age categories in an age-sex population pyramids. This is important because

this can tell what state this country it is whether in Stage 3 or Stage 5 in the demographic transition model.

Demographic equation: The formula that calculates population change. The formula finds the increase (or decrease) in a population. The formula is found by doing births minus deaths plus (or minus) net migration. This is important because it helps to determine which stage in the demographic transition model a country is in.

Demographic momentum: this is the tendency for growing population to continue growing after a fertility decline because of their young age distribution. This is important because once this happens a country moves to a different stage in the demographic transition model.

Demographic regions: Cape Verde is in Stage 2 (High Growth), Chile is in Stage 3 (Moderate Growth), and Denmark is in Stage 4 (Low Growth). This is important because it shows how different parts of the world are in different stages of the demographic transition.

Demographic Transition model: Has 5 steps. Stage 1 is low growth, Stage 2 is High Growth, Stage 3 is Moderate Growth, and Stage 4 is Low Growth and Stage 5 although not officially a stage is a possible stage that includes zero or negative population group. This is important because this is the way our country and others countries around the world are transformed from a less developed country to a more developed country.

Dependency ratio: The number of people who are too you or too old to work compared to the number of people in their productive years. This is important because this tells how many people each worker supports. For example the larger population of dependents, the greater financial burden on those who are working to support those who cannot.

Diffusion of fertility control: The diffusion of fertility control is spread throughout the world. In the U.S it's below 2.1 in much of Africa it is above 4, if South America is between 2 and 3, in Europe it is below 2.1, in China and Russia it is below 2.1, and in much of the Middle East it is above 4. This is important because its shows how many kids a mother is having thus helping to see where the countries are growing rapidly and where countries are levelling off.

Disease diffusion: There are two types, contagious and hierarchical. Hierarchical is along high density areas that spread from urban to rural areas. Contagious is spread through the density of people. This is important in determining how the disease spread so you can predict how it will spread.

Doubling time: The number of years needed to double a population, assuming a constant rate of natural increase. This is important because it can help project the countries population increase over the years and when its population will double.

Ecumene: The proportion of earths surface occupied by permanent human settlement. This is important because its tells how much of the land has been built upon and how much land is left for us to build on.

Epidemiological transition model: This is a distinctive cause of death in each stage of the demographic transition. This is important because it can explain how a countries population changes so dramatically and more.

Gendered Space

Infant mortality rate: (IMR) The annual number of deaths of infants under one year of age, compared with total live births. Its is expressed as the annual number of deaths among infants among infants per 1000 births rather than a percentage. This is important because it tell how developed a country is, if they have a high IMR they are an LDC and if it is low they are an MDC.

J-curve: This is when the projection population show exponential growth; sometimes shape as a j-curve. This is important because if the population grows exponential our resource use will go up exponential and so will our use as well as a greater demand for food and more.

Maladaptation: This is an adaptation that has become less helpful than harmful. This relates to human geography because it has become less and less suitable and more of a problem or hindrance in its own right, as time goes on. Which shows as the world changes so do the things surrounding it.

Malthus, Thomas: Was one of the first to argue that the worlds rate of population increase was far outrunning the development of food population. This is important because he brought up the point that we may be outrunning our supplies because of our exponentially growing population.

Mortality: There are two useful ways to measure mortality; infant mortality rate and life expectancy. The IMR reflect a country's health care system and life expectancy measures the average number of years a baby can expect to live. This is important because you can use a countries mortality rate to determine important features about a country.

Natality: (Crude Birth Rate) This is the ratio of live births in an area to the population of that area; it is expressed as number of birth in year to every 1000 people alive in the society. This is important because it tells you the rate a country is having babies as well as how fast you can expect that population to grow.

Neo-malthusian-theory that builds upon Malthus' thoughts on overpopulation. Takes into count two factors that Malthus did not: population growth in LDC's, and outstripping of resources other than food.

Recognizes that population growth in LDC's is from the transfer of medical talents from MDC's but not the wealth that would provide food and resources.

Overpopulation-relationship between the number of people on Earth, and the availability of resources.

Problems result when an area's population exceeds the capacity of the environment to support them at an acceptable standard of living.

Population density-the frequency with which something occurs in space is density.

* Arithmetic density: total number of objects in an area. *Used to compare distribution of population in different countries.*
* Physiological density: number of persons per unit of area suitable for agriculture. *Could mean a country has difficulty growing enough food*.
* Agricultural density: the number of farmers per unit of area of farmland. *May mean a country has inefficient agriculture.*

Population distributions-the arrangement of a feature in space is distribution. Geographers identify the three main properties as density, concentration, and pattern.

Used to describe how things and people are distributed across the earth.

Population explosion-a sudden increase or burst in the population in either a certain geographical area or worldwide.

Occurred in the late 18th and early 19th centuries because several countries moved on to stage 2 of the DTM. Can trace factors that lead to these explosions.

Population projection-predicts the future population of an area or the world.

Helps predict future problems with population such as overpopulation or under population of a certain race or ethnicity.

Population pyramid-population displayed by age and gender on a bar graph.

Shape is determined primarily by crude birth rate. Shows age distribution and sex ratio.

Rate of natural increase-the percentage by which a population grows in a year.

CBR-CDR = NIR Excludes migration.

Affects the population and a country's or area's ability to support that population.

S-curve-traces the cyclical movement upwards and downwards in a graph. So named for its shape as the letter "s".

Relates to growth and decline in the natural increase.

Sex ratio-the number of males per hundred females in the population.

Depends on birth and death rates, immigration. Men have higher death rates but also higher birth rates. Immigration usually means more males because they can make the journey.

Standard of living-refers to the quality and quantity of goods and services available to people and the way they are distributed within a population.

Higher standards of living are found in MDC's rather than LDC's. Can help trace development.

Sustainability-providing the best outcomes for human and natural environments both in the present and for the future.

Relates to development that meets today's needs without compromising the ability of future generations to meet their own needs.

Underpopulation-it is the opposition to overpopulation and refers to a sharp drop or decrease in a region's population.

Unlike overpopulation, it does not refer to resources but to having enough people to support the local economic system. If there are not enough tax payers, then the area cannot continue.

Zero population growth-when the crude birth rate equals the crude death rate and the natural increase rate approaches zero.

Often applied to countries in stage 4 of the demographic transition model.

Activity space-space allotted for a certain industry or activity.

Can apply to an area within a city or surrounding a central place.

Chain migration-when one family member migrates to a new country and the rest of the family follows shortly after.

Mostly seen from Mexico to the United States when guest workers set up homes and make money for their family to follow them.

Cyclic movement-trends in migration and other processes that have a clear cycle.

Distance Decay-When contact between two groups diminishes because of the distance between them.

Forced Migration-People removed from there countries and forced to live in other countries because of war, natural disaster, and government.

Gravity Model-Predicts that the optimal location of a service is directly related to the number of people in the area and inversely related to the distance people must travel to access it.

Internal Migration-Permanent movement within a particular country.

Intervening Opportunity-An environmental or cultural feature of the landscape that helps migration.

Migration Patterns

Intercontinental-Permanent movement from one country to a different country on the same continent.

Interregional-Permanent movement from one region of the country to another.

Rural-Urban-Permanent movement from suburbs and rural area to the urban city area.

Migratory Movement-

Periodic Movement-

Personal Space.

Place Utility-

Push-Pull Factors-Factors that induce people to leave old residence and move to new locations.

Refugee-People forced to migrate from their home country and cannot return for fear of persecution because of their race, religion, nationality, membership in social group, or political opinion.

Space-Time Prism-

Step Migration.

Transhumance-Seasonal migration of live stock between mountains and lowland pasture areas.

Transmigration-

Voluntary.

Cultural Patterns & Processes

Acculturation: Process of adopting only certain customs that will be to their advantage.

Assimilation: Process of less dominant cultures losing their culture to a more dominant culture.

Cultural Adaptation

Cultural core/periphery pattern: The core-periphery idea that the core houses main economic power of region and the outlying region or periphery houses lesser economic ties.

Cultural Ecology: The geographic study of human environmental relationships.

Cultural Identity: Ones belief in belonging to a group or certain cultural aspect.

Cultural Landscape: The visible imprint of human activity on the landscape.

Cultural Realm

Culture: The body of customary beliefs, social forms, and material traits that together constitute a group of people's distinct tradition.

Culture Region

Formal (Uniform): An area in which everyone shares in one or more distinctive characteristics.

Core-Centre of economic activity.

Periphery-Outlying region of economic activity.

Functional (Nodal): Region organized at a node or focal point.

Vernacular (perceptual-regional self-awareness): A place that people believe exists as part of their cultural identity.

Diffusion Types

Expansion-The spread of one feature from one place to another in a snowballing process.

Hierarchical-The spread of an idea from persons or nodes of authority or power to other persons or places.

Contagious-The rapid widespread diffusion of a characteristic throughout the population.

Stimulus-The spread of an underlying principle when the characteristic fails to diffuse.

Relocation-The spread of an idea through physical movement of people from one place to another.

Innovation Adoption: Study of how why and at what rate new technology spreads throughout a culture.

Maladaptive diffusion: Diffusion of a process with negative side effects or What works well in one region may not in another.

Sequence Occupancy: Refers to such cultural succession and its lasting imprint proposed by Derwent Whittlesey.

Religion-the faithfulness to codified beliefs and rituals that generally involve a faith in a spiritual nature. This is important to HG because man wars have been fought over it.

Animism: Belief that objects, such as plants and stones, or natural events, like thunderstorms and earthquakes, have a discrete spirit and life. This is important to Human Geography because a lot of cultures around the world believe in Animism.

Buddhism: The third of the world's major universalizing religions. It has 365 million adherents especially in China and Southeast Asia. It is important because a large percent of the earth's population follow Buddhism beliefs.

Cargo Cult Pilgrimage-Cargo Cult's believe western goods have been traded to

them by ancestral spirits. It takes place in Melanesia and is important go HG because it's a big religious movement by a large number of people.

Christianity-is a monotheistic religion centered on the life and teachings of Jesus of Nazareth as presented in the New Testament. It's important to HG because it's the most popular religion in the world.

Confucianism-Developed by earlier Chinese man Confucius, it's a complex system of moral, social, political, and religious thought. This is important to HG because it has affected Chinese Civilizations tremendously.

Ethnic Religion-A religion with a rather concentrated distribution whose principles are likely to be based on the physical characteristics of the particular location where its adherents are located. This is important to HG because most religions start off as a Ethnic Religion.

Exclave/Enclave-A enclave is a country or part of a country mostly surrounded by the territory of another country; an exclave is one which is geographically separated from the main part by surrounding alien territory. This is important to HG because a lot of countries are within other countries.

Fundamentalism-Literal interpretation and strict adherence to basic principles of a religion. This is important to HG because there are a lot of Fundamentalists in all religions.

Geomancy-is a method of prediction that interprets markings on the ground, or how handfuls of dirt land when someone tosses them. The Arabic tradition consists of sketching sixteen random lines of dots in sand. This is important to HG because most farmers use a form of Geomancy.

Hajj-The pilgrimage to Mecca for Islam followers. It's the fifth of the five pillars. It is important to HG because just about all Islam followers try the pilgrimage there.

Hinduism-Created in India, approximately one billion followers. Unlike other religions, heaven isn't always the ultimate goal in life. Third largest in world behind Christianity and Islam. Talk about Karma (what goes around comes around.) It is important to HG because such a large number of people follow the religion and it's unlike any other one.

Interfaith boundaries-the boundaries between the world's major faiths, such as Christianity, Muslim, and Buddhism. This isn't the same as Intrafaith boundaries which describes the boundaries within a major religion. This is important to HG because it separates different groups of people for different reasons.

Islam-It means the submission to the will of god. Its a monotheistic religion originating with the teachings of Muhammad, a key religious figure. It is the second largest religion in the world. This is important to HG because it has impacted the world greatly, especially boundaries.

Jainism-religion and philosophy originating in ancient India. Stresses spiritual independence and equality throughout all life. It affects HG because a lot of people believe in it in India.

Judaism-It is the religion of ancient Hebrews, said to be one of the first monotheistic faiths. This is important to HG because many other religions have been based off it.

Landscapes of the dead-The certain areas where people have commonly been buried. This is important to HG because it has always been important where people are buried.

Monotheism/polytheism-Monotheism this is the belief in one god and polytheism is the belief in many gods. This affects HG because many religions spread throughout the world fall under these two categories.

Mormonism: a term used to describe religious, ideological, and cultural aspects of the various denominations of the Latter Day Saint movement. It is important because a lot of people around the world practice Mormonism.

Muslim pilgrimage: If physically and financially able, a Muslim makes a pilgrimage to Makkah. (Mecca) They usually make the trip around Ramadan. This pilgrimage is also referred to as Hajj. It is important because Islam is one of the most popular religions practiced around the world.

Muslim population: It is the religion of 1.3 billion people in the world. It is the predominant religion of the Middle East from North Africa to Central Asia. Half of the world's Muslims live in four countries outside the Middle East: Indonesia, Pakistan, Bangladesh, and India. It is important because Islam is one of the most popular religions practiced around the world.

Proselytic Religion: Referred to as a Universalizing Religion, which is an attempt to be global, to appeal to all people, wherever they may live in the world, not just to those of one culture or location. There are three religions that practice this they are Christianity, Islam, and Buddhism. To proselytize is to try to convert another person to your religion. This important to HG because these are three of the biggest religions in the world they are practiced all over the world.

Reincarnation: The idea of reincarnation is that after this life you will come back in another life either as a plant, animal, or a human life. So basically what you do in this life will affect what your next life is like. This is commonly practiced by the Buddhists and the Hindus. This is important to HG because these two religions are very important in the world.

Religion (groups, places): One group is universalizing religions. These are Christianity, Islam, and Buddhism. All of these have different branches. There's also ethnic religions, such as, Hinduism, Daoism, and Confucianism. These religions are spread out throughout the world. This affects HG because all regions throughout the world have a general religion.

Religious architectural styles: These are the styles of architecture created by the religions. For example, Christians have always made temples, and Buddhists have always made a lot of religious statues. This is important to HG because these styles affected most of the future styles for other civilizations.

Religious Conflict-this is the conflicts between religions. One of these is Israel-Palestine. This consists of Roman Takeovers, Muslim conquests, and the crusades. This affects HG because there has been a lot of bloodshed over Religious Conflict.

Religious Culture Hearth: This is where most religions are born. Most major religions have come from the Middle East near Israel, but a few have come from India too. This is important to HG because where religions are created, civilizations are too.

Religious toponym: This refers to the origin and meaning of the names of religions. This is important to HG because many names mean significant things including beliefs of cultures.

Sacred space-Sacred space is the place where religious figures and congregations meet to perform religious ceremonies. This is important to HG because a lot of history has taken place at sacred spaces.

Secularism-This is the belief that humans should be based on facts and not religious beliefs. This is important to HG because this has caused conflicts in a lot of different places including politics.

Shamanism-This is the range of traditional beliefs and practices that claim the ability to cure, heal, and cause pain to people. This is important to HG because it is thought as good and bad.

Sharia law-it is the legal framework within which public and some private aspects of life are regulated for those living in a legal system based on Muslim principles. This is important to HG because it affects many people around Muslims around the world.

Shintoism-said to be the way of god. It is the native religion of Japan and was once its state religion. It involves the worship of *kami* (a god). Not very significant anymore and lost importance to today. This is important to HG because before WWII it was very popular and affected a lot of people in Japan.

Sikhism-is a religion that began in sixteenth century Northern India. The principal belief in Sikhism is faith in *Vâhigurû*. Emphasizes faith in god. This is important to HG because its another minor religion in India that affects a lot of people.

Political Organization of Space

Annexation: Incorporation of a territory into another geo-political entity.

Antarctica: Southernmost continent in the world. It has no permanent residents and doesn't belong to any country.

Apartheid: Afrikaans for apartness, it was the segregation of blacks in South Africa from 1948 to 1994. It was created to keep the white minority in power and allow them to have almost total control over the black majority.

Balkanization: The political term used when referring to the fragmentation or breakup of a region or country into smaller regions or countries. The term comes from the Balkan wars, where the country of Yugoslavia was broken up in to six countries between 1989 and 1992. It was the effect of the Balkan wars.

Border Landscape: There are two types, exclusionary and inclusionary. Exclusionary is meant to keep people out, such as the border between the U.S. and Mexico. Inclusionary is meant to facilitate trade and movement, such as the U.S.-Canada border.

Boundary disputes: Conflicts over the location, size, and extent of borders between nations. There is conflict over where exactly the border is between the U.S. and Mexico, especially along the Rio Grande because the river has changed course and moved, and it is the traditional border.

Boundary origin: Boundaries often originated from old tribal lands and lands won in war. They were meant to establish claims to land and were often smaller historically.

Boundary process: The process of creating boundaries.

Boundary type: Many boundaries are natural boundaries, formed by rivers, mountains, etc. There are also political boundaries. These are often formed through war and compromise in treaties and agreements. In countries often form cultural boundaries that used to belong to a groups cultural homeland. However, countries in Africa, the Middle East and elsewhere aren't arranged by culture but politics, and Western countries turned their former colonies into nations without respect for culture.

Buffer state: A country lying between two more powerful countries that are hostile to each other. An example is Mongolia, which serves as a buffer between Russia and China.

Capital: Principle city in a state or country. The best place to locate a capital is at the centre of a country, so it is a somewhat equal distance from all parts of the country.

Centrifugal: Religious, political, economic, conflict, etc. that causes disunity in a state.

Centripetal: An attitude that unifies people and enhances support for the state.

City-state: A region controlled by a city and that has sovereignty. They were more common in the middle ages and Renaissance in Europe.

Colonialism: The attempt by a country to establish settlements and impose political and economic control and principles. It was a big thing in the 17th through

20th century for countries in Europe to take areas around the world and make them into colonies.

Confederation: association of sovereign states by a treaty or agreement. It deals with issues such as defence, foreign affairs, trade, and a common currency.

Conference of Berlin: Regulated trade and colonization in Africa. It formalized the scramble to gain colonies in Africa and set up boundaries for each country's colonies.

Core/periphery: Core countries have high levels of development, a capacity at innovation and a convergence of trade flows. Periphery countries usually have less development and are poorer countries.

Decolonization: Decolonization is the movement of American / European colonies gaining independence. Some were peaceful struggles while others became violent.

Devolution: Devolution is the both the decentralization of a government from a unitary to a federal system or a fracturing of a government like Balkanization.

Domino theory: Domino theory is the idea that if one land in a region came under the influence of Communists, then more would follow in a domino effect. The domino theory was used by successive United States administrations during the Cold War, to justify American intervention around the world.

Exclusive Economic Zone: An Exclusive Economic Zone (EEZ) is a sea zone over which a state has special rights over the exploration and use of marine resources. The country that controls the EEZ has rights to the fishing, whaling, etc., as well as the raw material resources.

Electoral regions: Electoral regions are the different voting districts that make up local, state, and national regions.

Enclave/exclave: An enclave is a country or part of a country mostly surrounded by the territory of another country or wholly lying within the boundaries of another country (Lesotho). An exclave is a country which is geographically separated from the main part by surrounding alien territory (Azerbaijan).

Ethnic conflict: An ethnic conflict is a war between ethnic groups often as a result of ethnic nationalism or fight over natural resources. Ethnic conflict often includes genocide. It can also be caused by boundary disputes.

European Union: The European Union (EU) is a supranational and intergovernmental union of 27 democratic member states of Europe. The EU's activities cover most areas of public policy, from economic policy to foreign affairs, defence, agriculture and trade. The European Union is the largest political and economic entity on the European continent, with around 500 million people and an estimated GDP of US$13.4 trillion.

Federal: Federalism is a political philosophy in which a group or body of members are bound together with a governing representative head. Federalism is the system in which the power to govern is shared between the national & state governments.

Forward capital: A forward capital is a symbolically relocated capital city usually because of either economic or strategic reasons. A forward capital is sometimes used to integrate outlying parts of a country into the state. An example would be Brasília.

Frontier: A frontier is a zone where no state exercises complete political control. It is usually uninhabited or sparsely inhabited. It separates countries where a boundary cannot be found. A current example can be found between Saudi Arabia and Yemen.

Geopolitics: Geopolitics is the study that analyses geography, history and social science with reference to international politics. It examines the political and strategic significance of geography, where geography is defined in terms of the location, size, and resources of places.

Gerrymander: Gerrymandering is the process of redrawing legislative boundaries for the purpose of benefiting the political party in power. The process is usually used to turn "too close to call" states into a party's favour.

Global commons: Global commons is that which no one person or state may own or control and which is central to life. A global common contains an infinite potential with regard to the understanding and advancement of the biology and society of all life. (Forests, oceans, land mass and cultural identity).

Heartland/rimland: Heartland is the central region of a country or continent; especially a region that is important to a country or to a culture. Rimland is the maritime fringe of a country or continent.

Immigrant state: An immigrant state is a type of receiving state which is the target of many immigrants. Immigrant states are popular because of their economy, political freedom, and opportunity. One example would be the USA.

Agricultural & Rural Land Use

Adaptive Strategies

Agrarian: People or societies that are farmers therefore promote agricultural interest ext.

-Where agrarian people and societies are located is not generally near cities ext. but these types of people are essential to the way that we live and our ability to live in cities.

Agribusiness: Commercial agriculture characterized by integration of different steps in the food-processing industry, usually through ownership by large corporations.

It influences how things are grown and what people eat.

Agricultural Industrialization: The use of machinery in agriculture, like tractors ext.

Makes it a lot faster for farmers to yield crop.

Agricultural landscape: The land that we farm on and what we choose to put were on our fields.

Effects how much yield one gets from their plants.

Agricultural Location Model

Agricultural Origins: Through time nomadic people noticed the growing of plants in a cycle and began to domesticate them and use for there own use. Carl Sauer points out vegetative planting and seed agriculture as the original forms. He also points out that vegetative planting likely was originated in SE Asia and seed agriculture originated in W. India, N. China and Ethiopia.

Without the development of agriculture we would still have a relatively small and likely uneducated population.

Agriculture: The deliberate effort to modify a portion of Earth's surface through the cultivation of crops and the raising of livestock for subsistence or economic gain.

It has influenced the growth of areas and human society.

Animal Domestication: Domestication of animals for selling or using byproducts.

Helped us obtain meat with out having to go out and kill our food right before dinner.

Aquaculture: The cultivation of aquatic organisms especially for food.

Allowed us to use the sea and its abundant sources of food for our benefit.

Biorevolution: The revolution of biotechnology and the use of it in societies.

Biotechnology: Using living organisms in a useful way to produce commercial products like pest resistant crops.

Has helped the farmers grow a more bountiful harvest through the using of pesticides ext.

Collective Farm

Commercial Agriculture (intensive, extensive): Agriculture undertaken primarily to generate products for sale off the farm.

Allowed people to move away from farms-fuelled industrial revolution.

Core/Periphery: The areas in the world that include MDCs are called the core and the area of the world that contains the LDCs is referred to as the periphery.

Crop Rotation: The practice of rotating use of different fields from crop to crop each year, to avoid exhausting the soil.

Takes up large areas of land but keeps land usable for future generations.

Cultivation regions: Regions were there is agricultural activity.

Areas with agricultural activity generally are not a place were a big city would be located-affects locations of different areas.

Dairying: The "farming" and sale/distribution of milk and milk products.

Gets us calcium, allows for people to move to the city because there is a way of getting milk or milk products.

Debt-for-nature swap: When agencies such as the World Bank make a deal with third world countries that they will cancel their debt if the country will set aside a certain amount of their natural resources.

Diffusion: The process of spread of a feature or trend from one place to another over time.

Influences the development of some regions faster than others.

Double Cropping: Harvesting twice a year from the same land.

Can cause agricultural exhaustion making people move away from the land.

Economic activity (primary, secondary, tertiary, quaternary, quinary):

Primary: Involves jobs like lumber and mining.

Secondary: Manufacturing products and assembling raw materials.

Tertiary: The service sector that provides us with transportation, communication and utilities.

Quaternary:

Quinary:

All of these jobs are necessary in the world.

Environmental Modifications (pesticides, soil erosion, desertification): The destruction of the environment for the purpose of farming. (Using pesticides that drain in to the water and soil and pollute them overuse of land causing the desert like conditions of desertification (dust bowl).

Doing harm to the environment through pesticides and causing desertification have horrible long term effect on humans and their future.

Extensive subsistence agriculture (shifting cultivation, nomadic herding/pastoralism):

Shifting Cultivation: Use many fields for crop growing each field is used for a couple years then left fallow for a relatively long time.

Nomadic herding/pastorilism: Based on herding domesticated animals.

Effect the way that some in the world to live and were they fall in demographic transition.

Extractive Industry:

Farm crises:

Farming: see agriculture.

Feedlot: a plot of land on which livestock are fattened for market.

Essential to how we live and eat today-necessity for most people's diets.

First agricultural revolution: Around 8000 B.C. when humans first domesticated plants and animals.

This allowed for future generations to grow larger because they no longer we just a hunter gatherer society.

Fishing – The technique, occupation, or diversion of catching fish. Fishing provides a food source and employment to society.

Food Chain – A series of organisms interrelated in their feeding habits, the smallest being fed upon by a larger one, which in turn feeds a still larger one, etc.

Forestry – The science of planting and taking care of trees and forests. Trees provide building materials and fuel to society.

Globalized Agriculture – Diffusion of agriculture across the globe.

Green Revolution – Rapid diffusion of new agricultural technology, especially new high-yield seeds and fertilizer. Because of Green Revolution, agricultural productivity at a global scale has increased faster than the population.

Growing Season – The season in which crops grow best. Growing season can vary by location, societies rely on their growing season to which crops they can or can't grow at their latitude.

Hunting and Gathering – Before the agriculture, humans gained food by hunting for animals, fishing, or gathering plants. They lived in small groups (less than 50 people), travelled frequently following game and seasonal growth of plants.

Intensive Subsistence Agriculture – A form of subsistence agriculture in which farmers must expend a relatively large amount of effort to produce the maximum feasibly yield from a parcel of land. Popular in East, South, and Southeast Asia, because the ratio between farmers and arable land is so high, most of the work is done by the family by hand or by animal with processes refined over thousands of years.

Intertillage – Tillage between rows of crops of plants.

Livestock Ranching – commercial grazing of livestock over an extensive area. Practiced is semi-arid or arid land, where vegetation is too sparse or the soil to too

poor to support crops. Prominent in later 19th century in the American West; ranchers free roamed throughout the West, until the U.S. government began selling land to farmers who outlined their farms with barbed wire, forcing the ranchers to establish large ranches to allow their cattle to graze.

Market Gardening – The small scale production of fruits, vegetables, and flowers as cash crops sold directly to local consumers. Distinguishable by the large diversity of crops grown on a small area of land, during a single growing season. Labour is done manually.

Mediterranean Agriculture – Farming in the land surrounding the Mediterranean Sea (Southern Europe, North Africa, and Western Asia), also in lands with similar climates (California, central Chile, Southwestern South Africa, and Southwestern Australia). Sea winds provide moisture and moderate winter; land is hilly with mountains frequently plunging directly into sea. Growing fruits, vegetables, flowers, and tree crops are the main crops, while animals are grown under transhumance – kept on coastal plains in winter and moved to hills in the summer.

Mineral Fuels – Natural resources containing hydrocarbons, which are not derived from animal or plant sources.

Mining – Extraction of valuable minerals or other geological materials from the Earth, usually from an ore body, vein, or coal seam. Any material that cannot be grown from agricultural processes, or created artificially, is mined (mining in a wider sense then including extraction of petroleum, natural gas, and water).

Planned Economy –Economic system in which a single agency makes all decisions about the production and allocation of goods and services. Commonly used in which state or government controls the factors of production and makes all decisions about their use and about the distribution of income. Example: Economy of the Soviet Union, in the 80's and 90's government presiding over planned economies began deregulating and moving toward market basted economies by introducing market forces to determine pricing, distribution, and production. Today most economies are market or mixed economies, except those in Cuba or North Korea.

Renewable – Energy replaced continually within a human lifespan, has an essentially unlimited supply and is not depleted when used by people. Solar energy, hydroelectric, geothermal, fusion and wind, are the most widely used.

Non-Renewable – Energy formed so slowly that for practical purposes it cannot be renewed. The three main fossil fuels (petroleum, natural gas, and coal) plus nuclear energy are the most widely used, mostly because they are more cost efficient.

Rural Settlement – Sparsely settled places away from the influence of large cities. Live in villages, hamlets on farms, or in other isolated houses. Typically have an agricultural character, with an economy based on logging, mining, petroleum, natural gas or tourism.

Dispersed –Characterized by farmers living on individual farms isolated from neighbours rather than alongside other farmers in the area.

Nucleated – a number of families live in close proximity to each other, with fields surrounding the collection of houses and farm buildings.

Building Material – houses and buildings are typically built from materials that are abundant in the area.

Village Form –

Sauer, Carl O. – defined cultural landscape, as an area fashioned from nature by a cultural group. A combination of cultural features such as language and religion; economic features such as agriculture and industry; and physical features such as climate and vegetation. "Culture is the agent, the natural area is the medium, the cultural landscape is the result."

Second Agricultural Revolution – Precursor to Industrial Revolution in the 19th century, that allowed a shift in work force beyond subsistence farming to allow labour to work in factories. Started in United Kingdom, Netherlands, and Denmark, especially with the Enclosure Act, which consolidated land in Great Britain. Potatoes and corn diffused from America's to Europe, and other resources followed from colonial possessions to Europe.

Specialization – Third level of cities (behind World Cities, and Command and Control Centres), offer a narrow and highly specialized variety of services. Typically specialize in management, research and development of a specific industry (motor vehicles in Detroit), or are centres of government and education, notably state capitals that also have a major university (Albany, Lansing, Madison, or Raleigh-Durham).

Staple Grains – Maize, wheat, and rice are the most produced grains produced world wide, accounting for 87% of all grains and 43% of all food. Maize staple food of North America, South American, Africa, and livestock worldwide, wheat is primary in temperate regions, and rice in tropical regions.

Suitcase Farm –Individuals who live in urban areas a great distance from their land and drive to the country to care for their crops and livestock. This practice lends itself well to the growth of wheat. Allows families to continue their long relationships with the ancestral farm, but still enjoy the benefits of waged incomes in urban environments.

Survey Patterns

Long Lots (French) – Houses erected on narrow lots perpendicular along a river, so that each original settler had equal river access.

Metes and Bounds (English) – Uses physical features of the local geography, along with directions and distances, to define the boundaries of a particular piece of

land. Metes refers to boundary defined by a measurement of a straight run, bounds refers to a more general boundary, such as a waterway, wall, public road, or existing building.

Township-and-Range (U.S.A) – Survey's used west of Ohio, after the purchase of the Louisiana Purchase. Land is divided into six-mile square blocks (township), which is then divided into one-mile square blocks (range). Ranges were then broken into smaller parcels to be sold or given to people to develop.

Sustainable Yield – Ecological yield that can be extracted without reducing the base of capital itself, the surplus required to maintain nature's services at the same or increasing level over time. Example, in fisheries the basic natural capital decreases with extraction, but productivity increases; so the sustainable yield is within the ranch that the natural capital together with production are able to provide satisfactory yield.

Third Agricultural Revolution –'Green Revolution' Rapid diffusion of new agricultural techniques between 1970's and 1980's, especially new high-yield seeds and fertilizers. Has caused agricultural productivity at a global scale to increase faster than population growth.

Mechanization – Farmers need tractors, irrigation pumps, and other machinery to make the most effective use of the new miracle seeds. Farmer's in LDC's cannot afford this machinery or the fuel to run the equipment, so governments must allocate funds to subsidizing the cost of seeds, fertilizers and machinery.

Chemical Farming – Increased use of fertilizers with nitrogen, phosphorus, and potassium. The development of higher-yield crops has produced: a 'miracle wheat seed" which is shorter and stiffer, less sensitive to variation in day length, responds better to fertilizers, and matures faster; a similar miracle rice seed, that was heartier and has increased yields; a high-yield corn seed is currently being developed.

Food Manufacturing – the Green Revolution has increased production to avoid widespread famine. Allowing the world population to grow about four billion since stared, also allowing populations in developing nations to consume 25% more than before. This increase in diets is questioned by the content in diets; Asian farmers are eating more rice than fish and other vegetables because they can rely on rice to grow efficiently.

'Tragedy of the Commons' – social trap that involves a conflict over resources between interests and the common good.

Transhumance –pastoral practice of seasonal migration of livestock between mountains and lowland pasture areas.

Truck Farm – Commercial gardening and fruit farming, so named because truck was a Middle English word meaning bartering or the exchange of commodities. Predominant in Southeastern U.S.A, because of the long growing season and humid

climate, accessibility to large markets of New York, Philadelphian, and Washington. Truck farms grow many of the fruits and vegetables that consumers demand in developed societies. Truck farms sell some of their product to fresh markets, but mostly to large processors for canning or freezing. Truck farms are highly efficient and large-scale operations that take full advantage of machines at every stage of the growing process.

Von Thunen, Johann Heinrich – 1826, Northern Germany. When choosing an enterprise, a commercial farmer compares two costs; cost of the land versus the cost of transporting production to market. Identifies a crop that can be sold for more than the land cost, distance of land to market is critical because the cost of transporting varies by crop.

Also found that specific crops were grown in varying rings around city. Market-oriented gardens and milk producers in first ring, because of expense of transportation and perishability. In the next rings wood lots used for construction and fuel, because it is a heavy industry with high transportation costs. Next rings are used for various crops or pasture, with the outermost ring devoted to animal grazing. Von Thunen's theory disregards site or human factors.

Industrialization & Development

Agricultural labour force is the number of people who work in agriculture. This is important because a large value indicates that the country is likely an LDC dependent on agriculture, while a small value indicates that there are fewer people working in agriculture, meaning that the agriculture is more efficient.

Calorie consumption as a percentage of daily requirement is an important index of development. People in MDCs generally consume more than 130% of their daily requirements, but most people in LDCs barely get enough to sustain themselves. The problem is worst in Africa, where most people do not eat enough.

The **Core-periphery model** describes the pattern of distribution of the MDCs and LDCs. When the earth is viewed from the North Pole, the MDCs are clustered near the centre of the map while the LDCs are near the edges.

Cultural Convergence is the change in culture that occurs as diffusion of ideas and technology increases. An example is the culture of LDCs becoming more like that of their former colonial power (an MDC).

Dependency theory states that LDCs tend to have a higher dependency ratio, the ratio of the number of people under 15 or over 64 to the number in the labour force.

Development is the improvement in material conditions of a place as a result of diffusion of technology and knowledge. This is important because it is a main goal for most of the planet's regions and development will help solve many problems.

Energy consumption is an index of development. MDCs tend to consume much more energy per capita than do LDCs. This will be important in the future because as LDCs begin to industrialize, there will be a great strain on the world's energy supply.

Foreign direct investment is investment in the economies of LDCs by transnational corporations based in MDCs. However, all countries are not recipients of this investment. Brazil, China and Mexico were the LDCs that received most of the investment.

Gender is an important developmental factor. A great difference in development between the genders is found primarily in LDCs, especially in the Middle East. Differences exist primarily in income and in literacy rate.

Gross domestic product is the total value of goods and services produced in a year in a given country. The value varies greatly between MDCs and LDCs and is one of the best indicators of development. Fast growth of GDP is a major goal of all countries.

Gross national product is similar to GDP except that includes income that people earn abroad.

The **Human Development Index** is an aggregate index of development, which takes into account economic, social and demographic factors, using GDP, literacy and education, and life expectancy.

Levels of development that countries are classified into include MDCs (more developed countries) and LDCs (less developed countries).

Measures of development are used to distinguish LDCs from MDCs. They include GDP, literacy rate, life expectancy, caloric intake, etc.

Neocolonialism refers to the economic control that MDCs are sometimes believed to have over LDCs. Through organizations such as the IMF, the MDCs are able to dictate precisely what LDCs economic policies are, or are able to use their economic subsidies to put LDCs industries out of business.

The **Physical Quality of Life** index is another development index. It is based on literacy rate, infant mortality rate, and life expectancy at age one.

Purchasing power parity is an index of income related to GDP. Unlike GDP however, PPP takes into account price differences between countries. Usually goods in LDCs are priced lower, so this makes the difference between LDCs and MDCs less.

W.W. Rostow developed the "Stages of Growth" model of economic development.

Technology gap-The difference in technologies used and/or developed in two companies, countries, ethnic groups, etc., where one is more advanced than the other.

Important because it helps to explain the differences between MDCs and LDCs.

Technology transfer-process by which existing knowledge, facilities, or capabilities developed under federal research and development funding are utilized to fulfil public and private needs.

Important because it allows for knowledge to be utilized for various needs instead of being confined to a certain sector.

Third World-countries in the developing world independent of their political status (developing countries).

Important because it is a classification to explain differences between the countries of the world.

World Systems Theory-refers to perspective that seeks to explain the dynamics of the "capitalist world economy" as a "total social system".

Important because explains the power hierarchy in which powerful and wealthy "core" societies dominate and exploit weak and poor peripheral societies.

Bid rent theory-refers to how the price and demand on land changes as the distance towards the CBD increases.

Important because it provides an explanation as to the spatial distribution of urban areas.

Assembly line production/Fordism-industrial arrangement of machines, equipment, and workers for continuous flow of work pieces in mass production operations, each movement of material is made as simple and short as possible.

Important because it allowed for goods to be produced at a rate comparable to the demand for many of those products, made for more efficient manufacturing industries.

Air pollution-concentration of trace substances at a greater level than occurs in average air, human causes include mainly motor vehicles, industry, and power plants.

Important because it can damage property and adversely affect the health of people, other animals, and plants.

Agglomeration economies-refers to benefits or advantages (savings, cost reductions, etc.) resulting from the spatial clustering of activities and/or people.

Important because:

Acid rain-tiny droplets of sulfuric acid and nitric acid in the atmosphere that dissolve in water and return to Earth's surface.

Important because it has damaged lakes, killing fish and plants.

"Stages of Growth" Model-linear theory of development that developed countries go through a common pattern of structural change (1-Traditional Society, 2-Transitional Stage, 3-Take Off, 4-Drive to Maturity, 5-High Mass Consumption).

Important because it explains the development experience of Western countries and is a general model for many others.

Rostow, W. W.-economist, developed the "Stages of Growth" model in the late 1950s.

Important because he developed the model that is frequently referred to.

Aluminum industry-U.S. companies are the largest single producer with plants in 35 states producing about $39.1 billion in products and exports. U.S. supply is comprised of three sources, primary, imports and recycled.

Important because it is a large industry that is important in transportation, packaging and building and construction.

Cities & Urban Land Use

Agglomeration – a built up area consisting of central city and its surrounding suburbs (similar to the term "urbanized area", shows the extent of a city's influence).

Barriadas – a neighbourhood, usually a slum or lower class (many of the Latin American cities have these outside the central city).

Bid-rent Theory – explains that the price/demand for land increases closer to the CBD (explains the concentric zone model and why different levels of development are located at certain distances from the central city).

Blockbusting – the process of white families selling their homes because of fears that blacks would move in and lower the property value (explains the white flight of the 1950's and the growth of suburbs).

CBD – stands for central business district, location of skyscrapers and companies (would always be the centre of the 3 urban models, many people commute, few actually live there).

Census Tract – these are govt. designated areas in cities that each have ~5,000 people, they often times correspond to neighborhoods (data in census tracts is used to analyse urban patterns such as gentrification or white flight).

Centrality – the strength of dominance of an urban centre over its surrounding area, larger than the MSA or agglomeration (Twin Cities centrality extends up into northern MN, over into ND, SD, and western WI).

Centralization – the movement of people, capital, services, and govt. into the central city (opposite of suburban sprawl, happened to cities before WWII and is happening now).

Christaller, Walter – he created the Central Place Theory, which explains how services are distributed and why there are distinct patterns in this distribution (central place theory involves market area/hinterland and the threshold, which is the minimum number of customers needed to keep the business running).

City – centralized area with a mayor and local government, usually bigger than a town (cities started in the Greek / Roman times, more and more people live in cities, especially in LDC's).

Cityscapes – similar to a landscape, yet of a city (cityscapes often show the city's skyline, which is the CBD).

Colonial City – cities founded by colonial powers, such as Mexico City by the Spanish (these often contain plazas, large Catholic cathedrals, and historic architecture, most of these are in Latin America and in Southern Asia, in India).

Commercialization – the process of the increasing importance of business (advertisements in cities, development leans toward services).

Concentric Zone Model – created by E.W. Burgess, city grows outwards from a central area (CBD in middle, then zone of transition, then zone of workers' homes, then zone of residences, then commuter's zone).

Counterurbanization – a net migration from urban to rural areas (this only happens in very developed areas in North America and Western Europe).

Decentralization-the process of dispersing decision-making outwards from the centre of authority (We learned about how nation states break up and form their own political clouts).

Deindustrialization-process of social and economic change caused by removal of industry. (We learned about how MDCs moved on after the 1800s).

Early Cities-Cities of the ancient world (-3500 to-1200) (We learned about how agriculture and language began in this era.

Economic base-Communities collection of basic industry (We learned about job sectors).

Edge city-A new concentration of business in suburban areas consisting of suburbs (We learned about urban sprawl).

Emerging cities-City currently without much population but increasing in size at a fast rate (learned about cities that are growing at a fast rate).

Employment structure-graph showing how primary secondary and tertiary sector jobs are separated.

Entrepot-Trading centre where goods are exported and imported without cost. (We learned about centres of trade).

Ethnic neighbourhood-A neighbourhood with distinctive ethnic composition (We learned about segregation of cities into ethnic backgrounds.).

Favela-A shantytown or slum, especially in Brazil (We learned about the slum conditions faced by Latin American countries.

Female-headed household-A household dominated by a woman (We learned about how MCDs have different family structure.).

Festival landscape-a landscape of cultural festivities (We learned about the culture.)

Gateway City-a settlement which acts as a link between two areas. (We learned about primate cities, which are similar.).

Gender-a person's sex (We learned about differences that occur as a result of gender).

Gentrification-process in which low cost neighborhoods are renovated by middle class to increase property values. (We learned about the positives and negatives of this process).

Ghetto-A usually poor section of a city inhabited primarily by people of the same race, religion, or social background. (We learned about the worst parts of cities).

Globalization-Development of worldwide patterns of economic relationships (we learned about he future impact this will have.

5

The Von Thunen Model

The Von Thunen model of agricultural land use was created by farmer and amateur economist J.H. Von Thunen (1783-1850) in 1826 (but it wasn't translated into English until 1966). Von Thunen's model was created before industrialization and is based on the following limiting assumptions:

- The city is located centrally within an "Isolated State" which is self sufficient and has no external influences.
- The Isolated State is surrounded by an unoccupied wilderness.
- The land of the State is completely flat and has no rivers or mountains to interrupt the terrain.
- The soil quality and climate are consistent throughout the State.
- Farmers in the Isolated State transport their own goods to market via oxcart, across land, directly to the central city. Therefore, there are no roads.
- Farmers act to maximize profits.

In an Isolated State with the foregoing statements being true, Von Thunen hypothesized that a pattern of rings around the city would develop.

There are four rings of agricultural activity surrounding the city. Dairying and intensive farming occur in the ring closest to the city. Since vegetables, fruit, milk and other dairy products must get to market quickly, they would be produced close to the city (remember, we didn't have refrigerated oxcarts!).

Timber and firewood would be produced for fuel and building materials in the second zone. Before industrialization (and coal power), wood was a very important fuel for heating and cooking. Wood is very heavy and difficult to transport so it is located as close to the city as possible.

The third zone consists of extensive fields crops such as grains for bread. Since grains last longer than dairy products and are much lighter than fuel, reducing

transport costs, they can be located further from the city. Ranching is located in the final ring surrounding the central city. Animals can be raised far from the city because they are self-transporting. Animals can walk to the central city for sale or for butchering.

Beyond the fourth ring lies the unoccupied wilderness, which is too great a distance from the central city for any type of agricultural product.

Even though the Von Thunen model was created in a time before factories, highways, and even railroads, it is still an important model in geography. The Von Thunen model is an excellent illustration of the balance between land cost and transportation costs. As one gets closer to a city, the price of land increases. The farmers of the Isolated State balance the cost of transportation, land, and profit and produce the most cost-effective product for market. Of course, in the real world, things don't happen as they would in a model.

Geography of Agriculture

Around ten to twelve thousand years ago, human began to domesticate plants and animals for food. Before this first agricultural revolution, people relied on hunting and gathering to obtain food supplies. While there are still groups of hunters and gatherers in the world, most societies have switched to agriculture. The beginnings of agriculture did not just occur in one place but appeared almost simultaneously around the world, possibly through trail and error with different plants and animals or by long term experimentation. Between the first agricultural revolution thousands of years ago and the 17th century, agriculture remained pretty much the same.

In the seventeenth century, a second agricultural revolution took place which increased efficiency of production as well as distribution which allowed more people to move to the cities as the industrial revolution got under way. The eighteenth century's European colonies became sources of raw agricultural and mineral products for the industrializing nations.

Now, many of the countries which were once colonies of Europe, especially those in Central America, are still heavily involved in the same types of agricultural production as they were hundreds of years ago. Farming in the twentieth century has become highly technological in more developed nations with geographical technologies like GIS, GPS, and remote sensing while less developed nations continue with practices which are similar to those developed after the first agricultural revolution, thousands of years ago.

About 45% of the world's population makes their living through agriculture. The proportion of the population involved in agriculture ranges from about 2% in the United States to about 80% in some parts of Asia and Africa. There are two types of agriculture, subsistence and commercial. There are millions of subsistence farmers in the world, those who produce only enough crops to feed their families.

Many subsistence farmers use the slash and burn or swidden agricultural method. Swidden is a technique used by about 150 to 200 million people, and is especially prevalent in Africa, Latin America, and Southeast Asia. A portion of land is cleared and burned to provide at least one and up to three years of good crops for that portion of land. Once the land can no longer be utilized, a new patch of ground is slashed and burnt for another round of crops. Swidden is not a neat or well-organized method of agricultural production by it is effective for farmers who don't know much about irrigation, soil, and fertilization.

The second type of agriculture is commercial agriculture, where the primary purpose is to sell one's product at market. This takes place throughout the world and includes major fruit plantations in Central America as well as huge agribusiness wheat farms in the Midwestern United States.

Geographers commonly identify two major "belts" of crops in the U.S. The wheat belt is identified as crossing the Dakotas, Nebraska, Kansas, and Oklahoma. Corn, which is primarily grown to feed livestock, reaches from southern Minnesota, across Iowa, Illinois, Indiana, and Ohio.

J.H. Von Thunen developed a model in 1826 (which wasn't translated into English until 1966) for the agricultural use of land. It has been utilized by geographers since that time. His theory stated that the more perishable and heavier products would be grown closer to urban areas. By looking at the crops grown within metropolitan areas in the U.S., we can see that his theory still holds true. It is very common for perishable vegetables and fruits to be grown within metropolitan areas while less-perishable grain is predominantly produced in non-metropolitan counties.

Agriculture uses about a third of the land on the planet and occupies the lives of about two and a half billion people. It's important to understand where our food comes from.

Johann Heinrich von Thunen

Johann Heinrich von Thunen (24 June 1783, Wangerland – 22 September 1850) was a prominent nineteenth century economist. Von Thunen was a Mecklenburg (north German) landowner, who in the first volume of his treatise, *The Isolated State* (1826), developed the first serious treatment of spatial economics, connecting it with the theory of rent. The importance lies less in the pattern of land use predicted than in its analytical approach.

Von Thunen developed the basics of the theory of marginal productivity in a mathematically rigorous way, summarizing it in the formula in which:

$R = Y (p – c) – Y Fm$

where R=land rent; Y=yield per unit of land; c=production expenses per unit of

commodity; p=market price per unit of commodity; F=freight rate; m=distance to market.

The Von Thunen model of agricultural land, created before industrialization, made the following simplifying assumptions:

- The city is located centrally within an "Isolated State."
- The Isolated State is surrounded by wilderness.
- The land is completely flat and has no rivers or mountains.
- Soil quality and climate are consistent.
- Farmers in the Isolated State transport their own goods to market via oxcart, across land, directly to the central city. There are no roads.
- Farmers behave rationally to maximize profits.

The use of a piece of land is put to is a function of the cost of transport to market and the land rent a farmer can afford to pay (determined by yield, which is held constant here).

The model generated four concentric rings of agricultural activity. Dairying and intensive farming lies closest to the city. Since vegetables, fruit, milk and other dairy products must get to market quickly, they would be produced close to the city.

Timber and firewood would be produced for fuel and building materials in the second ring. Wood was a very important fuel for heating and cooking and is very heavy and difficult to transport so it is located as close to the city.

The third zone consists of extensive fields crops such as grain. Since grains last longer than dairy products and are much lighter than fuel, reducing transport costs, they can be located further from the city.

Ranching is located in the final ring. Animals can be raised far from the city because they are self-transporting. Animals can walk to the central city for sale or for butchering.

Beyond the fourth ring lies the wilderness, which is too great a distance from the central city for any type of agricultural product.

Von Thunen's rings proved especially useful to economic history, such as Fernand Braudel's *Civilization and Capitalism,* untangling the economic history of Europe and European colonialism before the Industrial Revolution blurred the patterns on the ground. In economics, von Thunen rent is an economic rent created by spatial variation or location of a resource. It is 'that which can be earned *above* that which can be earned at the margin of production'.

Model of Agricultural Land Use: The Thunen Rings

In his theory of "The Isolated State", he started out from Adam Smith's idea of "economic man": that the farmer is expected to maximize his profit ("economic rent")

from his farmland. Von Thunen, as a landlord, knew that such returns depends on an optimal use of the land surfaces and the transport costs. In concentrating on the effects of these two variables on profits, removal of other factors results in a homogeneous-and isolated-state: A circular, completely undilating plane with a single, dominant market in the centre and no interactions with the outside. The economy in the surrounding rural area would have to rearrange itself according to economic behaviour in such a way that each industry brings optimal profit in:

Transport cost depends on the distance from the market and different kind of products. The gain from farming per unit area (locational rent) decreases with increasing distance from the market. The minimum price of a commodity is calculated by locational rent, transport costs and fixed production costs-the profit is then the difference between the costs and the fixed market price. Idealized pattern of agricultural land use zones in von Thunen's model.

Locational rent, a term used by von Thunen in his argument, is to be understood as the equivalent to land value. It corresponds to the maximum amount a farmer could pay for using the land, without making losses. It can be defined as the equation below:

$L = Y(P - C) - YDF$ is...

- L: Locational rent (in DM / km^2)
- Y: Yield (in t / km^2)
- P: Market price of the crop (in DM / t)
- C: Production cost of the crop (in DM / t)
- D: Distance from the market (in km)
- F: Transport cost (in DM / t / km)

Take the locational rent of a product with a yield of 1,000 t / km^2, for example, with a fixed price of 100 DM / t in the market. Production and transport costs are respectively, 50 DM / t and 1 DM / t / km. The locational rent is 50,000 DM / km^2 at the market, 40,000 DM / km^2 10 km from the market and only 20,000 DM/km^2 30 km from the market. Since locational rent falls with increasing distance from the market, the amount each farmer is willing to pay for agricultural land will shrink and the price of land will eventually decline.

Von Thunen concluded that the cultivation of a crop is only worthwhile within certain distances from the city: beyond that, either the cost of the land becomes too high, with increasing distances transport costs also increase, or, if there is another product having greater yield or lower transport costs. After a distance from the market (the city) the production of a crop becomes unprofitable, either because its profits drop to zero or the profits earned by other crops are higher, as von Thunen calculated them for products having different intensities (cattle, wood, grain, eggs,

milk, etc.): For each product there is a certain distance from the city where its production would be worthwhile. Since Thunen referred transport costs directly to the market ("Luftlinie"), circular land use zones arises-the Thunen rings. Representation of the locational rent relationship between two agricultural goods.

The farmers of these products compete against each other, plant their crops concentrically around the market according to the locational rent curves of their own crops. Products having low yields with high price and high transport costs relative to its weight or distance due to its weight, will have higher locational rent close to the market than a product having lower transport costs. Locational rent is the highest possible amount one will pay for the use of the land for a certain cultivation, and is a relative indicator of competitiveness of it in the market.

Weaknesses & Criticism

The model was developed in an isolated state and did not take into consideration differences in sites (local physical conditions). It can be modified by relaxing some of the conditions set forth by Von Thunen:

- differential transportation costs. Example: boats are the cheapest mode of transportation
- variations in topography
- soil fertility
- changes in demand or price of the commodity

However, the model tends to hold true in most instances

The theory may break down somewhat in industrial and post-industrial economies as urban expansion / sprawl occur. For example, modern refrigerators enable perishable products to be transported longer distances.

Like many other models in geography, von Thunen's model was criticized frequently due to its restrictive nature. The basic conditions of the model, however, could be approximated by slight modifications of the respective reality. The circular pattern, which can be attributed to only one market and excluding transport costs gradients running from the centre, is for example only one of many conceivable geometrical starting situations. If other natural landscapes or transportation routes are present, the land use zones would be stripe-formed. If several markets were present, groups of zones would be formed around each market.

A justified objection against it is the reference to the absence of any productive profit. In von Thunen's theory different agricultural uses compete for the optimal location, which results from the product-specific supply / expenditure relation. The competitive power becomes indirectly measurable over locational rent. After deducting production costs and location-specific transport costs, however, nothing more remains of the market profits. The von Thuenen model leads to the idea of complete self-

sufficiency among farmers. Thunen's idea of "economic rent" attempted-while ignoring other characteristics-to explain the use of zones controlled solely by economically rational perception. Possible consumers play, finally, the crucial role for the choice of location. At the same time evaluation of all potential locations is released, which leads to a zoning of the possible offers. This simply developed space restaurant model reacts however sensitively to changes of the space overcoming costs. It possesses however due to its universality nevertheless a high value within geographical questions and methodology.

Von Thunen's Regional Land Use Model

If modern economics began with Adam Smith, modern location economics began with Von Thunen (1826). He was the first to develop a basic analytical model of the relationships between markets, production, and distance. For this purpose he looked upon the agricultural landscape. The relative costs of transporting different agricultural commodities to the central market determined the agricultural land use around a city. The most productive activities will thus compete for the closest land to the market and activities not productive enough will locate further away. The model has a set of basic assumptions which reflects agricultural conditions around a city in the early 19th century:

- Isolation. There is one isolated market in an isolated state having no interactions (trade) with the outside.
- Ubiquitous land characteristics. The land surrounding the market is entirely flat and its fertility uniform.
- Transportation. It is assumed there are no transport infrastructures such as roads or rivers and that farmers are transporting their production to the market using horses and carts. Transportation costs are dependent of the type of commodity being transported to the market as well as the distance involved.

The model compares the relationships between production cost, the market price and the transport cost of an agricultural commodity and is expressed as follows:

$R = Y(p-c)-Yfm$

- R = Rent per unit of land.
- Y = Yield per unit of land.
- p = market price per unit of yield.
- c = Average production costs per unit of yield.
- m = Distance from market (in kilometres or miles).
- of = Freight rate per unit of yield and unit of distance.

All agricultural land uses are maximizing their productivity (rent), which in this case is dependent upon their location from the market (Central City). The role of

farmer is to maximize his profit which is simply the market price minus the transport and production costs. The most productive activities (gardening or milk production) or activities having high transport costs (firewood) locate nearby the market. The above figure provides an overview of Von Thunen's agricultural land use model with the basic assumptions being applied (isolation, ubiquity, transportation). It can be divided in two parts:

- The pure isolated state over an isotropic plain. In this case, the model takes a shape of perfect concentric circles.
- The potential impacts of modified transport costs (a navigable river) and the presence of a competing centre.

The relationships between agricultural land use and market distance are very difficult to establish in the contemporary context. However, a strong relationship between the transport system and regional agricultural land use patterns can be acknowledged at the continental level in North America.

Location Theory

Economics and geography in economics and geography, theory concerned with the geographic location of economic activity; it has become an integral part of economic geography, regional science, and spatial economics. Location theory addresses the questions of what economic activities are located where and why. The location of economic activities can be determined on a broad level such as a region or metropolitan area, or on a narrow one such as a zone, neighbourhood, city block, or an individual site.

Johann Heinrich von Thunen (Thunen, Johann Heinrich von), a Prussian landowner, introduced an early theory of agricultural location in *Der isolierte Staat* (1826) (*The Isolated State*). The Thunen model suggests that accessibility to the market (town) can create a complete system of agricultural land use. His model envisaged a single market surrounded by farmland, both situated on a plain of complete physical homogeneity. Transportation costs over the plain are related only to the distance traveled and the volume shipped. The model assumes that farmers surrounding the market will produce crops which have the highest market value (highest rent) that will give them the maximum net profit (the location, or land, rent). The determining factor in the location rent will be the transportation costs. When transportation costs are low, the location rent will be high, and vice versa. This situation produces a rent gradient along which the location rent decreases with distance from the market, eventually reaching zero. The Thunen model also addressed the location of intensive versus extensive agriculture in relation to the same market. Intensive agriculture will possess a steep gradient and will locate closer to the market than extensive agriculture. Different crops will possess different rent gradients. Perishable crops (vegetables and dairy products) will possess steep gradients while less perishable crops (grains) will

possess less steep gradients. In 1909 the German location economist Alfred Weber formulated a theory of industrial location in his book entitled *Über den Standort der Industrien* (*Theory of the Location of Industries,* 1929). Weber's theory, called the location triangle, sought the optimum location for the production of a good based on the fixed locations of the market and two raw material sources, which geographically form a triangle. He sought to determine the least-cost production location within the triangle by figuring the total costs of transporting raw material from both sites to the production site and product from the production site to the market.

The weight of the raw materials and the final commodity are important determinants of the transport costs and the location of production. Commodities that lose mass during production can be transported less expensively from the production site to the market than from the raw material site to the production site. The production site, therefore, will be located near the raw material sources. Where there is no great loss of mass during production, total transportation costs will be lower when located near the market.

Once a least-transport-cost location had been established within the triangle, Weber attempted to determine a cheap-labour alternate location. First he plotted the variation of transportation costs against the least-transport-cost location. Next he identified sites around the triangle that had lower labour costs than did the least-transport-cost location. If the transport costs were lower than the labour costs, then a cheap-labour alternative location was determined.

Another major contribution to location theory was Walter Christaller's formulation of the central place theory, which offered geometric explanations as to how settlements and places are located in relation to one another and why settlements function as hamlets, villages, towns, or cities.

William Alonso (*Location and Land Use: Toward a General Theory of Land Rent,* 1964) built upon the Thunen model to account for intra-urban variations in land use. He attempted to apply accessibility requirements to the city centre for various types of land use (housing, commercial, and industry). According to his theory, each land use type has its own rent gradient or bid rent curve. The curve sets the maximum amount of rent any land use type will yield for a specific location. Households, commercial establishments, and industries compete for locations according to each individual bid rent curve and their requirements for access to the city centre. All households will attempt to occupy as much land as possible while staying within their accessibility requirements. Since land is cheaper at the fringe of the city, households with less need for city centre accessibility will locate near the fringe; these will usually be wealthy households. Poor households require greater accessibility to the city centre and therefore will locate near the centre, competing with commercial and industrial establishments. This will tend to create a segregated land use system, because households will not pay commercial and industrial land prices for central locations.

The Thunen, Weber, Alonso, and Christaller models are not the sole contributors to location theory, but they are its foundation. These theories have been expanded upon and refined by geographers, economists, and regional scientists.

Weber's Theory of Industrial Location (The Pure Theory)

Alfred Weber, a German economist, enunciated a systematic theory of industrial location in 1909. Weber's theory of location is purely deductive in its approach. He analysed the factors that determine the location of industry and classified these factors into two divisions. These are:

(i) Primary causes of regional distribution of industry (regional factors)

(ii) Secondary causes (agglomerative and deglomerative factors) that are responsible for redistribution of industry

(i) Primary Causes (Regional Factors)

According to Weber, transport costs and labour costs are the two regional factors on which his pure theory is based. Assuming that there are no other factors that influence the distribution of industry, except transportation costs. Then it is clear that the location of industry will be pulled to those locations which have the lowest transportation costs. The key factors that determine transportation costs are:

(i) the weight to be transported.

(ii) the distance to be covered.

Weber lists some more factors which influence the transportation costs such as – (a) the type of transportation system and the extent of its use, (b) the nature of the region and kinds of roads, (c) the nature of goods themselves, i.e., the qualities which, besides weight, determine the facility of transportation.

However, the location of the place of production must be determined in relation to the place of consumption and to the most advantageously located material deposits. Thus, 'locational figures' are created. These locational figures depends upon (a) the type of material deposits and (b) the nature of transformation into products.

Weber classifies and calls those raw materials, which are available practically everywhere as 'ubiquities' (like brick-clay, water, etc) and 'localised' (like iron-ore, minerals, wood, etc) which are available only in certain regions. It is clear that localized materials play a more important role on the industry than the ubiquities. Further, regarding the nature of the transformation of materials into products, Weber categorized the raw materials as 'pure' and 'weight losing'. Pure materials impart their total weight to the products (eg. cotton, wool, etc) and the materials are said to be 'weight losing' if only a part enters into the product (eg. wood, coal, etc.). Hence, the location of industries using weight-losing materials is drawn towards their deposits and that of industries using pure-materials towards the consumption centres.

Weber further examines the cause of deviation of industrial location from the centres of least transport costs. The existence of differences in labour costs leads an industry to deviate from the optimal point of transport orientation. Geographical distribution of the population would give rise to differences in wages for labour. Naturally, the transport oriented location of an industry is drawn out and attracted towards the cheaper labour centres. Such migration of an industry from a point of minimum transport costs to a cheaper labour centre may be likely to occur only where the savings in the cost of labour are larger than the additional costs of transport which it ought to incur.

Secondary Causes (Agglomerative and Deglomerative Factors)

An agglomerative factor is an advantage or a cheapening of production or marketing which results from the fact that production is carried on at one place. A deglomerative factor is a cheapening of production which results from the decentralization of production i.e., production in more than one place. To some extent these agglomerative and deglomerative factors also contribute to local accumulation and distribution of industry. These factors will operate only within the general framework formed by the two regional factors, i.e., costs of transportation and costs of labour. The advantages which could be derived in this context are external economies.

The pulls which the agglomerative factors possess to attract an industry to a particular point are mainly dependent on two factors. Firstly, on 'the index of manufacture' (the proportion of manufacturing costs to the total weight of the product) and secondly, on the 'locational weight' (the total weight to be transported during all the stages of production). To deduce a general principle, Weber uses the concept of "co-efficient of manufacture" which is the ratio of manufacturing cost to locational weight. Agglomeration is encouraged with high co-efficient of manufacture and deglomeration with low co-efficient of manufacture and these tendencies are inherent in their nature.

Split Location: Productive activities could be divided depending on the nature of raw-materials, industry and market. Weber considers the location for an industry at more than one place. According to Weber, a split of production into several locations will be the rule for productive process which can technically be split. For instance, the first stage of production may be near the raw material deposits and the subsequent stages near the place of final consumption. Likewise, in a paper industry the manufacture of pulp may be carried on near the supplies of the raw materials and the second stage of paper manufacture near the consumption outlet.

Locational Coupling: Weber also conceived the advantages of setting up different types of industries in the same locality. The production of quite different articles may be combined in one plant because several raw materials may diverge from a common source. This may be either due to technical or economic reasons: for instance, certain chemical industries, garments factories which manufacture over-coats, shawls, blouses,

etc. Locational coupling may also occur due to connection through materials. If the by-product of an industry happens to be the raw material of another industry, then the two industries may select a single place of location. For instance, the dye-stuff industry is connected with other industries using coke, because coal tar (upon which the dye-stuff industry is based) is a by-product of the burning coke.

Criticisms: Weber's theory of location has been criticized on various grounds which may be summarized as follows:

1. Weber has been criticized for his unrealistic approach and deductive reasoning. According to Sargant Florence, vague generalizations cannot provide suitable solutions to the theory of location as non-economic considerations will also influence which are not mentioned in the pure theory. He says that Weber's theory fails to explain locations resulting from historical and social forces.
2. A.Predohl criticizes Weber's theory as more a selective theory than a deductive theory. The very distinction between primary and secondary is itself artificial, illogical and arbitrary.
3. Weber assumes fixed labour centres and unlimited supplies of labour which are unrealistic. The rise of industry may create new labour centres and we cannot assume unlimited labour supplies at any centre.
4. In a competitive market structure, the assumption of fixed points of consumption is unrealistic. Country-wise scattering, usually, of consuming public is a reality and there may be a shift in the consuming centres with a shift in industrial population.
5. A. Robinson also considers Weber's division of raw materials into 'ubiquities' and 'localised' as artificial.

Weber's deductive theory of location, in spite of the shortcomings, is the only theory which has been enjoying the universal acceptance and application, as all the other alternative suggestions are neither complete nor comprehensive.

Factors Influencing Industrial Location

Generally, location of industries is influenced by economic considerations though certain non-economic considerations also might influence the location of some industries. Maximisation of profit which also implies cost minimization is the most important goal in their choice of particular places for the location of industries. There are several factors which pull the industry to a particular place. Some of the major factors influencing location are discussed below:

1. Availability of raw materials: In determining the location of an industry, nearness to sources of raw material is of vital importance. Nearness to the sources of raw materials would reduce the cost of production of the industry. For most of the major industries, the cost of raw materials form the bulk of

the total cost. Therefore, most of the agro-based and forest-based industries are located in the vicinity of the sources of raw material supply.

2. Availability of Labour: Adequate supply of cheap and skilled labour is necessary for and industry. The attraction of an industry towards labour centres depends on the ratio of labour cost to the total cost of production which Weber calls 'Labour cost of Index'. The availability of skilled workers in the interior parts of Bombay region was one of the factors responsible for the initial concentration of cotton textile industry in the region.
3. Proximity to Markets: Access to markets is an important factor which the entrepreneur must take into consideration. Industries producing perishable or bulky commodities which cannot be transported over long distance are generally located in close proximity to markets. Industries located near the markets could be able to reduce the costs of transport in distributing the finished product as in the case of bread and bakery, ice, tins, cans manufacturing, etc. Accessibility of markets is more important in the case of industries manufacturing consumer goods rather than producer goods.
4. Transport Facilities: Transport facilities, generally, influence the location of industry. The transportation with its three modes, i.e., water, road, and rail collectively plays an important role. So the junction points of water-ways, roadways and railways become humming centres of industrial activity. Further, the modes and rates of transport and transport policy of Government considerably affect the location of industrial units. The heavy concentration of cotton textile industry in Bombay has been due to the cheap and excellent transportation network both in regard to raw materials and markets.
5. Power: Another factor influencing the location of an industry is the availability of cheap power. Water, wind, coal, gas, oil and electricity are the chief sources of power. Both water and wind power were widely sought at sources of power supply before the invention of steam engine. During the nineteenth century, nearness to coal-fields became the principal locating influence on the setting up of new industries, particularly, for heavy industries. With the introduction of other sources of power like electricity, gas, oil, etc. the power factor became more flexible leading to dispersal and decentralization of industries.
6. Site and Services: Existence of public utility services, cheapness of the value of the site, amenities attached to a particular site like level of ground, the nature of vegetation and location of allied activities influence the location of an industry to a certain extent. The government has classified some areas as backward areas where the entrepreneurs would be granted various incentives like subsidies, or provision of finance at concessional rate, or supply of power a cheaper rates and provision of education and training facilities. Some

entrepreneurs induced by such incentives may come forward to locate their units in such areas.

7. Finance: Finance is required for the setting up of an industry, for its running, and also at the time of its expansion. The availability of capital at cheap rates of interests and in adequate amount is a dominating factor influencing industrial location. For instance, a review of locational history of Indian cotton textile industry indicates that concentration of the industry in and around Bombay in the early days was mainly due to the presence of rich and enterprising Parsi and Bhatia merchants, who supplied vast financial resources.
8. Natural and Climatic Considerations: Natural and climatic considerations include the level of ground, topography of a region, water facilities, drainage facilities, disposal of waste products, etc. These factors sometimes influence the location of industries. For instance, in the case of cotton textile industry, humid climate provides an added advantage since the frequency of yarn breakage is low. The humid climate of Bombay in India and Manchester in Britain offered great scope for the development of cotton textile industry in those centres.
9. Personal Factors: In deciding location of industrial units, sometimes an entrepreneur may have personal preferences and prejudices against certain localities. For instance, Mr. Ford started to manufacture motor cars in Detroit simply because it was his home-town. In such cases, personal factor dominates other considerations. However, this kind of domination is rare.
10. Strategic Considerations: In modern times, strategic considerations are playing a vital role in determining industrial location. During war-time a safe location is assuming special significance. This is because in times of war the main targets of air attacks would be armament and ammunition factories and industries supplying other commodities which are required for war. The Russian experience during the Second World War provides and interesting example.
11. External Economies: External economies also exert considerable influence on the location of industries. External economies arise due to the growth of specialized subsidiary activities when a particular industry is mainly localized at a particular centre with port and shipping facilities. External economies could also be enjoyed when a large number of industrial units in the same industry were located in close proximity to one another.
12. Miscellaneous Factors: Historical incidents also play a dominating role in determining the location of industries in certain cases. The development of cotton-textile industry in Lancashire provides an interesting example for this. Further, the size of and industrial unit would also have much influence in choosing location. This is because the size of industrial units depends upon

the radius of the circle within which they can profitably distribute their goods and upon the density of population living within the circle.

Alfred Weber

Alfred Weber (30 July 1868 – 2 May 1958) was a German economist, sociologist and theoretician of culture whose work was influential in the development of modern economic geography.

Life

Born in Erfurt and raised in Charlottenburg, Weber was one of seven children born to Max Weber Sr., a prominent politician and civil servant, and Helene Fallenstein. Weber Sr.'s engagement with public life immersed the family home in politics, as his salon received many prominent scholars and public figures. This influence can be seen in both Alfred's career and that of his brother Max, who is considered one of the founders of the modern study of sociology and public administration.

From 1907 to 1933, Weber was a professor at the University of Heidelberg until his dismissal following criticism of Hitlerism. Weber lived in Nazi Germany during the Second World War, but was a leader in intellectual resistance. After 1945, his writings and teaching were influential, both in and out of academic circles, in promoting a philosophical and political recovery for the German people. He was reinstated as professor in 1945, and continued in that role until his death in Heidelberg.

Work

Weber supported reintroducing theory and causal models to the field of economics, in addition to using historical analysis. In this field, his achievements involve work on early models of Industrial location. He lived during the period when sociology became a separate field of science.

Weber maintained a commitment to the "philosophy of history" traditions. He contributed theories for analyzing social change in Western civilization as a confluence of civilization (intellectual and technological), social processes (organizations) and culture (art, religion, and philosophy). He went to St. Joseph's Convent in Bideford, Maine on 13 April 1928 He conducted empirical and historical analyses of the growth and geographical distribution of cities and capitalism.

Least Cost Theory

Leaning heavily on work developed by the relatively unknown Wilhelm Launhardt, Alfred Weber formulated a least cost theory of industrial location which tries to explain and predict the locational pattern of the industry at a macro-scale. It emphasizes that firms seek a site of minimum transport and labour cost.

The point for locating an industry that minimizes costs of transportation and labour requires analysis of three factors:

Material Index

The point of optimal transportation based on the costs of distance to the "material index"-the ratio of weights of the intermediate products (raw materials) to the finished product. In one scenario, the weight of the final product is less than the weight of the raw material going into making the product — the *weight losing industry*. For example, in the copper industry, it would be very expensive to haul raw materials to the market for processing, so manufacturing occurs near the raw materials. (Besides mining, other primary activities (or extractive industries) are considered material oriented: timber mills, furniture manufacture, most agricultural activities, etc.. Often located in rural areas, these businesses may employ most of the local population. As they leave, the local area loses its economic base.)

In the other, the final product is heavier than the raw materials that require transport. Usually this is a case of some ubiquitous raw material, such as water, being incorporated into the product. This is called the *weight-gaining industry*.

Labour

The labour distortion: sources of lower cost labour may justify greater transport distances and become the primary determinant in production.

A. Unskilled Labour –industries such as the garment industry require cheap unskilled laborers to complete activities that are not mechanized. They are often termed "ubiquitous" meaning they can be found everywhere. Its pull is due to low wages, little unionization and young employees.
B. Skilled Labour-High tech firms, such as those located in Silicon Valley, require exceptionally skilled professionals. Skilled labour is often difficult to find.

Agglomeration and Deglomeration

Agglomeration is the phenomenon of spatial clustering, or a concentration of firms in a relatively small area. The clustering and linkages allow individual firms to enjoy both internal and external economies. Auxiliary industries, specialized machines or services used only occasionally by larger firms tend to be located in agglomeration areas, not just to lower costs but to serve the bigger populations.

Deglomeration occurs when companies and services leave because of the diseconomies of industries' excessive concentration. Firms who can achieve economies by increasing their scale of industrial activities benefit from agglomeration. However, after reaching an optimal size, local facilities may become over-taxed, lead to an offset of initial advantages and increase in PC. Then the force of agglomeration may eventually be replaced by other forces which promote deglomeration.

Globalization

Similarly, industrial activity is considered a secondary economic activity, and is

also discussed as manufacturing. Industrial activity can be broken down further to include the following activities: processing, the creation of intermediate parts, final assembly. Today with multinational corporations, the three activities listed above may occur outside MDCs.

Weber's theory can explain some of the causes for current movement, yet such discussion did not come from Weber himself. Weber found industrial activity the least expensive to produce. Least cost location then implies marketing the product at the least cost to the consumer, much like retailers attempt to obtain large market shares today. Economically, it is explained as one way to make a profit; creating the cheapest product for the consumer market leads to greater volume of sales and hence, greater profits. Therefore, companies that do not take the time to locate the cheapest inputs or the largest markets would not succeed, since their product costs more to produce and costs the consumer more.

His theory has five assumptions. His first assumption is known as the isotropic plain assumption. This means the model is operative in a single country with a uniform topography, climate, technology, economic system. His second assumption is that only one finished product is considered at a time, and the product is shipped to a single market. The third assumption is raw materials are fixed at certain locations, and the market is also a known fixed location. The fourth assumption is labour is fixed geographically but is available in unlimited quantities at any production site selected. The final assumption is that transport costs are a direct function of weight of the item and the distance shipped.

In use with his theory he created the locational triangle. His triangle is used with one market and two sources of material. This illustrated that manufacturing that utilizes pure materials will never tie the processing location to the material site. Also industries utilizing high weight loss materials will tend to be pulled toward the material source as opposed to the market. Furthermore many industries will select an intermediate location between market and material. The last generalization is considered to be wrong because he never takes into account terminal costs and therefore is considered biased toward intermediate locations.

To further explore the location of firms Weber also created two concepts. The first is of an isotim, which is a line of equal transport cost for any product or material. The second is the isodapane which is a line of total transport costs. The isodapane is found by adding all of the isotims at a location. The reason for using isodapanes is to systematically introduce the labour component into Weber's locational theory.

Weber has received much criticism. It has been said that Weber did not effectively and realistically take into account geographic variation in market demand, which is considered a locational factor of paramount influence. Also his treatment of transport did not recognize that these costs are not proportional to distance and weight, and that intermediate locations necessitate added terminal charges. Labour is not always

available in unlimited quantity at any location and is usually quite mobile through migration. Plus most manufacturing plants obtain a large number of material inputs and produce a wide range of products for many diverse markets, so his theory doesn't easily apply. Furthermore he underestimated the effect of agglomeration.

Weber's Location Triangle

Alfred Weber's work (1909) is considered to have established the foundations of modern location theories. One of his core assumption is that firms will chose a location in view to minimize their costs. This involves a set of simplifications, namely that location takes place in an isolated region (no external influences) composed of one market, that space is isotropic (no variations in transport costs except a simple function of distance) and that markets are located in a specific number of centres. Those conditions are quite similar to those behind Von Thunen's agricultural land use model elaborated almost one hundred years earlier. The model also assumes perfect competition, implying a high number of firms and customers, small firm sizes (to prevent disruptions created by monopolies and oligopolies) and a perfect knowledge of market conditions, both for the buyers and suppliers. Several natural resources such are water are ubiquitous (available everywhere) while many production inputs such as labour, fuel and minerals are available at specific locations. According to Weber, three main factors influence industrial location; transport costs, labour costs and agglomeration economies. Location thus imply an optimal consideration of these factors.

Solving Weber's location model often implies three stages; finding the least transport cost location and adjusting this location to consider labour costs and agglomeration economies. Transportation is the most important element of the model since other factors are considered to only have an adjustment effect. To solve this problem, Weber uses the location triangle within which the optimal is located. The problem resides in finding an optimal factory location P located at the respective distances of d(M), d(S1) and d(S2). Several methodologies can be used to solve this problem such as drawing an analogy to a system of weights and pulleys (Varignon's solution) or using trigonometry. Another way preferred among geographers, particularly with GIS, is to use cost surfaces which are overlaid.

Weber's location theory explains well the location of heavy industries, particularly from the industrial revolution until the mid twentieth century (the sector that Weber was looking at). Activities having a high level of use of raw materials tend to locate near supply sources, such as aluminum factories will locate near energy sources (electricity) or port sites. Activities using ubiquitous raw materials, such as water, tend to locate close to markets. To assess this issue, Weber developed a material index which is simply the weight of the inputs divided by the weight of the final product (output). If the material index is higher than 1, location tends to be toward material

sources. If it is less than 1, location tends to be toward the market. Contemporary developments in manufacturing, the reduction of transport costs and new economic sectors (high technology) has changed locational behaviour substantially as it locates without much consideration to Weber's principles. Still, these principles apply well for industries with a very high material index.

Alfred Weber and Subsequent Developments in Industrial Location Theory

1. Why are we interested in Weberian Triangles? Objectives of deriving Weberian principles and least-cost location theory
2. Setting up the theoretical model: the constraints (assumptions)
    - the "objective function" for this "normative" model
        * starting with transport cost minimization
        * overall cost minimization
        * profit maximization
3. Assumptions related to organization and, importantly, to production process and production function [the initial Weber model is based on linear production relations with no scale economies and no input substitution; "Leontief-type production function"
4. Assumptions related to inputs used in the production process
    * Inputs are available in unlimited supply at given (fixed) prices independent of location
    * These material inputs are either localized at a limited number of locations or ubiquitously available (= everywhere)
    * "pure" versus weight-losing materials
    * competitive price formation on markets for inputs (input prices are independent of production levels)
5. Assumptions related to markets
    * Demand is fixed and concentrated at a limited number of (known) locations (points)
    * competitive price formation (given prices at market, i.e. prices are independent of production levels)
6. Assumptions related to the transportation process and transport costs
        * Transport is possible in any direction, and transport costs for materials or products are linear, i.e. directly proportional to weight and distance, implying a "flat plain" and an absence of networks and distance & scale economies in transportation

7. Precursor: Wilhelm Launhardt
   — Carl Wilhelm Friedrich Launhardt, 1832-1918
8. One-Dimensional Model: One Localized Input/ One Localized Market
9. Location Triangle: Two Localized Inputs, one Localized Market
   — Solution procedures (The optimal Weberian location is NOT the centre of gravity [which minimizes the sum of the squares of distances], but the spatial median, which minimizes the sum of the distances).
      * Varignon Frame
      * Geometric solution (use of weight triangle)
      * Iterative algorithm starting with centre of gravity (arithmetic mean) in search for the spatial median
      * Isodapane method (iterative procedure; intermittant partial equilibria; spatial substitution between 'transport inputs'(Isard)
   — Conditions for interior versus corner locations
      * material index
      * locational weight
      * "orientation"
10. Isodapane-based Extensions of the Triangle Solution
    • Critical Isodapane
      * Labour and other spatial cost variations, "labour coefficients"
      * Agglomeration, "joint action spaces" (Weber Ch.5)
11. David Smith:
12. Analysis of Effects of Variations in Selected Assumptions
    • non-linear transport cost functions
      * non-linear production functions

Steps :

* Definition and kinds of "production function"
* Nature of linear, fixed-factor production function (right angled isoquants)
* Linear homogenous production function of the frist degree (constant returns to scale)
* Non-linear production functions (Leon Moses)
* increasing returns to scale
* decreasing returns to scale
* Equilibrium at any one isoquant/ level of output: iso-cost lines
* Iso-cost lines and scale: structure of factor input markets

The Moses Model [JSTORS-PDF] :

* Structure of the model;
* "Spatial iso-outlay lines"
* Inseparability of optimum output, input combination and location

Product (Output) Oriented Production Function (Hamilton 1974), allowing for :

* Multiple products
* Substitution between outputs
* Continuous substitution among and between inputs & outputs
* Environmental interpretations and insights

Considerations related to the structure and price mechanisms on on input and output markets and the market for transport services.

Entrepreneurship, organization, subsidies, chance, personal factors.

Externalities, agglomeration, "business climates", etc.

Multi-location problems (warehouses, branch plants etc.)

Location Models

Factors affecting Industrial Location:

- Accessibility – Good in & out e.g. motorway access
  - Access for workforce
  - Access to airports to be global
- Land – Cheap & flat
  - Room for expansion
- Availability of skilled workforce
- Location of raw materials
- Services-power & water i.e. Infrastructure
- Cost-tax, land costs, rates etc.
- Government policy – influence location of industry

Raw Materials

- Near to raw materials due to immobility – transport inefficient
- Today industry is footloose – grater efficiency and parts may be made elsewhere
- Industries with heavy, bulky materials & low costs locate near raw materials

Power Supply

Early industry needed to be near power supply, however now more efficient.

Often locate along transport routes. With the National Grid power is transported cheaply and easily over the whole country. Some industries may use HEP for cheap power e.g. Aluminium production in Canada.

Transport

Once a major consideration. Transport now more efficient with the use of containers. It now doesn't really matter where located due to ease of transport to markets. 2 types of transport costs:

1. Terminal costs – Time & equipment needed to store goods.
2. Long haul costs – Cost of actually moving goods.

Markets

Pull of markets now very important. Locate near new markets if :

- Bulky products and many linking industries
- Perishable goods
- Market is very large
- The market is wealthy

Labour Supply

The cost of labour in the UK is high therefore mechanisation used. Labour generally immobile & expect industry to come to them. Industry tries to locate where there are specific skills. Mechanisation means they are freer to locate. Often locate near unskilled labour or high skilled for R&D. May want to locate in a good environment to attract high-skilled, well-paid workers e.g. SE England. This is particularly true for quaternary industry e.g. Microsoft in Seattle.

Capital

- Working Capital – money from profits, shareholders etc. Doesn't usually affect location.
- Fixed capital – Buildings & equipment. Not mobile.
- Social Capital – Houses, shops, hospitals etc. which may attract industry.

Government Policy

Aim to even out differences in employment e.g. regional assistance in the UK. Try to encourage industry to areas of high unemployment and restrict others. Offer grants by exemption from rates. Government may control industry completely.

Land

19th C – Large areas of flat land needed. Today prefers cheap, uncongested land with better accessibility. Government is trying to attract to inner city areas.

Environment

People want to work in better environment so look for smaller towns near the country.

Chance Factors

E.g. Industries begun by individuals in their home town.

Industrial Location Models

2 main approaches to location theory in a "free market" or Capitalist system :

- The industrialist seeks the Lowest Cost Location – Weber
- The industrialist seeks the area which will give the highest profit – Smith

Weber's Model – Classical Location Theory

Weber was a German spatial economist who created his model in 1909.

Assumptions made:-

- There was a uniform transport system, culture, climate, economic & political situation
- Not all materials were evenly distributed across the plain
  - Ubiquitous – were evenly distributed
  - Localised – not evenly distributed, may be gross or pure
- Size and location of markets fixed
- Transport costs were a function of mass & distance moved.
- Labour found at fixed locations with the same rates and skills
- There was perfect competition so no industry would influence prices and revenue would be similar

The best site would be 1 with minimal production costs – Least Cost Location.

MI = material Index.

MI > 1 – weight loss in manufacture

Raw materials said to be gross

Locate near to raw materials

MI < 1 – Weight gain in manufacture

Locate near markets

MI = 1 – Raw materials must be pure as its weight remains constant

Can locate near market, raw materials or anywhere in between

Spatial Distribution of Transport Costs

Isotim – A line joining all places with equal transport costs for moving either the raw material or the product.

Isodapane – A line joining all places with equal total transport costs. (The sum of the costs of transporting raw materials and products).

Effect of Labour Costs and Agglomeration Economies

There were also factors that influence production costs in the Weber model.

Labour Costs

This considered whether moving to an area of cheaper labour would offset the increase in transport costs. He introduced the critical isodapane at the point where savings on labour costs equalled the loss by greater transport costs. If cheap labour was within the critical Isodapane then it would be worth moving there.

Agglomeration Economies

It would be profitable for 3 firms to locate within the intersection of their 3 isodapanes. However, it would not be profitable to agglomerate if none of the isodapanes overlapped.

Criticisms of Weber's Model

- The model no longer relates to modern conditions e.g. improvements in transport, the changing organisation of industry
- Different countries may be at different stages of economic development
- The material index is crude and only works for primary industry or where there is a very high or low index
- There are misconceptions in the assumptions
  - Price and demand change over time
  - There is no perfect competition
  - Decisions are not always rational
  - Transport costs vary
- Weber over-emphasised the importance of transport costs.

Benefits of the Model

- Many TNCs seek cheap labour
- Iron & steel industries are in LCLs
- Agglomeration occurs in hi-Tec industry e.g. Cambridge Science Park
- The global importance of TNCs creates a rational being (maximising profits)

Smith's area of Maximum Profit

Put forward by Smith in 1971, suggesting profits are made anywhere where total revenue exceeded total costs. There would therefore be a wider area where production is profitable. Firms rarely locate to LCL but usually between the profit margins. Firms

choose a sub-optimal location because they don't have perfect knowledge and don't always act rationally. These are also called spatial margins. Often people want satisfactory profits in exchange for better working conditions so are within the profit area.

Spatial Margins of Profitability

Problems:

- The model may be too simple.
- It is possible to define perfect knowledge and ability.
- Can behaviour be due only to ability and knowledge?

Benefits:

- Areal distribution – spatial margins are lines within which production is profitable.
- Considerable value in predicting the constraints on locational choice where information on spatial variations in cost and revenue are limited.
- Used to show the effect of government legislation on industrial location.
- Focuses on limits of freedom of locational choice for planners with imperfect ability and knowledge rather than a point of maximum profit.

Behavioural Models

Location is explained as much by social & cultural factors as by economic factors.

Criticises Least Cost Location as people's behaviours vary. E.g. some want maximum profit; some want a good standard of living. Some people aim for satisfactory location as they obtain "psychological income".

Behavioural Matrix

Consists of two axes.

- 1 indicating ability
- The other the quantity and quality of information available

With increasing knowledge and ability the decision-maker is more likely to make the right decisions. Behavioural matrix accounts for the behaviour of the entrepreneur rather than the location so there may be deviations from the LCL.

Structuralist Models

Location is explained by the underlying structures of society. Therefore location is driven by changes in the National & World economy.

6

Industrial Location Theory

Alfred Weber identified the most significant factors which determine industrial location in 1929. It is important to remember the ideas which Von Thunen, Christaller and Losch focused upon in order to understand Weber's focus upon transportation influencing industrial location.

Similarly, please remember that INDUSTRIAL activity is considered a secondary economic activity, and is also discussed as manufacturing. Industrial activity can be broken down further to include the following activities:

* processing
* the creation of intermediate parts
* final assembly.

When Alfred Weber writes about industrial location in the 1920's, he is examining large companies whose industrial activities ALL take place within our national borders. This is not true of industrial activity today. With multi-national corporations, the three activities listed above may occur outside of the United States. Prior to 1983, all final assembly had to be completed in the USA or American corporations who imported finished products back into the USA paid a large duty (which inflated the cost of good to the consumer). Since 1983, beginning with the Maquiladora industries and continuing further with the NAFTA treaty, American corporations now pay little, or no duty at all, on the importation of finished products to the American market. We will discuss the impact of this legislation more deeply in the next lecture as we focus upon the globalization of manufacturing. But, please keep in mind that Weber's theory did not address industrial activity outside the United States.

His theory does have validity in explaining some of the causes for current movement, yet such discussion would not come from Weber himself.

Weber's industrial location theory demonstrates the unique way that geographers look at the world. I often tell my classes that Weber asked the following question;

"where is industry located? Why there?" much in the same manner that a physical geographer asks why a mountain is located where it is. Given this focus, Weber found through his examination of industrial activity that similar industries located in the area where he found it the least cost to produce. What this means is that you find industries that produce the same good, clustered in regions that enable them to reduce their costs of manufacturing (materials, labour, transport) and locate the largest market. This strategy would earn them the greatest volume of sales.

Least cost location then implies marketing the product at the least cost to the consumer. Much like retailers, such as Walmart, Target, and Costco, attempt to obtain large market shares today. It is explained economically as one way to make a profit, creating the cheapest product for the consumer market would lead to greater volume of sales and hence, greater profits. Therefore, companies which did not take the time to locate the cheapest inputs or the largest markets, would go out of business since their product would cost more to produce and cost the consumer more at the market.

It is important to understand that Weber's industrial location theory is an example of classical economic theory. As Smith and other classical economists thought, the consumer was ECONOMICALLY RATIONAL. What this means is that the market economy is based upon the concept that all consumers have total market knowledge (what goods are available and how much they cost to produce). Given such knowledge, EACH TIME you consume (demand) a product, you RATIONALLY select the good that is the cheapest cost of equal quality. This allows the market price to reach an equilibrium between supply and demand. Do you do this EVERY time you purchase something?

Given consumer rationality and corporate wants for profits, Weber found that the most successful businesses had located in regions which allowed the least cost of production to be actualized. Like businesses often followed the lead of pioneering companies, leading to the clustering of similar activities. Weber called this clustering agglomeration. We have previously examined this clustering when discussing suburban sprawl and megalopolis development (such as Boswash). Just remember agglomeration as the clustering of people and industry. Which type of cluster comes first is important when we start attempting to locate cheap labour or we begin to evaluate the impact of corporate movement from one location to another.

Weber's main point was that the cost of transport (another theory on this) determined the location of industry. Therefore, he uses Von Thunen's idea (that the cost of transport determines crop selection) and applies it to industry. Similar to Von Thunen, the weight of the raw materials and the weight of the end product (this difference is known as the material index) will determine the site of production depending upon how much the industry is willing to pay to get its product to the market (connecting to Christaller's ideas of market area).

Weber's theory rest primarily on four such sites, what he calls industrial orientations.

* Material orientation
* Labour orientation
* Transport orientation
* Market orientation.

Material Oriented Industries-These industries are called "weight-losing" or bulk industries, and it would be very expensive to haul raw materials to the market for processing, so that manufacturing occurs near the raw materials. Besides mining, other primary activities (or extractive industries) are considered material oriented; timber mills, furniture manufacture, most agricultural activities. The Pacific Northwest is a good example of material oriented industries. Often located in rural areas, these businesses may employ most of the population. As they leave (such as Weyerhauser, Pacific Timber...) entire cities lose their economic base.

Labour Oriented Industries-these are industries, such as the garment industry in New York or San Francisco which require cheap unskilled labour to complete activities that have not been mechanized, or high tech firms, such as those located in Silicon Valley, which require exceptionally skilled professionals. Either type of industry is finds that labour (quality or cost) is the primary determinant in production.

Unskilled Labour-is often termed "ubiquitous" by geographers. What this means is that unskilled labour is found everywhere. If your factory is dependent upon unskilled labour, when you build the factory, they will come (similar to the movie "field of Dreams"). These types of industries may be retail (malls, theaters), communications (such as MCI), and even in high tech industry, most circuit board manufacture requires unskilled labour. If you are the businessman evaluating potential locations in this type of industry, you will be trying to find the areas which provide:

* low wages
* little unionization
* young employees (few healthcare costs)
* female employees (they have babies which allows you to keep wages low, and they tend not to be as demanding)

Pages 354 and 355 give some specific examples of this comparing hourly labour costs in many countries, and demonstrating true labour costs choices which may determine business location.

Skilled Labour-industries founded on the use of skilled labour faced an opposite dilemma, skilled labour is very scarce and often difficult to find. Silicon Valley is a good example of this. In 1992, the Chronicle published an article which listed the percentage of residents throughout cities in the South Bay who possessed a bachelor's

degree. I recall that 70% of Mountain View residents had BA's and 88% of Los Altos residents possessed a BA. Compared to the national average of 12.5%, these numbers demonstrate the pull for industries using skilled labour to the Bay Area. Other activities which reflect skilled labour industries are corporate headquarters, and research and development centres. Skilled labour requires higher education so this allows first world countries to retain jobs in these industries.

Transport Oriented Industries-these are industries whose primary production cost is transport. Geographers term the location for where these industries locate as "break-in-bulk" locations. This term is used to describe a location where two or more modes of transportation may connect.

San Francisco is a great example of this, having port, rail, air, and highway linkages. Such diversity allows the business to use the cheapest transport available. Current examples of such industrial agglomeration is illustrated on the inset map 2.8, which portrays the clustering of businesses along Interstate 65 and 75. A similar pattern exists in Tracy where food distributors have built large warehouses (cheap land) for Northern California distribution. Tracy provides rail access and highway linkages to most market, so that relatively little trucking (2-3 hours) is required.

Market Oriented Industries-these are industries whose product may be weight-gaining, breakable or perishable. Despite the high costs of land and labour in market regions (large cities), it is to the advantage of such industries to locate as close to the consumer as possible. Such industries are Coca-cola or bottling industries, auto assembly plants, such as NUMI, dairies, such as Berkeley Farms in Sonoma (close to urban market due to perishability), and computers and television assembly (due to break-ability).

Overall, Weber's theory is amazing because he was able to demonstrate how major production costs (labour, materials, transport) determined industrial location. He clearly articulated the major factors which businessmen used to identify prime locations. These factors are still the dominant factors which explain industrial location today.

Weber's theory also allows us to ask questions concerning how industry could be relocated. Could one stimulate economic development in a depressed region by investing in transportation? or education? What is the role of the market in generating economic development?

Introduction to Human Geography

Location of Industry

Location of industry is concerned with the least cost location, so that again transport costs are a crucial element in the location decision. A German economist, Alfred Weber devised the theory of industrial location, in 1909. In the early part of

the industrial revolution factories developed in areas that were already producing manufactured goods. These were the places where woollen textiles were produced in farmhouses on farms that bred the sheep. The shift was from the farmhouse to a mill, in the same area. By chance many of these textile mills were on coalfields, so that when the shift from water powered to steam coal powered mills occurred, the transition was in the same place. The same thing happened as steel production was also shifted from local forges to coal powered mills. The early industrial revolution saw mills and factories develop on coalfields, and remain entrenched there for more than a century.

By the end of the 19th century, these raw material locations were losing their ascendancy. At the beginning of the industrial revolution roads were poor quality and slow. Canals were rapidly constructed to move heavy industrial materials, but nowhere did these form a really convenient network. It was the development of railways into extensive networks by the end of the 19th century, that enabled industrial location to free itself from raw material sites. This trend continued with roads and vehicles in the twentieth century, but water transport and especially the sea, remained dominant for long distance transport of industrial goods. Weber's analysis came at the point where railway networks had developed to their ultimate extent. He was therefore concerned with the balance of location between raw material site, the market for manufactured goods, and transport.

Some definitions

The factory or plant is an individual building or premises that produces manufactured goods. A company may own several factories, probably in different locations. The industry comprises many factories, or plants, and a number of independent companies. Industrial location is primarily concerned with the siting of a single factory, rather than the whole industry, although the location of the industry is in itself a locational factor. The concepts of site and situation play separate roles, although we may use the word site in relation to location when we are really looking at the situation of the factory.

The site of a factory, or group of factories, is the actual physical location, or block of land. There are some basic locational constraints for the site, such as a plentiful supply of flat land, access to transport, power and water, availability of labour, and capital and finance facilities. Almost all cities will possess appropriate industrial sites and these will be zoned by councils.

It is therefore the situation, or the relative location, in relation to other factories and the industry that is important. Like agricultural and central place location theories, Weber makes assumptions that simplify reality, but unlike these other theories, he does not assume an equal distribution. Rather he assumes that raw materials are unequally distributed in fixed locations.

Assumptions

1. There is an uneven distribution of natural resources on the plain. Raw materials are concentrated in specific sites.
2. The size and location of markets are given at fixed points on the plain.
3. There are fixed locations of labour where wage rates are fixed and labour is immobile and unlimited (capitalists love that).
4. The area has a uniform culture, climate and political system.
5. Entrepreneurs minimise costs of production.
6. Perfect competition exists.
7. Costs of land, structures, equipment and capital do not vary regionally.
8. There is a uniform system of transport over a flat surface.

Raw Material or Market Site

In the first instance we consider whether to locate an industry in the raw material or market location. If there is no weight loss or weight gain in production, you site your factory at either location, because the transport costs are the same each way. The diagram on the handout illustrates how isotims of equal cost distance are constructed around each site, thereby creating intersection points of equal but higher values, termed isodapanes.

As transport costs are not identical for raw material and manufactured goods a relative weighting must be calculated. Weber did this with a material index, whereby the relative weight gain or loss is calculated.

Material index = total weight of materials used to manufacture the product.

Total weight of the finished product.

If the product is a pure material its index will be 1. If the index is less than 1 the final product has gain weight in manufacture, thus favouring production at the market place. The weight gain is most likely to come from the addition of ubiquitous materials, like water, that we can expect to occur anywhere. Such a product would be a drink, soft drinks or beer, where a small quantity of usually dried materials are added to water and bottles to make a much heavier and more fragile final product. Most products lose weight in manufacture, such as a metal being extracted from an ore. Thus their material index will be more than 1, thus favouring the raw material site.

The significance of the material index is in calculating precisely the difference between the unit transport costs of raw materials and finished products. The number of the index is used to calculate a relative weighting, which is then applied to the spacing/radius of the isotims. The locational triangles on the handout are small examples of the weighting of more than one material.

While the drawing of isotims and isodapanes is very straightforward for 2 sites, it is in adding a number of material sites and markets that the spatial model both increases in complexity and begins to provide a useful method for calculating the least cost location. As well as weight loss or gain the material index and weighting of transport costs can also take account of loss or gain in transport, of features such as perishability, fragility and hazard.

Qualifications to Industrial Location

Companies make economies of scale and benefit from shared facilities, labour force, infrastructure, services and raw materials if they are sited in the same place as existing factories. This process of agglomeration concentrates many factories into industrial regions or zones.

At the other end of the process is industrial inertia. As changes occur in raw materials, decline of locational advantages and transport infrastructure, old industrial areas decline, but do not lose all of their industry or population. Factories close and unemployment soars, but people remain because it is their home. These areas decline gradually over a very long period and may regenerate with new light industry.

Light manufacturing or footloose industry is increasingly controlled by Trans National Corporations who make global location decisions based on economies in transport, wages and infrastructure. They poured investment into South East Asia, and then pulled it out. They drive much of the world economy and benefit from extremely efficient transportation that skews the locational decision towards the largest markets. While Indonesia riots, the capitalists race for China.

Rostovian Take-off Model

The Rostovian take-off model (also called "Rostow's Stages of Growth") is one of the major historical models of economic growth. It was developed by W. W. Rostow. The model postulates that economic growth occurs in five basic stages, of varying length.

1. Traditional society
2. Preconditions for take-off
3. Take-off
4. Drive to maturity
5. Age of High mass consumption

Rostow claimed that these stages of growth were designed to tackle a number of issues. Some of which he identified himself and wrote:

> *"Under what impulses did traditional, agricultural societies begin the process of their modernization? When and how did regular growth become a built in feature of each society? What forces drove the process of sustained growth along*

and determined its contours? What common social and political features of the growth process may be discerned at each stage? What forces have determined relations between the more developed and less developed areas; and what relation if any did the relative sequence of growth bear to outbreak of war? And finally where is compound interest taking us? Is it taking us to communism; or to the affluent suburbs, nicely rounded out with social overhead capital; to destruction; to moon; or where?"

Rostow asserts that countries go through each of these stages fairly linearly, and set out a number of conditions that were likely to occur in investment, consumption and social trends at each state. Not all of the conditions were certain to occur at each stage, however, and the stages and transition periods may occur at varying lengths from country to country, and even from region to region.

Rostow's model is one of the more structuralist models of economic growth, particularly in comparison with the 'backwardness' model developed by Alexander Gerschenkron. The two models are not necessarily mutually exclusive, however, and many countries seem to follow both models rather adequately.

Beyond the structured picture of growth itself, another important part of the model is that economic take-off must initially be led by a few individual sectors. This belief echoes David Ricardo's comparative advantage thesis and criticizes Marxist revolutionaries push for economic self-reliance in that it pushes for the 'initial' development of only one or two sectors over the development of all sectors equally. This became one of the important concepts in the theory of modernization in the social evolutionism.

Theoretical Framework

Rostow's model is a part of the liberal school of economics, laying emphasis on the efficacy of modern concepts of free trade and the ideas of Adam Smith. It disagrees with Friedrich List's argument which states that economies which rely on exports of raw materials may get "locked in", and would not be able to diversify, regarding this Rostow's model states that economies may need to depend on raw material exports to finance the development of industrial sector which has not yet of achieved superior level of competitiveness in the early stages of take-off. Rostow's model does not disagree with John Maynard Keynes regarding the importance of government control over domestic development which is not generally accepted by some ardent free trade advocates. The basic assumption given by Rostow is that countries want to modernize and grow and that society will agree to the materialistic norms of economic growth.

Stages

Traditional Societies

Traditional societies are marked by their pre-Newtonian understanding and use

of technology. Economy in this stage has a limited production function which could barely attains basic minimum level of output. It does not entirely mean that the economy's production level is static. The output level can still be increased as there is surplus of cultivable land which can be used for increasing agricultural produce. The state as well as the farmers were aware of the irrigation methods and expanded this facility. There were technological innovations but only on ad hoc basis. All this resulted in increase in output but there always existed an upper limit which could never be crossed which basically was because there was lack of application and constant development of modern science and technology. Trade was done in barter system and the monetary system was not well developed. The investment level is less than 5%.

There were numerous changes in size of population and quality of life because of wars, famines due to crop failures and epidemics like plague. Volume fluctuation in trade was due to political stability or instability. Manufacturing sector and other industries had a tendency to grow but were limited by the inadequate scientific knowledge and backward frame of mind which also lead to low labour productivity. In this stage self sufficient regions are present.

Due to agricultural dominance there was establishment of a hierarchical social structure in which there was no vertical mobility. This resulted in concentration of political power in the hand of land owners and family became a major institution with high importance placed on clan connection in the social organisation. This social structure was feudalistic in nature.

Pre-conditions to Take-off

In the second stage of economic growth the economy undergoes a process of change for building up of conditions for growth and take off. Walt Whitman Rostow said that these changes in society and the economy had to be of fundamental nature in the socio political structure and production technique. This pattern was followed in Europe, parts of Asia, Middle East and Africa. There is also a second pattern in which he said that there was no need for change in socio political structure because these economies were not deeply caught up in social and political structures. The only changes required were in economic and technical dimensions. The nations which followed this pattern were North America, New Zealand and Australia.

There are three important dimensions to this transition are firstly shift from agrarian to industrial or manufacturing society. Secondly trade and other commercial activities of the nation should broaden market reach to not only local areas but also international markets. Lastly the surplus attained should not be wasted on the conspicuous consumption of the land owners but should be spent on development of industries, infrastructure and also preparation of self sustained growth. Further more agriculture becomes commercialized and mechanized to to technological advancement and there is growth of entrepreneurship.

The strategic factor is that investment level should be above 5% of the national income. This rise in investment rate depends on many sectors of the economy. According to Walt Whitman Rostow capital formation depends on productivity of agriculture and creation of social overhead capital. Agriculture plays a very important role in this transition process as the surplus quantity of the produce is to be utilized to support urban population and also be a major part of export to earn foreign exchange for development. Increase in agricultural productivity will lead to expansion of domestic markets for manufacturing commodities which will further lead to induced investment in the industrial sector. Further agricultural surplus should be invested in modernization.

Social overhead capital creation would have to be taken only by the government. The Government Would have to play a major role in it as social overhead capital is lumpy in nature, it has long gestation period and indirect routes of pay-offs thus the private sector would not be interested in playing any role in it's development.

All these changes would be effective only if there is basic change in attitude of society towards risk taking, changes in work environment and changes in social and political organisations and structures. Basically the pre-conditions of take-off is beginning of the industrial revolution.

Take-off

This stage is characterized by dynamic economic growth. Which was as Rostow suggests began due to a sharp stimulus of either economic, political or technological in nature. The main feature of this is the self sustained growth. Take-off then occurs when sector led growth becomes common and society is driven more by economic processes than traditions. At this point, the norms of economic growth are well established and growth becomes a nations second nature. In discussing the take-off, Rostow is noted to have adopted the term "transition", which is to describe the process change of a traditional economy to a modern economy. After take-off, a country will take as long as fifty to one hundred years to reach maturity. Globally, this stage occurred during the Industrial Revolution. As per Rostow the requirements for a nation to take-off are the following three main conditions:

1. *"The rate of productive investment should rise from approximately 5% to over 10% of national income or net national product (this happened in Canada before the 1890's and Argentina before 1914);*
2. The development of one or more substantial manufacturing sectors, with a high rate of growth;
3. The existence or quick emergence of a political, social and institutional framework which exploits the impulses to expansion in the modern sector and the potential external economy effects of the take-off". *This condition means capability of mobilizing capital from domestic resources.*

Industrialization becomes a crucial phenomenon as it helps to prepare the basic structure for structural changes in a massive scale. Rostow says that this transition does not follow a set trend as there are a variety of different motivations or stimulus which began this growth process. Take off requires a large and sufficient amount of loanable funds for expansion of the industrial sector which generally come from two sources which are:

1. Shifts in income flows by way of taxation, implementation of land reforms and various other fiscal measures.
2. Ploughing back profits which depends on profits earned from foreign trade as noticed in many countries the other ways are expanding of demand of various domestically produced goods in the domestic markets. This was noticed in mostly all the countries except in US, Canada, Russia and Sweden where in their major source was by capital imports. This entire process of expansion of the industrial sector yields an increase in rate of return to some individuals who save at high rates and invest their savings in the industrial sector activities. The economy exploit their underutilized natural resources to increase their production.

The take-off also needs a group of entrepreneurs in the society who would induce innovation and help produce growth in the economy. The entrepreneurial activity is conducive to firstly a proper set up of a valuation system. Secondly, they need to feel that they can not secure prestige and power in the society and lastly, the society should tolerate the unorthodox path of attainment of materialistic and political goals. The success of passing through this stage depends on the following major factors:

- Existence of enlarged effective demand for the product of key sectors.
- Introduction of new production function in these sectors.
- Society's increased capacity to generate enough funds to initiate the take-off transition.
- The activities in the key sector should induce chain of growth activity in other sectors of the economy.

In the table note that Take-off periods of different countries are the same as the industrial revolution in those countries.

Drive to Maturity

After take-off there follows a long interval of sustained if fluctuating progress, as the now regularly growing economy drives to extend modern technology over the whole front of its economic activity. Some 10-20% of the national income is steadily invested, permitting output regularly to outstrip the increase in population. The make-up of the economy changes unceasingly as technique improves, new industries accelerate, older industries level off. The economy finds its place in the international

economy: goods formerly imported are produced at home; new import requirements develop, and new export commodities to match them. The society makes such terms as it will with the requirements of modern efficient production, balancing off the new against the older values and institutions, or revising the latter in such ways as to support rather than to retard the growth process. The drive to maturity refers to the need for the economy itself to diversify. The sectors of the economy which lead initially begin to level off, while other sectors begin to take off. This diversity leads to greatly reduced rates of poverty and rising standards of living, as the society no longer needs to sacrifice its comfort in order to strengthen certain sectors.

Age of High Mass Consumption

The age of high mass consumption refers to the period of contemporary comfort afforded many western nations, wherein consumers concentrate on durable goods, and hardly remember the subsistence concerns of previous stages. Rostow uses the Buddenbrooks dynamics metaphor to describe this change in attitude.

In Thomas Mann's novel, Buddenbrooks, a family is chronicled for three generations. The first generation is interested in economic development, the second in its position in society. The third, already having money and prestige, concerns itself with the arts and music, worrying little about those previous, earthly concerns. So too, in the age of high mass consumption, a society is able to choose between concentrating on military and security issues, on equality and welfare issues, or on developing great luxuries for its upper class. Each country in this position chooses its own balance between these three goals.

Of particular note is the fact that Rostow's "Age of High Mass Consumption" dovetails with (occurring before) Daniel Bell's hypothesized "Post-Industrial Society." The Bell and Rostovian models collectively suggest that economic maturation inevitably brings on job-growth which can be followed by wage escalation in the secondary economic sector (manufacturing), which is then followed by dramatic growth in the tertiary economic sector (commerce and services).

In the Bell model, the tertiary economic sector rises to predominance, encompassing perhaps 65 to 75 percent of the employment in a given economy. Maturation can then bring-on deindustrialization as manufacturers reorient to cheaper labour markets, and deindustrialization can, in turn, destabilize the tertiary sector. The suggestion is that mature economies may implicitly destabilize and cycle back-and-forth between the final stages of the Rostovian-Bell developmental phases as they rebalance themselves, over time, and re-evolve their economic base.

Criticism of the Model

1: Rostow is historical in the sense that the end result is known at the outset and is derived from the historical geography of a developed, bureaucratic society.

2: Rostow is mechanical in the sense that the underlying motor of change is not disclosed and therefore the stages become little more than a classificatory system based on data from developed countries.

3: His model is based on American and European history and defines the American norm of high mass consumption as integral to the economic development process of all industrialized societies.

4: His model assumes the inevitable adoption of Neoliberal trade policies which allow the manufacturing base of a given advanced polity to be relocated to lower-wage regions.

Rostow's thesis is biased towards a western model of modernization, but at the time of Rostow the world's only mature economies were in the west, and no controlled economies were in the "era of high mass consumption." The model de-emphasizes differences between sectors in capitalistic vs. communistic societies, but seems to innately recognize that modernization can be achieved in different ways in different types of economies.

The most disabling assumption that Rostow is accused of is trying to fit economic progress into a linear system. This charge is correct in that many countries make false starts, reach a degree of transition and then slip back, or as is the case in contemporary Russia, slip back from high mass consumption (or almost) to a country in transition. On the other hand, Rostow's analysis seems to emphasize success because it is trying to explain success. To Rostow, if a country can be a disciplined, uncorrupt investor in itself, can establish certain norms into its society and polity, and can identify sectors where it has some sort of advantage, it can enter into transition and eventually reach modernity. Rostow would point to a failure in one of these conditions as a cause for non-linearity.

Another problem that Rostow's work has is that it considers mostly large countries: countries with a large population (Japan), with natural resources available at just the right time in its history (Coal in Northern European countries), or with a large land mass (Argentina). He has little to say and indeed offers little hope for small countries, such as Rwanda, which do not have such advantages. Neo-liberal economic theory to Rostow, and many others, does offer hope to much of the world that economic maturity is coming and the age of high mass consumption is nigh. But that does leave a sort of 'grim meathook future' for the outliers, which do not have the resources, political will, or external backing to become competitive.

Rostow's Model [Based on a Virtual Developing Country]

The Virtual Developing Country is a case study of Zambia. There are a series of field trips available looking at different issues connected with economic development. This trip is the Copper Tour and this page looks at Rostow''s Model-the Stages of Economic Development.

Theories

Rostow's Model-the Stages of Economic Development

Next theory-Models of Demographic Transition

In 1960, the American Economic Historian, WW Rostow suggested that countries passed through five stages of economic development.

Stage 1 Traditional Society : The economy is dominated by subsistence activity where output is consumed by producers rather than traded. Any trade is carried out by barter where goods are exchanged directly for other goods. Agriculture is the most important industry and production is labour intensive using only limited quantities of capital. Resource allocation is determined very much by traditional methods of production.

Stage 2 Transitional Stage (the preconditions for takeoff) : Increased specialisation generates surpluses for trading. There is an emergence of a transport infrastructure to support trade. As incomes, savings and investment grow entrepreneurs emerge. External trade also occurs concentrating on primary products.

Stage 3 Take Off : Industrialisation increases, with workers switching from the agricultural sector to the manufacturing sector. Growth is concentrated in a few regions of the country and in one or two manufacturing industries. The level of investment reaches over 10% of GNP.

The economic transitions are accompanied by the evolution of new political and social institutions that support the industrialisation. The growth is self-sustaining as investment leads to increasing incomes in turn generating more savings to finance further investment.

Stage 4 Drive to Maturity : The economy is diversifying into new areas. Technological innovation is providing a diverse range of investment opportunities. The economy is producing a wide range of goods and services and there is less reliance on imports.

Stage 5 High Mass Consumption : The economy is geared towards mass consumption. The consumer durable industries flourish. The service sector becomes increasingly dominant.

According to Rostow development requires substantial investment in capital. For the economies of LDCs to grow the right conditions for such investment would have to be created. If aid is given or foreign direct investment occurs at stage 3 the economy needs to have reached stage 2. If the stage 2 has been reached then injections of investment may lead to rapid growth.

Limitations

Many development economists argue that Rostows's model was developed with

Western cultures in mind and not applicable to LDCs. It addition its generalised nature makes it somewhat limited. It does not set down the detailed nature of the pre-conditions for growth. In reality policy makers are unable to clearly identify stages as they merge together. Thus as a predictive model it is not very helpful. Perhaps its main use is to highlight the need for investment. Like many of the other models of economic developments it is essentially a growth model and does not address the issue of development in the wider context.

Describe and Critically Analyse Rostow's Theory of Growth?

The Rostovian take-off model (also called "Rostow's Stages of Growth") is one of the major historical models of economic growth. It was developed by W. W. Rostow. The model postulates that economic modernization occurs in five basic stages, of varying length.

1. Traditional society
2. Preconditions for take-off
3. Take-off
4. Drive to maturity
5. Age of High mass consumption

Rostow asserts that countries go through each of these stages fairly linearly, and set out a number of conditions that were likely to occur in investment, consumption and social trends at each state. Not all of the conditions were certain to occur at each stage, however, and the stages and transitions periods may occur at varying lengths from country to country, and even from region to region.

Rostow's model is one of the more structuralist models of economic growth, particularly in comparison with the 'backwardness' model developed by Alexander Gerschenkron. The two models are not necessarily mutually exclusive, however, and many countries seem to follow both models rather adequately.

Beyond the structured picture of growth itself, another important part of the model is that economic take-off must initially be led by a few individual sectors. This belief echoes David Ricardo's comparative advantage thesis and criticizes Marxist revolutionaries push for economic self-reliance in that it pushes for the 'initial' development of only one or two sectors over the development of all sectors equally. This became one of the important concepts in the theory of modernization in the social evolutionism.

Theoretical Framework **:** Rostow's model is descendent from the liberal school of economics, emphasizing the efficacy of modern concepts of free trade and the ideas of Adam Smith. It also denies Friedrich List's argument that countries reliant on exporting raw materials may get "locked in", and be unable to diversify, in that Rostow's model states that countries may need to depend on a few raw material

exports to finance the development of manufacturing sectors which are not yet of superior competitiveness in the early stages of take-off. In that way, Rostow's model does not deny John Maynard Keynes in that it allows for a degree of government control over domestic development not generally accepted by some ardent free trade advocates. Although empirical at times, Rostow is hardly free of normative discourse. As a basic assumption, Rostow believes that countries want to modernize as he describes modernization, and that the society will ascent to the materialistic norms of economic growth.

Traditional Societies **:** Traditional societies are marked by their pre-Newtonian understanding and use of technology. These are societies which have pre-scientific understandings of gadgets, and believe that gods or spirits facilitate the procurement of goods, rather than man and his own ingenuity. The norms of economic growth are completely absent from these societies.

Preconditions to Take-off **:** The preconditions to take-off are, to Rostow, that the society begins committing itself to secular education, that it enables a degree of capital mobilization, especially through the establishment of banks and currency, that an entrepreneurial class form, and that the secular concept of manufacturing develops, with only a few sectors developing at this point. This leads to a take off in ten to fifty years. At this stage, there is a limited production function, and therefore a limited output. There are limited economic techniques available and these restrictions create a limit to what can be produced.

Take-off **:** Take-off then occurs when sector led growth becomes common and society is driven more by economic processes than traditions. At this point, the norms of economic growth are well established. In discussing the take-off, Rostow's is a noted early adopter of the term "transition", which is to describe the passage of a traditional to a modern economy. After take-off, a country will take as long as fifty to one hundred years to reach maturity.

Drive to Maturity **:** The drive to maturity refers to the need for the economy itself to diversify. The sectors of the economy which lead initially begin to level off, while other sectors begin to take off. This diversity leads to greatly reduced rates of poverty and rising standards of living, as the society no longer needs to sacrifice its comfort in order to strengthen certain sectors.

Age of High Mass Consumption **:** The age of high mass consumption refers to the period of contemporary comfort afforded many western nations, wherein consumers concentrate on durable goods, and hardly remember the subsistence concerns of previous stages. Rostow uses the Buddenbrooks dynamics metaphor to describe this change in attitude. In Thomas Mann's novel, Buddenbrooks, a family is chronicled for three generations. The first generation is interested in economic development, the second in its position in society. The third, already having money and prestige, concerns itself with the arts and music, worrying little about those

previous, earthly concerns. So too, in the age of high mass consumption, a society is able to choose between concentrating on military and security issues, on equality and welfare issues, or on developing great luxuries for its upper class. Each country in this position chooses its own balance between these three goals.

Criticism of the Model

Rostow's thesis assumes a strong bias towards a western model of modernization. It de-emphasizes any difference between how leading sectors develop in free and controlled markets. However, Rostow's consideration of non-western cases such as China shows that to some extent, modernization can be achieved in different ways and through free market or controlled economic means and still fit into his model. It is more at his description of the final age, the age of high mass consumption, where controlled economies seem most to find no niche in Rostow's work. Even there, though, it could be said that the society seeks out economic equality at the complete detriment of any luxury.

The most disabling assumption that Rostow is accused of is trying to fit economic progress into a linear system. This charge is correct in that many countries make false starts, reach a degree of transition and then slip back, or as is the case in contemporary Russia, slip back from high mass consumption (or almost) to a country in transition. On the other hand, Rostow's analysis seems to emphasize success because it is trying to explain success. To Rostow, if a country can be a disciplined, uncorrupt investor in itself, can establish certain norms into its society and polity, and can identify sectors where it has some sort of advantage, it can enter into transition and eventually reach modernity. Rostow would point to a failure in one of these conditions as a cause for non-linearity.

Another problem that Rostow's work has is that it considers mostly large countries: countries with a large population (Japan), with natural resources available at just the right time in its history (Coal in Northern European countries), or with a large land mass (Argentina). He has little to say and indeed offers little hope for small countries, such as Rwanda, which do not have such advantages. Neo-liberal economic theory to Rostow, and many others, does offer hope to much of the world that economic maturity is coming and the age of high mass consumption is high. But that does leave a sort of 'grim meathook future' for the outliers, which do not have the resources, political will, or external backing to become competitive.

International Trade Model of Development

Other names: Rostow's Development Model.

Creator: Walt Whitman Rostow 1916-2003 was an American economist who proposed his five stage model of development in the 1950's, the ideas of which stemmed from modern free trade and Adam Smith. Rostow's model does not deny

John Maynard Keynes in that it allows for a degree of government control over domestic development not generally accepted by some ardent free trade advocates. Although empirical at times, Rostow is hardly free of normative discourse. As a basic assumption, Rostow believes that countries want to modernize as he describes modernization, and that society will assent to the materialistic norms of economic growth.

Purpose: requires a country to identify its distinctive or unique economic resources. The model puts forth the idea that a country can develop economically by concentrating on resources in short supply to expand beyond local industries to reach the global market and finance the country's further development.

The Five Stages of Development:

1. Traditional Society-Refers to a country that has yet to begin developing, where a high percentage of people are involved with agriculture and a high percentage of the country's wealth is invested in activities such as the military and religion, seen as "nonproductive" by Rostow. These are societies which have pre-scientific understandings of gadgets, and believe that gods or spirits facilitate the procurement of goods, rather than man and his own ingenuity.
2. Transitional Stage-AKA the preconditions for takeoff. Under the model, the process of development begins when an elite group initiates innovations economic activities. Under the influence of these well-educated leaders, the country starts to invest in new technology and infrastructure, such as water supplies and transportation systems. These projects will ultimately stimulate an increase in productivity likely increasing the GDP. There is a limited production function, and therefore a limited output. There are limited economic techniques available and these restrictions create a limit to what can be produced. Increased specialization generates surpluses for trading. There is an emergence of a transport infrastructure to support trade. External trade also occurs concentrating on primary products.
3. Takeoff-Rapid growth is generated in a limited number of economic activities, such as textiles or food products. These few, takeoff industries achieve technical advances and become productive, whereas other sectors of the economy remain dominated by traditional practices. After take-off, a country will take as long as fifty to one hundred years to reach maturity. Globally, this stage occurred during the Industrial Revolution. Industrialization increases, with workers switching from the agricultural sector to the manufacturing sector. The level of investment reaches over 10% of GNP. The growth is self-sustaining as investment leads to increasing incomes in turn generating more savings to finance further investment.
4. Drive to maturity-Modern technology, previously confined to a few takeoff industries, diffuses to a wide variety of industries, which then experience

rapid growth comparable to the takeoff industries. Workers become more skilled and specialized. The economy is diversifying into new areas the economy is producing a wide range of goods and services and there is less reliance on imports.

5. High Mass Consumption-AKA age of mass consumption. The economy shifts from production of heavy industry such as steel and energy, to consumer goods, such as motor vehicles and refrigerators. Of particular note is the fact that Rostow's "Age of High Mass Consumption" dovetails with (occurring before) Daniel Bell's hypothesized "Post-Industrial Society." The Bell and Rostovian models collectively suggest that economic maturation inevitably brings on job-growth which can be followed by wage escalation in the secondary economic sector (manufacturing), which is then followed by dramatic growth in the tertiary economic sector (commerce and services).

Main Points & Examples

Rostow's development model was based on two factors. First, the developed countries of Western Europe and Anglo-America? had been joined by others in Southern and Eastern Europe and Japan. Second, many LDCs contain an abundant supply of raw materials sought by manufacturers and producers in MDCs. In the past, European colonial powers extracted many of these resources without paying compensation to the colonies, as core countries do to periphery. In a global economy, the sale of these raw materials could generate funds for LDCs to promote development.

According to the model, each country is in one of these five stages of development. With MDC's in stage 4 or 5, whereas LDCs are in one of the three earlier stages. The model asserts that today's MDC's passed through the other stages in the past. For example, the U.S. was in stage 1 prior to independence, stage 2 during the 1st half of the 1800's, stage 3 during the middle of the 1880's, and stage 4 during the late 1800's, before entering stage 5 during the early 1900's. The model assumes that LDCs will achieve development by moving along from an earlier to a later stage.

A country that concentrates on international trade benefits from exposure to consumers in other countries. To remain competitive, the takeoff industries must constantly evaluate changes in international consumer preferences, marketing strategies, production engineering, and design technologies.

Examples of countries adopting this method of development include areas in East/Southeast Asia and Arabian Peninsula.

In Southeast Asia, a group of countries, Singapore, Taiwan, South Korea, and the former British colony of Hong Kong came to be known as the "four dragons" after adopting the international trade approach. They were lacking in natural resources so they promoted development by concentrating on producing a handful of manufactured goods, especially clothing and electronics. Low labour costs enabled these countries

to sell products inexpensively in MDCs. The countries of the Arabian Peninsula, which includes Saudi Arabia, Kuwait, Bahrain, Oman, and the United Arab Emirates, went from LDC's to some of the wealthiest countries almost overnight due to increased petroleum prices during the 1970's. Arabian Peninsula countries have used petroleum revenues to finance large-scale projects, such as housing, highways, airports, universities, and telecommunications networks.

Weaknesses:

1: Rostow is 'historical in the sense that the end result is known in the outset and is derived from the historical geography of developed society.
2: Rostow is mechanical in the sense the underlying motor of change is not disclosed and therefore the stages become little more than a classificatory system based on data from developed countries.
3: His model is based on American and European history and aspiring to American norm of high mass consumption.
4: His model represents a "non-communist manifesto" or we can say a "capitalist manifesto".

Rostow's thesis is biased towards a western model of modernization, but at the time of Rostow the world's only mature economies were in the west, and no controlled economies were in the "era of high mass consumption." The model de-emphasizes differences between sectors in capitalistic vs. communistic societies, but seems to innately recognize that modernization can be achieved in different ways in different types of economies.

The most disabling assumption that Rostow is accused of is trying to fit economic progress into a linear system. This charge is correct in that many countries make false starts, reach a degree of transition and then slip back, or as is the case in contemporary Russia, slip back from high mass consumption (or almost) to a country in transition. On the other hand, Rostow's analysis seems to emphasize success because it is trying to explain success. To Rostow, if a country can be a disciplined, uncorrupt investor in itself, can establish certain norms into its society and polity, and can identify sectors where it has some sort of advantage, it can enter into transition and eventually reach modernity. Rostow would point to a failure in one of these conditions as a cause for non-linearity.

Another problem is that Rostow's work considers mostly large countries: countries with a large population (Japan), with natural resources available at just the right time in its history (Coal in Northern European countries), or with a large land mass (Argentina). He has little to say about small countries, such as Rwanda, which do not have such advantages. Neo-liberal economic theory to Rostow, and many others, does offer hope to much of the world that economic maturity is coming and the age of high mass consumption is nigh.

Limitations

Many development economists argue that Rostows's model was developed with Western cultures in mind and not applicable to LDCs. It addition its generalized nature makes it somewhat limited. It does not set down the detailed nature of the pre-conditions for growth. In reality, policy makers are unable to clearly identify stages as they merge together. Thus as a predictive model it is not very helpful. Perhaps its main use is to highlight the need for investment. Like many of the other models of economic developments it is essentially a growth model and does not address the issue of development in the wider context.

Rostow's thesis is biased towards a western model of modernization, but at the time of Rostow the world's only mature economies were in the west, and no controlled economies were in the "era of high mass consumption." The model de-emphasizes differences between sectors in capitalistic vs. communistic societies, but seems to innately recognize that modernization can be achieved in different ways in different types of economies.

The most disabling assumption that Rostow is accused of is trying to fit economic progress into a linear system. This charge is correct in that many countries make false starts, reach a degree of transition and then slip back, or as is the case in contemporary Russia, slip back from high mass consumption (or almost) to a country in transition. On the other hand, Rostow's analysis seems to emphasize success because it is trying to explain success. To Rostow, if a country can be a disciplined, uncorrupt investor in itself, can establish certain norms into its society and polity, and can identify sectors where it has some sort of advantage, it can enter into transition and eventually reach modernity. Rostow would point to a failure in one of these conditions as a cause for non-linearity.

Another problem that Rostow's work has is that it considers mostly large countries: countries with a large population (Japan), with natural resources available at just the right time in its history (Coal in Northern European countries), or with a large land mass (Argentina). He has little to say and indeed offers little hope for small countries, such as Rwanda, which do not have such advantages. Neo-liberal economic theory to Rostow, and many others, does offer hope to much of the world that economic maturity is coming and the age of high mass consumption is nigh. But that does leave a sort of 'grim meathook future' for the outliers, which do not have the resources, political will, or external backing to become competitive.

Self-sufficiency

China, India and most African and Eastern European countries adopted this strategy at one time. The idea is to protect local, fledgling businesses from large, international competition. This also helps to make your country independent of the MDCs and not at the whim of TNCs.

Elements of self-sufficiency approach-Import limitation.

Higher taxes on imported goods (tariffs).

Set quotas on imports.

Import-license requirements.

India once did all of these and even made it illegal to exchange their money on currency exchanges.

The government wanted businesses to produce for India only (local businesses that is). If private companies could not make a profit, the government subsidized them.

Nicholas J. Spykman

Nicholas John Spykman (1893 – 1943) was a Dutch-American geostrategist, known as the "godfather of containment." As a political scientist he was one of the founders of the classical realist school in American foreign policy, transmitting Eastern European political thought into the United States. A Sterling Professor of International Relations, teaching as part of the Institute for International Studies at Yale University, one of his prime concerns was making his students geographically literate—geopolitics was impossible without geographic understanding. He was married to the children's novelist E. C. Spykman. He died of cancer at the age of 49.

Spykman published two books on foreign policy. *America's Strategy in World Politics* was published in 1942 near the entry of the United States into World War II. Concerned with balance of power, he argues that isolationism, relying on the oceans to protect the United States ("hemispheric" or "quarter defence"), was bound to fail. His object was to prevent a U.S. retreat, similar to U.S. policy following World War I. *The Geography of the Peace* was published the year after Spykman's death. In it he lays out his geostrategy, arguing that the balance of power in Eurasia directly affected United States security.

In his writings concerning geography and foreign policy, Spykman was somewhat of a geographical determinist. Since geography was "the most fundamentally conditioning factor because of its relative permanence," it was of primary relevance in analyzing a state's potential foreign policy.

Spykman's Geostrategy

N.J. Spykman could be considered as a disciple and critic of both geostrategists Alfred Mahan, of the United States Navy, and Halford Mackinder, the British geographer. His work is based on assumptions similar to Mackinder: the unity of world politics, and the unity of the world sea. He extends this to include the unity of the air. The exploration of the entire world means that the foreign policy of any nation will affect more than its immediate neighbours; it will affect the alignment of nations throughout the world's regions. Maritime mobility opened up the possibility of a new geopolitical structure: the overseas empire.

Spykman adopts Mackinder's divisions of the world, renaming some:

- the Heartland;
- the Rimland (analogous to Mackinder's "inner or marginal crescent").

Heartland

At the same time, because he gives credit to the strategic importance of maritime space and coastal regions, Spykman's analysis of the heartland is markedly different from Mackinder's. He does not see it as a region which will be unified by powerful transportation or communication infrastructure in the near future. As such, it won't be in a position to compete with the United States' sea power. Spykman agrees that the heartland offers a uniquely defensive position, but that is all Spykman grants the occupier of the heartland.

While the USSR encompassed a great expanse of land, its arable land remained in a small portion of its territory, mostly in the West. Indeed, the Soviet's raw materials were largely located to the West of the Ural mountains as well. Since the political and material centre of gravity was in the Western part of the USSR, Spykman sees little possibility of the Soviets exerting much power in Central Asia.

Still, Russia was to remain the greatest land power in Asia, and could be a peacekeeper or a problem.

Rimland

The Rimland (Mackinder's "Inner or Marginal Crescent") was divided into three sections:

- the European coast land;
- the Arabian-Middle Eastern desert land;
- the Asiatic monsoon land.

While Spykman accepts the first two as defined, he rejects the simple grouping of the Asian countries into one "monsoon land." India, the Indian Ocean littoral, and Indian culture were geographically and civilizationally separate from the Chinese lands. The Rimland's defining characteristic is that it is an intermediate region, lying between the heartland and the marginal sea powers. As the amphibious buffer zone between the land powers and sea powers, it must defend itself from both sides, and therein lies its fundamental security problems. Spykman's conception of the Rimland bears greater resemblance to Alfred Thayer Mahan's "debated and debatable zone" than to Mackinder's inner or marginal crescent.

The Rimland has great importance coming from its demographic weight, natural resources, and industrial development. Spykman sees this importance as the reason that the Rimland will be crucial to containing the Heartland (whereas Mackinder had believed that the Outer or Insular Crescent would be the most important factor in the Heartland's containment).

Offshore Continents

There are two offshore continents flanking Eurasia: Africa and Australia. Spykman

sees the two continents' geopolitical status as determined respectively by the state of control over the Mediterranean Sea and the "Asiatic Mediterranean." Neither has ever been the seat of significant political power — chaos prevents Africa from harnessing the resources of its tropical regions; Australia hasn't enough arable territory to serve as a base of power.

Other than the two continents there are offshore islands of significance are Britain and Japan, while the New World, buffered by the Atlantic and Pacific Oceans.

Eurasian Dynamics

Again, Spykman differs from Mackinder. Mackinder sees Eurasian wars as historically pitting the heartland against the sea powers for control of the rimland, establishing a land power-sea power opposition. Spykman states that historically battles have pitted Britain and rimland allies against Russia and its rimland allies, or Britain and Russia together against a dominating rimland power. In other words, the Eurasian struggle was not the sea powers containing the heartland, but the prevention of any power from ruling the rimland.

Spykman recalls Mackinder's famous dictum,

Who controls eastern Europe rules the Heartland;

Who controls the Heartland rules the World Island;

Who rules the World Island rules the World,

But disagrees, refashioning it thus:

Who controls the rimland rules Eurasia;

Who rules Eurasia controls the destinies of the world.

Therefore, British, Russian, and U.S. power would play the key roles in controlling the European littoral, and thereby, the essential power relations of the world.

U.S. Strategic Goals

Spykman thought that it was in U.S. interests to leave Germany strong after World War II in order to be able to counter Russia's power. Strategically, there was no difference between Germany dominating all the way to the Urals, or Russia controlling all the way to Germany; both scenarios were equally threatening to the U.S.

Spykman predicted that Japan would lose the war in the Pacific, while China and Russia would remain to struggle against one another over boundaries. He also forecast the rise of China, becoming the dominant power in Asia, causing the U.S. to take responsibility for Japan's defence.

Spykman was opposed to European integration and argued that U.S. interests favored balanced power in Europe rather than integrated power. The U.S. was fighting

a war against Germany to prevent Europe's conquest—it would not make sense to federalize and thereby unify Europe after a war fought to preserve balance.

John Foster Dulles and the founders of U.S. containment strategy would borrow heavily from Spykman, as well as Mackinder, when forging U.S. Cold War strategy.

Quotations

- "Geography is the most fundamental factor in foreign policy because it is the most permanent." —from *The Geography of the Peace*
- "Plans for far-reaching changes in the character of international society are an intellectual by-product of all great wars." —from *America's Strategy in World Politics*
- "There are not many instances in history which show great and powerful states creating alliances and organizations to limit their own strength. States are always engaged in curbing the force of some other state. The truth of the matter is that states are interested only in a balance which is in their favour. Not an equilibrium, but a generous margin is their objective. There is no real security in being just as strong as a potential enemy; there is security only in being a little stronger. There is no possibility of action if one's strength is fully checked; there is a chance for a positive foreign policy only if there is a margin of force which can be freely used. Whatever the theory and rationalization, the practical objective is the constant improvement of the state's own relative power position. The balance desired is the one which neutralizes other states, leaving the home state free to be the deciding force and the deciding voice." —from *America's Strategy in World Politics*
- "[A] political equilibrium is neither a gift of the gods nor an inherently stable condition. It results from the active intervention of man, from the operation of political forces. States cannot afford to wait passively for the happy time when a miraculously achieved balance of power will bring peace and security. If they wish to survive, they must be willing to go to war to preserve a balance against the growing hegemonic power of the period." —from *America's Strategy in World Politics*
- "Nations which renounce the power struggle and deliberately choose impotence will cease to influence international relations either for evil or good"
- "Geographic facts do not change, but their meaning for foreign policy will"

7

Introductory Notes to Heartlands and Rimlands

In reading the two articles in this section, there are a number of things that you should keep in mind.

In 1904, Sir Halford Mackinder published a theory about political strength in Eurasia. He analysed the strengths and weaknesses of the various regions and concluded that the Russian Core and areas east of the core contained the potential to become a world power. In 1919 he revised his theory to include Eastern Europe and the theory became known as Mackinder's Heartland Theory.

Basically stated:

- Who rules East Europe commands the heartland.
- Who rules the heartland commands the World Island (Eurasia and Africa).
- Who rules the World Island commands the World.

In 1942, Nicholas Spykman proposed a theory which countered Mackinder's Heartland Theory. Spykman stated that Eurasia's Rimland, the coastal areas or buffer zone, is the key to controlling the World Island, not the heartland.

Nicholas Spykman's book, "America's Strategy in World Politics" was published during World War II and deals with political and military strength of the United States at that time. Today we look at the Rimland in terms of its economic strength and potential. While the book does deal with economic issues, what has become known as the Rimland Theory deals primarily with military intervention, control and conquest of the Old World. The chapter entitled "The Geography of the Second World War" deals with this military situation on a global geographic basis. It is in this chapter that we find the foundations of the Rimland Theory of global domination. The theory was later expanded and refined in a series of lectures which were transcribed into the book "The Geography of the Peace".

The Cold War: The Geography of Containment

The conflict between the United States and the Soviet Union known as the Cold War has been examined in many different ways. The term Cold War was used to indicate that while it was a real conflict, the threat of nuclear weapons made direct armed confrontation too dangerous to contemplate. The exact beginning and ending dates of the conflict cannot be precisely defined, nor can questions of motive or blame be exactly determined. Most people believe that the conflict began soon after the end of the Second World War, although many trace the origins of the conflict to World War I. Estimates of a date for the end of the Cold War range from January of 1989 to December 25, 1991, when the Soviet Union was officially dissolved.

A study of the relevance of some of the traditional theories of political geography to the pursuit of the United States policy of containment must first consider the scope and origins of the containment policy, and the theories which may be applied to the study of containment. Specifically, the Mackinder Heartland theory, the Spykman Rimland theory, and the organic state theory are applicable to the study of containment.

Origins of Containment

The battles of the Cold War were fought on nearly every continent and in nearly every country around the globe. At the heart of the conflict were two very different world-views held by the two nations and their allies. The Soviet Union viewed capitalism as a monster, which, if unchecked, would consume the entire world with hedonistic abandon. America viewed Communism as an inherently evil mechanism designed to destroy the rights and liberties of all mankind. Both sides believed that the other was seeking world domination. In this way, each side held views that were essentially in line with Friedrich Ratzel's organic state theory, and, to an extent, with Rudolf Kjellen's extensions of the organic state theory.

Both states were acting to secure their own survival and security. To do so, they must create a system in which their own power was maximized, while that of their opponents must be reduced. It was necessary to "absorb politically valuable territory" by direct control or indirect influence in order to ensure the security and survival of the state. One analyst wrote, "Such a drive to achieve absolute security in a system in which no state could achieve that aim short of total domination left other states insecure and contributed to the volatility of the international system throughout the Cold War." Thus, in their efforts to pursue the Cold War, the superpowers were involved in a number of "hot" conflicts around the world.

These conflicts can be characterized by several alternating periods of aggression and reaction. The Soviets took the lead, with a grab for territory at the end of World War II. The success of the Soviets in achieving dominance in Eastern Europe, and the apparent desire of the Soviet Union to continue to expand it's sphere of influence into Greece, Turkey, and Iran caused great alarm in the United States. It was at this

time that the policy of containment was crafted, and the United States began to implement elements of the containment strategy.

The United States drafted its strategy for meeting the post-war Soviet threat in 1947. George Kennan, an American diplomat serving in Moscow, proposed a strategy in an anonymous article in the July 1947 edition of *Foreign Affairs* magazine. In his article, Mr. Kennan proposed, "a long-term patient but firm and vigilant containment of Russian expansive tendencies" through "counter-force at a series of constantly shifting geographical and political points." This soon became the impetus for the development of the containment policy by the Truman administration. President Truman identified two key components to the containment strategy: the formation of regional alliances and providing economic and military assistance to other nations to prevent communist expansion.

American administrations from Truman to George H.W. Bush pursued the strategy of containment in one form or another. The actions of the Soviet Union during the Cold War and the American response to those actions through the containment policy can be viewed in relation to the Heartland concept of Sir Alfred J. Mackinder, and the Rimland theory of Nicholas Spykman, as well as the organic state theory. The key question regarding the Cold War which faces the student of geography is:

Do the actions and outcomes of the Cold War validate or invalidate traditional theories of political geography?

To answer this question, it is necessary to view the events of the Cold War in the context of these theories.

Containment in Mackinder's Heartland and Spykman's Rimland

Sir Halford Mackinder first proposed his Heartland theory in 1904. In his view, the entire continent of Eurasia was viewed as one landmass, with Europe being one of its peninsulas. He also described the combined continents of Eurasia and Africa as the World-Island. Mackinder designated the north central area of Eurasia as the Heartland. He also referred to it as the "Pivot Area," since control of this area was essential to control of the entire continent. He predicted that Russia, the owner of the Heartland, with its massive land power and abundant natural resources would have the capacity to challenge British and Japanese sea power and obtain control of all of Eurasia. Outside of the central Eurasian "Pivot Area" lay the "Inner or Marginal Crescent," which, on Mackinder's Eurasia-centered map, spanned from the Baltics around the European and Asian coasts to eastern Siberia. Control of this area would ensure the dominance of the Heartland power over the entire world. He proclaimed: "Who rules East Europe commands the Heartland, Who rules the Heartland commands the World-Island, Who rules the World-Island commands the world."

As can be seen from this illustration, the Heartland includes virtually all of the Soviet Union. Mackinder's later revisions move the Heartland westward, to include

Eastern Europe, but including less of Siberia. The Siberian area, called Lenaland after the Lena River, was not considered to be part of the Heartland, but was considered as a resource base for the Heartland power.

At the end of World War II, it seemed as though Mackinder's prediction was coming true. One War Department analyst wrote: "'With Germany crushed, there is no power in Europe to oppose her [the Soviet Union's] tremendous military forces... The conclusions from the foregoing are obvious... without question she will dominate Europe on the defeat of the Axis.'" The Soviet Union had also made large territorial gains in Asia, obtaining portions of Manchuria, North Korea, and some Japanese islands.

Mackinder was countered in 1944 by Spykman's Rimland theory. Spykman contended that the Heartland power would be immobilized by difficulties with internal lines of communication and lack of mobility to expand beyond the physical barriers along its borders. Spykman felt that the domination of the Rimland (Mackinder's Marginal Crescent) would guarantee a naval and air power complete dominance over the earth. Spykman concluded, "Who controls the rimland rules Eurasia; who rules Eurasia controls the destinies of the world."

However, the emphasis was on the Rimland and it's position between the Heartland and the seas. The Rimland must be capable of contending both with the Heartland's land power, and the naval powers on its coasts. This unique advantage of the Rimland to compete with both forms of power is a significant advantage. In addition, the Rimland contains large material and population resources.

The starting point for our analysis will be the map of the post-war world devised by the three great powers at the Yalta Conference in July 1943. At Yalta, Churchill, Roosevelt, and Stalin foresaw the eventual Allied victory in World War II. It was necessary for them to define the roles they each would play in the defeat and occupation of the Axis nations. Key components of the Yalta Agreement were:

1. The British and French colonial empires would not be dismantled.
2. A United Nations organization was to be formed following the war.
3. The colonies of the Axis powers were to be governed under a protectorate system by the United Nations.
4. Under terms of the Secret Agreement between Roosevelt and Stalin, American and Soviet troops were to meet at the Elbe River, which would separate Germany into eastern and western occupation zones. In addition, the Soviets were to maintain primary influence over Eastern Europe as a buffer zone. However, Stalin assured Roosevelt that he would support self-government and free elections in the nations of Eastern Europe.
5. Three days following the defeat of Germany, the Soviet Union would enter the war on the eastern front against Japan. The Soviets would liberate Manchuria

and Korea north of the 38th parallel, while the U.S. would occupy Korea south of the 38th parallel. In addition, the Soviets would gain control of the Japanese Kuril Islands, effectively placing the entire Sea of Okhotsk under Soviet control.

The Soviet Union took the advantage early in the Cold War. In the closing days of World War II, the Soviets sought to occupy as much territory as they possibly could. Most of the new Soviet territory is within the Rimland. In Eastern Europe, the Communists quickly set up puppet governments and placed communists in virtually every position of power. The West now understood that Stalin's concept of "free and open government" was quite different from Roosevelt's. The Soviets incorporated territory gained in Asia into its republics. North Korea and Iran remained under Soviet Army occupation.

At the beginning of the Cold War, the Soviets had established their sphere of control over portions of the Rimland. As of yet, the Allies were unaware that a confrontation lay ahead. They still believed that the Soviets would honor the terms of the Yalta Agreements in withdrawing its occupation forces and assisting in the establishment of independent governments in formerly occupied territories.

The Soviets established the Rimland as the key battleground of the Cold War. Spykman argued that the key to maintaining post-war stability lay in preventing any power from obtaining control of the Rimland. Kennan concluded that the key to the success of containment was industrial might. There were five industrial centres in the world: the United States, Britain, West Germany, Japan, and the Soviet Union. The US and its allies owned four of the five, and the Soviets only one. To be successful, containment must limit the Soviets to that one. Two of the four industrial centres controlled by the US and its allies were in the Marginal Crescent, the others, Japan and the United States, were in the outer. In addition, key territory within the Rimland included the oilfields of the Middle East, and warm-water seaports in the Mediterranean, the Indian Ocean, and the Pacific.

The first contests of the Cold War were in Turkey, Greece, and Iran. The Soviet Union refused to remove its troops from Iran by March of 1946, the date agreed upon by the governments of Iran, Great Britain, and the Soviet Union. Instead, the Soviet Union sought to exploit the oil resources of Iran, as well as using its position there to gain influence and concessions from Turkey and Iran. Confronted with American and British pressure in the United Nations, the Soviets agreed to withdraw their troops by May 1946. This was the Allies first victory against the Soviet Union. Meanwhile, Britain was financially unable to continue its support for the anti-Communist governments of Turkey and Greece. Great Britain requested the assistance of the United States in the prevention of communist expansion into these countries.

In response, President Truman in March of 1947 made a speech to Congress requesting aid to the governments of Greece and Turkey. In this speech, he outlined the elements of the Truman Doctrine: the United States must act to "help free peoples

to maintain their free institutions and their national identity against aggressive movements that seek to impose upon them totalitarian regimes" and that these regimes "undermine the foundations of international peace and hence the security of the United States." The United States must therefore "assist free peoples to work out their own destinies in their own way." The administration then began to put Kennan's containment strategy into action.

As discussed earlier, the United States attempted to counter the Soviet threat through the formation of regional alliances and by providing economic and military assistance to nations vulnerable to Soviet penetration. Initially, these efforts were concentrated on the Rimland, and then later moved into the outer crescent.

Between 1947 and 1960, the United States formed mutual security agreements with forty-six nations. These included agreements with most of the Rimland nations, as well as many nations in the outer crescent. The only areas not covered by these pacts were Africa and most nations in the Middle East. Africa remained largely under control of the colonial powers of Western Europe and within the West's sphere of influence. And while most of the Middle Eastern nations were not included in regional security alliances, President Eisenhower issued the Eisenhower Doctrine, promising to oppose any communist encroachment in the Middle East.

These alliances surrounded the Soviet Union and its allies with a ring of nations pledged to come to each other's aid if any member became a victim of Soviet aggression. This effectively limited the paths of Soviet expansion.

The first test of containment came in Korea in June of 1950. The communist regime of Kim Il Sung, with the approval of Stalin and Mao Tse Dong, sought to forcefully reunite North and South Korea. The United States, though completely surprised and unprepared, reacted quickly to the communist aggression. Troops from the army of occupation in Japan were sent to assist the withering South Korean army. The United States was able to garner the support of the United Nations, and a multinational force with troops from several Asian nations was sent to Korea. When American forces seeking to liberate North Korea approached the Chinese border, China sent more than a million troops to stop the capitalist aggression. Eventually, a stalemate was reached in 1953, with the lines near the 38th parallel, where they began. While neither side gained any additional power or territory, the Korean conflict demonstrated the United States' resolve to back up its containment policy with force.

The Soviet Union, however, had found a way to bypass the containment ring. Since 1919, the Soviet Union had been operating Communist International, or Comintern, an organization dedicated to spreading communist revolution throughout the world. In 1943, in an effort to gain favour with the West, Stalin officially disbanded the Comintern, although it continued to function informally. During and following the war, communist parties in nations throughout the world continued to support

the implementation of communism through violent revolution. The communists were well received in many parts of the world, particularly in southeast Asia and Latin and South America. The weapons of ideological warfare were able to pass through the American blockade where troops and tanks could not. American military intervention became necessary to stop the communist incursions. The United States acted to remove governments with communist leanings in Iran in 1953, in Guatemala in 1954, and in Lebanon in 1958.

Moving into the 1960's, the Cold War opened fronts in both the Rimland and the outer crescent. The colonial empires of the European nations began to crumble. The European powers attempted to institute democratic governments in their former colonies, and to maintain their economic connections to their former colonies.

As governments were being restructured in the former colonies, the communists sought to expand their influence to these newly independent nations. The governments of these nations, now in a position of insecurity and weakness, sought help in stabilizing their frail economies and political structures. Many aligned themselves with either the Soviets or the Americans as a means of securing economic and military assistance for their nations. As a result, in the second half of the Cold War, the areas of conflict expanded rapidly throughout the outer crescent. Most of the nations on the superpowers bankroll maintained military dictatorships or other forms of authoritarian rule, whether rightist or leftist in ideology.

The most intense of the "hot" conflicts of the Cold War was the Vietnam War. This conflict began as a struggle for independence from French colonial rule. The French fought to maintain control of their colony from the end of World War II until their defeat at Dienbienphu in May of 1954. An agreement reached by the Five-Power Conference (France, The United States, Great Britain, China, and the Soviet Union) divided the country into a communist North above the 17th parallel, and a free South Vietnam, with elections and reunification to occur in July 1956. Recognizing that elections were likely to be won by Ho Chi Minh, the United States blocked the proposed elections. Eventually, conflict broke out between the North and the South. Millions of American troops were sent to Vietnam to support the South, while the Soviets and the Chinese sent troops and equipment to aid the North. Despite the superiority of the American fighters in training, technology, and equipment, they were unable to defeat the communist guerillas. In 1973, a peace agreement was reached and the American troops withdrew from Vietnam. Fighting continued, however, and two years later, the Northern forces overran the South, and Vietnam was reunited under communist rule.

Vietnam was the most significant Cold War loss for the West. The war was costly in men and money, and destroyed the American perception that the US could not be defeated in war. It caused deep rifts within the country, and caused many to question the American ideals used to justify the war. Disagreements regarding Vietnam

also strained relations within the alliance systems, especially in NATO, where France, Britain, and the US differed on Vietnam policy, and in SEATO.

From 1979 to 1989, the Soviet Union pursued a war to maintain a communist government it had installed in Afghanistan. The conflict became the USSR's Vietnam. The war was fought against mujihadeen rebels who were trained and supplied by the US Central Intelligence Agency. The Soviets also discovered how difficult it is for a modern mechanized army to fight a determined rebel force in hostile terrain. The war was a tremendous drain on the Soviet military and economy, and was very unpopular on the home front. The Soviet withdrawal in 1989 resulted in the collapse of the communist government, and plunged the country into a civil war that continued for twelve more years.

The Cold War ended with a rapid series of dramatic events. Soviet President Mikhail Gorbachev recognized the tremendous strains the conflict was having on his nation. The Soviet Union had been suffering for many years from a number of resource pressures resulting from its extremely inefficient economic system. Mikhail Gorbachev sought to reduce the demands on his nation's resources by withdrawing some of the Soviet Union's support for Cold War battles around the world. Significantly, the Soviet Union withdrew its forces from Afghanistan, and made agreements with the United States on nuclear arms reduction. Mikhail Gorbachev introduced reforms in the Soviet Union designed to relax state control of the economy and increase personal freedoms. He reduced Soviet control of the Eastern Bloc nations. These actions soon resulted in the fall of the Berlin Wall in December of 1989 and the collapse of the Soviet Union in December of 1991.

With the Cold War hostilities officially ended, the superpowers rapidly abandoned their Cold War outposts. Support for aligned governments in Africa, Latin America, and Asia ended. Almost overnight, the Cold War battlefields grew silent.

Analysis: Containment and the Heartland/Rimland Concepts.

How do the events of the Cold War correlate with the theories of Mackinder and Spykman?

In the first place, it seems that the conduct of the Cold War was not based on the principles espoused by either Mackinder or Spykman. There seems to be no indication that American planners considered the advice of either man during the Cold War. American actions during the Cold War were almost purely reactive. Whereas Kennan had argued for a policy of selective engagement, where the US would contend with the Soviets only for key strategic locations, the American policy came to be an effort to oppose communism wherever it appeared. American politicians and military analysts frequently portrayed the communist expansion as a centrally planned and controlled operation under the direction of Moscow. However, there is no evidence to support this. The Soviets desire was to spread revolution through

training of communist leaders by the Comminform. Rather than having a strategic plan for obtaining control of certain areas, communist revolutions seemed to occur haphazardly, as the opportunity presented itself, and as conditions were ripe. It is unclear what Soviet geopolitical thought was at the time, but it does seem as though the ideas of Mackinder were considered.

It seems that whether planned or not, both sides emphasized the value of Mackinders' Marginal Crescent and Spykman's Rimland. The following map indicates the areas of the most intense superpower conflict. As can be seen, most of the conflicts of the Cold War occurred in the Rimland / Marginal Crescent.

To what degree are Mackinder and Spykman's ideas supported or invalidated by the events of the Cold War?

The end of the Cold War resulted in the presence of a single superpower, the United States. This outcome was not predicted by either Mackinder or Spykman. However, it does correlate with Spykman's thinking. Spykman held that the nation that could maintain control of the Rimland could dominate the world political and economic order. However, neither side was able to obtain control of the Rimland. The intense conflict within the Rimland throughout the Cold War demonstrates the geopolitical and economic value of this territory. The unwillingness of the West to allow their opponent complete control over the Rimland is likely a significant factor in the eventual collapse of the Soviet Union. The United States victory in the Cold War occurred in much the same way that George Kennan had theorized: the Soviet system was incapable of self-sustainment, and by limiting the industrial and economic resources available to the Soviet Union, the system collapsed upon itself. By limiting the Soviets access to the resources available in the Rimland, the United States brought the fulfillment of Kennan and Spykman's theories.

Mackinder's theories are in some ways validated and invalidated by the outcome of the Cold War. Mackinder, in the height of World War II, wrote:

All things considered, the conclusion is unavoidable that if the Soviet Union emerges from this war as conqueror of Germany, she must rank as the greatest land Power on the globe. Moreover, she will be the Power in the strategically strongest defensive position. The Heartland is the greatest natural fortress on earth. For the first time in history it is manned by a garrison sufficient both in number and quality.

Mackinder also extols the value of the natural resources afforded to the Soviet Union. He states that the Soviet Union is capable of producing all of the food, industrial output, and power that it needs, and is basically self-sufficient within its borders. Mackinder also believed that the natural inclination of the Heartland power would be to establish a domain which included the Marginal Crescent. Based on Mackinder's determination, the Soviet Union should have had the power and resources to become the dominant world power.

For a number of reasons, Mackinder's concept largely did not come to pass. One key factor to consider is that Mackinder did not correlate the availability of resources with the ability to exploit those resources. The Soviet economic system was unable to capitalize on the resources available to it. The Soviet system was unable to create a robust economy. In fact, the Soviet Union eventually became dependent on imports of many commodities it should have been capable of producing within its own borders. Supporting expensive military operations around the globe became a great burden on the weak Soviet economic system.

Mackinder also failed to account for the rise of the United States as a world power. While he acknowledges that the United States had risen to the level of a great power following the First World War, he does not consider that the United States would eclipse the powers of Europe on the world stage. And once again, Mackinder considers the resources of the Soviet Union to equal the combined resources of the US and Canada, but does not account for the ability or inability to exploit those resources.

It could therefore be said that the United States military power, fuelled by its economic power, prevented Soviet domination of the Rimland, which prevented Soviet domination of the World-Island of Eurasia and Africa. It must be noted that military defeat of the Heartland power may still have been impossible. The West had no land power equal to that of the Soviet Union, and the Soviet Union was impenetrable with Western sea power. The only counter Europe had to the Soviet Union's massive military might in Europe was the threat of nuclear weapons. And, as Mackinder noted, even with airpower, the wide plains of Eastern Europe held the only plausible path for invasion of the Soviet Union. The United States and the Soviet Union had few options for direct military conflict outside of nuclear war.

An important note in his favour is Mackinder's advocacy of an Atlantic alliance. Mackinder had proposed an alliance of Great Britain and the nations of Western Europe with the United States as early as 1904. He considered this alliance of sea and land powers as an essential part of preventing the heartland power from expanding beyond its borders into Europe and Africa. The NATO alliance, formed after World War II, essentially performed this function as Mackinder had proposed.

Victory in the Cold War came as George Kennan predicted, by the collapse of the Soviet system due to its inferior economic power.

Analysis: The Cold War and Organic State Theory.

The organic state theory was originally proposed by the German geographer Friedrich Ratzel in 1896. According to his theory, which is based on the theories of Darwin, the State can be compared to an organism in nature. Glassner summarized Ratzel's idea in this way: The State is land, with man on the land, linked by the State idea and conforming to natural laws, with development tied to the natural environment.

Therefore, for example, States, like plants and people, do not do well in desert or polar regions. States need food in the form of *Lebensraum* (living space) and resources, and they constantly compete for them. States, like organisms, must grow or die. They live through stages of youth, maturity, and old age, with possible rejuvenation. The vitality of a State can generally be gauged by its size at a given time.

The Swedish geographer Rudolf Kjellin took Ratzel's ideas to an extreme. He declared that the State, in fact, *is* an organism. Kjellin maintained that it was absolutely necessary for States to grow and expand their territory, population, and resource base for survival. The larger States would consume the weaker ones, until only a few large States remained.

Karl Haushoffer, another German geographer, took the ideas of Ratzel and Kjellin even further. He convoluted their ideas to create justifications for the expansion of Nazi Germany. His teachings, known as *Geopolitik*, demonstrated the need for aggressive expansion of the German State as a requirement for obtaining the *Lebensraum* Germany needed to survive. It is largely the association of Geopolitik with Nazi aggression that discredited the study of geopolitics for many years.

Modern geographers generally do not believe that the State is literally an organism, and give little credence to the deterministic points of view put forward by Ratzel, Kjellin, and Haushoffer. It is not generally held that the expansion of the State is a requirement for survival. In fact, some geographers discredit the concept of the organic state theory altogether.

While the more extreme views put forward by Kjellin and Haushoffer cannot be used to create or justify the actions of States, certain aspects of the organic state theory can be applied to an understanding of the Cold War. Ratzel's original concept, especially, is useful in understanding the actions of the Soviet Union and the United States during the Cold War. The actions of both the Soviet Union and the United States during the Cold War clearly indicate that both parties sought to expand the territory under their control in an effort to expand their base of power while denying critical resources to the other.

The motivation for these actions was rooted in the concept of a bipolar world: a world consisting only of the Soviet Union and its allies, and the United States and its allies. This view, in a way, demonstrates Kjellin's theory of superstates. It could be viewed under the bipolar world concept that the Cold War globe was thus divided into two superstates. This view that every state must be aligned with one side or the other was so strongly held that the U.S. Secretary of State John Foster Dulles declared, "Neutralism is immoral."

Both states felt that the other was pursuing a program designed to completely eliminate their competitor from the map of the world. An American strategist wrote, "'the Soviets cannot in all prudence be expected to have ambitions for their imperial

security restricted to Eurasia. At stake ultimately in the contest is control…of the world.'" According to Glassner, the United States "assumed that the Soviets (and all communists and most socialists everywhere) were, and are, unqualifiedly evil, that they are bent on world domination, that they are fiendishly clever, and that any small victory by them would automatically result in many more." The Soviets felt that they must perpetuate the "'inevitable and irreconcilable struggle'" with the capitalist states and that it was necessary to break out of the "capitalist encirclement." It was therefore necessary to pursue the path of expansionism and interventionism to ensure their own survival.

Viewed in light of the organic state theory, the actions of the superpowers during the Cold War were as could have been expected. While neither side utilized the organic state theory for formulation or justification for their foreign policies, their actions were in accordance with its principles. The superpowers acted to expand their control of "politically valuable positions," to continually expand their spheres of influence, and to dominate smaller and weaker states in a struggle for survival. In the end, the United States proved to be the "fittest," and became the sole surviving superpower.

The Cold War was a long and intense struggle between two great nations, a struggle that reached nearly every corner of the globe. The events and outcomes of the Cold War have validated elements of three main classical geopolitical theories: the Mackinder Heartland concept, Spykman's Rimland doctrine, and the organic state theory. Mackinder's Heartland theory held merit, but failed to survive the realities brought about by the rapid pace of change in technology and balance of power in the world following World War II. The organic state theory proves to be useful as a generalization only of the behaviour of states in the pursuit of national survival. Of these theories, Spykman's Rimland doctrine most closely correlated with the policy of containment that the United States pursued successfully to the end of the Cold War.

Geostrategy

Geostrategy, a subfield of geopolitics, is a type of foreign policy guided principally by geographical factors as they inform, constrain, or affect political and military planning. As with all strategies, geostrategy is concerned with matching means to ends — in this case, a country's resources (whether they are limited or extensive) with its geopolitical objectives (which can be local, regional, or global). Strategy is as intertwined with geography as geography is with nationhood, or as Gray and Sloan state it, "[geography is] the mother of strategy."

Geostrategists, as distinct from geopoliticians, advocate proactive strategies, and approach geopolitics from a nationalist point-of-view. As with all political theories, geostrategies are relevant principally to the context in which they were devised: the

nationality of the strategist, the strength of his or her country's resources, the scope of his or her country's goals, the political geography of the time period, and the technological factors that affect military, political, economic, and cultural engagement. Geostrategy can function normatively, advocating foreign policy based on geographic factors, analytical, describing how foreign policy is shaped by geography, or predictive, predicting a country's future foreign policy decisions on the basis of geographic factors.

Many geostrategists are also geographers, specializing in subfields of geography, such as human geography, political geography, economic geography, cultural geography, military geography, and strategic geography. Geostrategy is most closely related to strategic geography.

Especially following World War II, some scholars divide geostrategy into two schools: the uniquely German organic state theory; and, the broader Anglo-American geostrategies.

Critics of geostrategy have asserted that it is a pseudoscientific gloss used by dominant nations to justify imperialist or hegemonic aspirations, or that it has been rendered irrelevant because of technological advances, or that its essentialist focus on geography leads geostrategists to incorrect conclusions about the conduct of foreign policy.

Defining Geostrategy

Academics, theorists, and practitioners of geopolitics have agreed upon no standard definition for "geostrategy." Most all definitions, however, emphasize the merger of strategic considerations with geopolitical factors. While geopolitics is ostensibly neutral, examining the geographic and political features of different regions, especially the impact of geography on politics, geostrategy involves comprehensive planning, assigning means for achieving national goals or securing assets of military or political significance.

Coining the Term

The term "geo-strategy" was first used by Frederick L. Schuman in his 1942 article "Let Us Learn Our Geopolitics." It was a translation of the German term "*Wehrgeopolitik*" as used by German geostrategist Karl Haushofer. Previous translations had been attempted, such as "defence-geopolitics." Robert Strausz-Hupé had coined and popularized "war geopolitics" as another alternate translation.

Modern Definitions

- "[G]eostrategy is about the exercise of power over particularly critical spaces on the Earth's surface; about crafting a political presence over the international system. It is aimed at enhancing one's security and prosperity; about making the international system more prosperous; about shaping rather than being

shaped. A geostrategy is about securing access to certain trade routes, strategic bottlenecks, rivers, islands and seas. It requires an extensive military presence, normally coterminous with the opening of overseas military stations and the building of warships capable of deep oceanic power projection. It also requires a network of alliances with other great powers who share one's aims or with smaller 'lynchpin states' that are located in the regions one deems important." —James Rogers and Luis Simón, "Think Again: European Geostrategy"

- "[T]he words *geopolitical, strategic,* and *geostrategic* are used to convey the following meanings: *geopolitical* reflects the combination of geographic and political factors determining the condition of a state or region, and emphasizing the impact of geography on politics; *strategic* refers to the comprehensive and planned application of measures to achieve a central goal or to vital assets of military significance; and geostrategic merges strategic consideration with geopolitical ones." —Zbigniew Brzezinski, *Game Plan* (emphasis in original)
- "For the United States, Eurasian geostrategy involves the purposeful management of geostrategically dynamic states and the careful handling of geopolitically catalytic states, in keeping with the twin interests of America in the short-term preservation of its unique global power and in the long-run transformation of it into increasingly institutionalized global cooperation. To put it in a terminology that hearkens back to the more brutal age of ancient empires, the three grand imperatives of imperial geostrategy are to prevent collusion and maintain security dependence among the vassals, to keep tributaries pliant and protected, and to keep the barbarians from coming together." —Zbigniew Brzezinski, *The Grand Chessboard*
- *Geostrategy* is the geographic direction of a state's foreign policy. More precisely, geostrategy describes where a state concentrates its efforts by projecting military power and directing diplomatic activity. The underlying assumption is that states have limited resources and are unable, even if they are willing, to conduct a *tous asimuths* foreign policy. Instead they must focus politically and militarily on specific areas of the world. Geostrategy describes this foreign-policy thrust of a state and does not deal with motivation or decision-making processes. The geostrategy of a state, therefore, is not necessarily motivated by geographic or geopolitical factors. A state may project power to a location because of ideological reasons, interest groups, or simply the whim of its leader. —Jakub J. Grygiel, *Great Powers and Geopolitical Change* (emphasis in original)
- "It is recognized that the term 'geo-strategy' is more often used, in current writing, in a global context, denoting the consideration of global land-sea distribution, distances, and accessibility among other geographical factors in strategic planning and action... Here the definition of geo-strategy is used in a more limited regional frame wherein the *sum of geographic factors interact to*

influence or to give advantage to one adversary, or intervene to modify strategic planning as well as political and military venture." —Lim Joo-Jock, *Geo-Strategy and the South China Sea Basin.* (emphasis in original)

- "A science named "geo-strategy" would be unimaginable in any other period of history but ours. It is the characteristic product of turbulent twentieth-century world politics."-Andrew Gyorgi, *The Geopolitics of War: Total War and Geostrategy* (1943)
- "'Geostrategy,'—a word of uncertain meaning—has... been avoided." —Stephen B. Jones, "The Power Inventory and National Strategy"

History of Geostrategy

Precursors

As early as Herodotus, observers saw strategy as heavily influenced by the geographic setting of the actors. In *History*, Herodotus describes a clash of civilizations between the Egyptians, Persians, Scythians, and Greeks—all of which he believed were heavily influenced by the physical geographic setting.

Dietrich Heinrich von Bulow proposed a geometrical science of strategy in the 1799 *The Spirit of the Modern System of War.* His system predicted that the larger states would swallow the smaller ones, resulting in eleven large states. Mackubin Thomas Owens notes the similarity between von Bulow's predictions and the map of Europe after the unification of Germany and of Italy.

Golden Age

Between 1890 and 1919 the world became a geostrategist's paradise, leading to the formulation of the classical geopolitical theories. The international system featured rising and falling great powers, many with global reach. There were no new frontiers for the great powers to explore or colonize—the entire world was divided between the empires and colonial powers. From this point forward, international politics would feature the struggles of state against state.

Two strains of geopolitical thought gained prominence: an Anglo-American school, and a German school. Alfred Thayer Mahan and Halford J. Mackinder outlined the American and British conceptions of geostrategy, respectively, in their works *The Problem of Asia* and "*The Geographical Pivot of History*". Friedrich Ratzel and Rudolf Kjellén developed an organic theory of the state which laid the foundation for Germany's unique school of geostrategy.

World War II

The most prominent German geopolitician was General Karl Haushofer. After World War II, during the Allied occupation of Germany, the United States investigated many officials and public figures to determine if they should face charges of war

crimes at the Nuremberg trials. Haushofer, an academic primarily, was interrogated by Father Edmund A. Walsh, a professor of geopolitics from the Georgetown School of Foreign Service, at the request of the U.S. authorities. Despite his involvement in crafting one of the justifications for Nazi aggression, Fr. Walsh determined that Haushofer ought not stand trial.

Cold War

After the Second World War, the term "geopolitics" fell into disrepute, because of its association with Nazi *geopolitik*. Virtually no books published between the end of World War II and the mid-1970s used the word "geopolitics" or "geostrategy" in their titles, and geopoliticians did not label themselves or their works as such. German theories prompted a number of critical examinations of *geopolitik* by American geopoliticians such as Robert Strausz-Hupé, Derwent Whittlesey, and Andrew Gyorgy.

As the Cold War began, N.J. Spykman and George F. Kennan laid down the foundations for the U.S. policy of containment, which would dominate Western geostrategic thought for the next forty years.

Alexander de Seversky would propose that airpower had fundamentally changed geostrategic considerations and thus proposed a "geopolitics of airpower." His ideas had some influence on the administration of President Dwight D. Eisenhower, but the ideas of Spykman and Kennan would exercise greater weight. Later during the Cold War, Colin Gray would decisively reject the idea that airpower changed geostrategic considerations, while Saul B. Cohen examined the idea of a "shatterbelt", which would eventually inform the domino theory.

Post-Cold War

Since the fall of the Berlin Wall, for most NATO or former Warsaw Pact countries, Geopolitical strategies have generally followed the course of either solidifying security obligations or accesses to global resources; however, the strategies of other countries have not been as palpable.

Notable Geostrategists

The below geostrategists were instrumental in founding and developing the major geostrategic doctrines in the discipline's history. While there have been many other geostrategists, these have been the most influential in shaping and developing the field as a whole.

Alfred Thayer Mahan

Alfred Thayer Mahan was an American Navy officer and president of the U.S. Naval War College. He is best known for his *Influence of Sea Power upon History* series of books, which argued that naval supremacy was the deciding factor in great power warfare. In 1900, Mahan's book *The Problem of Asia* was published. In this volume

he laid out the first geostrategy of the modern era. The *Problem of Asia* divides the continent of Asia into 3 zones:

- A northern zone, located above the 40th parallel north, characterized by its cold climate, and dominated by land power;
- The "Debatable and Debated" zone, located between the 40th and 30th parallels, characterized by a temperate climate;
- A southern zone, located below the 30th parallel north, characterized by its hot climate, and dominated by sea power.

The Debated and Debatable zone, Mahan observed, contained two peninsulas on either end (Asia Minor and Korea), the Isthmus of Suez, Palestine, Syria, Mesopotamia, two countries marked by their mountain ranges (Persia and Afghanistan), the Pamir Mountains, the Tibetan Himalayas, the Yangtze Valley, and Japan. Within this zone, Mahan asserted that there were no strong states capable of withstanding outside influence or capable even of maintaining stability within their own borders. So whereas the political situations to the north and south were relatively stable and determined, the middle remained "debatable and debated ground."

North of the 40th parallel, the vast expanse of Asia was dominated by the Russian Empire. Russia possessed a central position on the continent, and a wedge-shaped projection into Central Asia, bounded by the Caucasus mountains and Caspian Sea on one side and the mountains of Afghanistan and Western China on the other side. To prevent Russian expansionism and achievement of predominance on the Asian continent, Mahan believed pressure on Asia's flanks could be the only viable strategy pursued by sea powers.

South of the 30th parallel lay areas dominated by the sea powers—Britain, the United States, Germany, and Japan. To Mahan, the possession of India by Britain was of key strategic importance, as India was best suited for exerting balancing pressure against Russia in Central Asia. Britain's predominance in Egypt, China, Australia, and the Cape of Good Hope was also considered important.

The strategy of sea powers, according to Mahan, ought to be to deny Russia the benefits of commerce that come from sea commerce. He noted that both the Dardanelles and Baltic straits could be closed by a hostile power, thereby denying Russia access to the sea. Further, this disadvantageous position would reinforce Russia's proclivity toward expansionism in order to obtain wealth or warm water ports. Natural geographic targets for Russian expansionism in search of access to the sea would therefore be the Chinese seaboard, the Persian Gulf, and Asia Minor.

In this contest between land power and sea power, Russia would find itself allied with France (a natural sea power, but in this case necessarily acting as a land power), arrayed against Germany, Britain, Japan, and the United States as sea powers. Further, Mahan conceived of a unified, modern state composed of Turkey, Syria, and

Mesopotamia, possessing an efficiently organized army and navy to stand as a counterweight to Russian expansion.

Further dividing the map by geographic features, Mahan stated that the two most influential lines of division would be the Suez and Panama canals. As most developed nations and resources lay above the North-South division, politics and commerce north of the two canals would be of much greater importance than those occurring south of the canals. As such, the great progress of historical development would not flow from north to south, but from east to west, in this case leading toward Asia as the locus of advance.

Halford J. Mackinder

Halford J. Mackinder His major work, Democratic ideals and reality: a study in the politics of reconstruction, appeared in 1919. It presented his theory of the Heartland and made a case for fully taking into account geopolitical factors at the Paris Peace conference and contrasted (geographical) reality with Woodrow Wilson's idealism. The book's most famous quote was: "Who rules East Europe commands the Heartland; Who rules the Heartland commands the World Island; Who rules the World Island commands the World." This message was composed to convince the world statesmen at the Paris Peace conference of the crucial importance of Eastern Europe as the strategic route to the Heartland was interpreted as requiring a strip of buffer state to separate Germany and Russia. These were created by the peace negotiators but proved to be ineffective bulwarks in 1939 (although this may be seen as a failure of other, later statesmen during the interbellum). The principal concern of his work was to warn of the possibility of another major war (a warning also given by economist John Maynard Keynes).

Mackinder was anti-Bolshevik, and as British High Commissioner in Southern Russia in late 1919 and early 1920, he stressed the need for Britain to continue her support to the White Russian forces, which he attempted to unite. Significance of Mackinder.

Mackinder's work paved the way for the establishment of geography as a distinct discipline in the United Kingdom. His role in fostering the teaching of geography is probably greater than that of any other single British geographer.

Whilst Oxford did not appoint a professor of Geography until 1934, both the University of Liverpool and University of Wales, Aberystwyth established professorial chairs in Geography in 1917. Mackinder himself became a full professor in Geography in the University of London (London School of Economics) in 1923.

Mackinder is often credited with introducing two new terms into the English language : "manpower", "heartland". Influence on Nazi strategy.

The Heartland Theory was enthusiastically taken up by the German school of Geopolitik, in particular by its main proponent Karl Haushofer. Whilst Geopolitik

was later embraced by the German Nazi regime in the 1930s, Mackinder was always extremely critical of the German exploitation of his ideas. The German interpretation of the Heartland Theory is referred to explicitly (without mentioning the connection to Mackinder) in The Nazis Strike, the second of Frank Capra's Why We Fight series of American World War II propaganda films. Influence on American strategy.

The Heartland theory and more generally classical geopolitics and geostrategy were extremely influential in the making of US strategic policy during the period of the Cold War.

Influence on Later Academics

Evidence of Mackinder's Heartland Theory can be found in the works of geopolitician Dimitri Kitsikis, particularly in his geopolitical model "Intermediate Region".

Friedrich Ratzel

Influenced by the works of Alfred Thayer Mahan, as well as the German geographers Karl Ritter and Alexander von Humboldt, Friedrich Ratzel would lay the foundations for *geopolitik*, Germany's unique strain of geopolitics.

Ratzel wrote on the natural division between land powers and sea powers, agreeing with Mahan that sea power was self-sustaining, as the profit from trade would support the development of a merchant marine. However, his key contribution were the development of the concepts of *raum* and the organic theory of the state. He theorized that states were organic and growing, and that borders were only temporary, representing pauses in their natural movement. *Raum* was the land, spiritually connected to a nation (in this case, the German peoples), from which the people could draw sustenance, find adjacent inferior nations which would support them, and which would be fertilized by their *kultur* (culture). Ratzel's ideas would influence the works of his student Rudolf Kjellén, as well as those of General Karl Haushofer.

Rudolf Kjellén

Rudolf Kjellén was a Swedish political scientist and student of Friedrich Ratzel. He first coined the term "geopolitics." His writings would play a decisive role in influencing General Karl Haushofer's *geopolitik*, and indirectly the future Nazi foreign policy. His writings focused on five central concepts that would underlie German *geopolitik*:

1. *Reich* was a territorial concept that was composed of *Raum* (*Lebensraum*), and strategic military shape;
2. *Volk* was a racial conception of the state;
3. *Haushalt* was a call for autarky based on land, formulated in reaction to the vicissitudes of international markets;

4. *Gesellschaft* was the social aspect of a nation's organization and cultural appeal, Kjellén anthropomorphizing inter-state relations more than Ratzel had;
5. *Regierung* was the form of government whose bureaucracy and army would contribute to the people's pacification and coordination.

General Karl Haushofer

Karl Haushofer's geopolitik expanded upon that of Ratzel and Kjellén. While the latter two conceived of geopolitik as the state-as-an-organism-in-space put to the service of a leader, Haushofer's Munich school specifically studied geography as it related to war and designs for empire. The behavioral rules of previous geopoliticians were thus turned into dynamic normative doctrines for action on lebensraum and world power.

Haushofer defined geopolitik in 1935 as "the duty to safeguard the right to the soil, to the land in the widest sense, not only the land within the frontiers of the Reich, but the right to the more extensive Volk and cultural lands." Culture itself was seen as the most conducive element to dynamic expansion. Culture provided a guide as to the best areas for expansion, and could make expansion safe, whereas solely military or commercial power could not.

To Haushofer, the existence of a state depended on living space, the pursuit of which must serve as the basis for all policies. Germany had a high population density, whereas the old colonial powers had a much lower density: a virtual mandate for German expansion into resource-rich areas. A buffer zone of territories or insignificant states on one's borders would serve to protect Germany. Closely linked to this need was Haushofer's assertion that the existence of small states was evidence of political regression and disorder in the international system. The small states surrounding Germany ought to be brought into the vital German order. These states were seen as being too small to maintain practical autonomy (even if they maintained large colonial possessions) and would be better served by protection and organization within Germany. In Europe, he saw Belgium, the Netherlands, Portugal, Denmark, Switzerland, Greece and the "mutilated alliance" of Austro-Hungary as supporting his assertion.

Haushofer and the Munich school of geopolitik would eventually expand their conception of lebensraum and autarky well past a restoration of the German borders of 1914 and "a place in the sun." They set as goals a New European Order, then a New Afro-European Order, and eventually to a Eurasian Order. This concept became known as a pan-region, taken from the American Monroe Doctrine, and the idea of national and continental self-sufficiency. This was a forward-looking refashioning of the drive for colonies, something that geopoliticians did not see as an economic necessity, but more as a matter of prestige, and of putting pressure on older colonial powers. The fundamental motivating force was not be economic, but cultural and

spiritual. Beyond being an economic concept, pan-regions were a strategic concept as well. Haushofer acknowledged the strategic concept of the Heartland put forward by the Halford Mackinder. If Germany could control Eastern Europe and subsequently Russian territory, it could control a strategic area to which hostile sea power could be denied. Allying with Italy and Japan would further augment German strategic control of Eurasia, with those states becoming the naval arms protecting Germany's insular position.

Nicholas J. Spykman

Nicholas J. Spykman was an Dutch-American geostrategist, known as the "godfather of containment." His geostrategic work, *The Geography of the Peace* (1944), argued that the balance of power in Eurasia directly affected United States security.

N.J. Spykman based his geostrategic ideas on those of Sir Halford Mackinder's Heartland theory. Spykman's key contribution was to alter the strategic valuation of the Heartland vs. the "Rimland" (a geographic area analogous to Mackinder's "Inner or Marginal Crescent"). Spykman does not see the heartland as a region which will be unified by powerful transport or communication infrastructure in the near future. As such, it won't be in a position to compete with the United States' sea power, despite its uniquely defensive position. The rimland possessed all of the key resources and populations—its domination was key to the control of Eurasia. His strategy was for Offshore powers, and perhaps Russia as well, to resist the consolidation of control over the rimland by any one power. Balanced power would lead to peace.

George F. Kennan

George F. Kennan, U.S. ambassador to the Soviet Union, laid out the seminal Cold War geostrategy in his *Long Telegram* and *The Sources of Soviet Conduct*. He coined the term "containment", which would become the guiding idea for U.S. grand strategy over the next forty years, although the term would come to mean something significantly different from Kennan's original formulation.

Kennan advocated what was called "strongpoint containment." In his view, the United States and its allies needed to protect the productive industrial areas of the world from Soviet domination. He noted that of the five centres of industrial strength in the world—the United States, Britain, Japan, Germany, and Russia—the only contested area was that of Germany. Kennan was concerned about maintaining the balance of power between the U.S. and the USSR, and in his view, only these few industrialized areas mattered.

Here Kennan differed from Paul Nitze, whose seminal Cold War document, NSC-68, called for "undifferentiated or global containment," along with a massive military buildup. Kennan saw the Soviet Union as an ideological and political challenger rather than a true military threat. There was no reason to fight the Soviets throughout Eurasia, because those regions were not productive, and the Soviet Union was already

exhausted from World War II, limiting its ability to project power abroad. Therefore, Kennan disapproved of U.S. involvement in Vietnam, and later spoke out critically against Reagan's military buildup.

Henry Kissinger

Henry Kissinger implemented two geostrategic objectives when in office: the deliberate move to shift the polarity of the international system from bipolar to tripolar; and, the designation of regional stabilizing states in connection with the Nixon Doctrine. The regional stabilizers were pro-American states which would receive significant U.S. aid in exchange for assuming responsibility for regional stability. Among the regional stabilizers designated by Kissinger were Zaire, Iran, and Indonesia.

Zbigniew Brzezinski

Zbigniew Brzezinski laid out his most significant contribution to post-Cold War geostrategy in his 1997 book *The Grand Chessboard*. He defined four regions of Eurasia, and in which ways the United States ought to design its policy toward each region in order to maintain its global primacy. The four regions (echoing Mackinder and Spykman) are:

- Europe, the Democratic Bridgehead
- Russia, the Black Hole
- The Middle East, the Eurasian Balkans
- Asia, the Far Eastern Anchor

In his subsequent book, *The Choice*, Brzezinski updates his geostrategy in light of globalization, 9/11 and the intervening six years between the two books.

Criticism of Geostrategy

"Few modern ideologies are as whimsically all-encompassing, as romantically obscure, as intellectually sloppy, and as likely to start a third world war as the theory of 'geopolitics.'" —Charles Clover, "Dreams of the Eurasian Heartland".

Geostrategy encounters a wide variety of criticisms. It has been called a crude form of geographic determinism. It is seen as a gloss used to justify international aggression and expansionism—it is linked to Nazi war plans, and to a perceived U.S. creation of Cold War divisions through its containment strategy. Marxists and critical theorists believe geostrategy is simply a justification for American imperialism.

Some political scientists argue that as the importance of non-state actors rises, the importance of geopolitics concomitantly falls. Similarly, those who see the rise of economic issues in priority over security issues argue that geoeconomics is more relevant to the modern era than geostrategy. Most international relations theory that is critical of realism in international relations is likewise critical of geostrategy because

of the assumptions it makes about the hierarchy of the international system based on power.

Further, the relevance of geography to international politics is questioned because advances in technology alter the importance of geographical features, and in some cases make those features irrelevant. Thus some geographic factors do not have the permanent importance that some geostrategists ascribe to them.

Border

Borders define geographic boundaries of political entities or legal jurisdictions, such as governments, sovereign states, federated states and other subnational entities. Some borders—such as a state's internal administrative borders, or inter-state borders within the Schengen Area—are open and completely unguarded. Other borders are partially or fully controlled, and may be crossed legally only at designated border checkpoints. Some, mostly contentious, borders may even foster the setting up of buffer zones.

Definitions

In the past, many borders were not clearly defined lines, but were neutral zones called marchlands. This has been reflected in recent times with the neutral zones that were set up along part of Saudi Arabia's borders with Kuwait and Iraq (however, these zones no longer exist). In modern times the concept of a marchland has been replaced by that of the clearly defined and demarcated border. For the purposes of border control, airports and seaports are also classed as borders. Most countries have some form of border control to restrict or limit the movement of people, animals, plants, and goods into or out of the country. Under international law, each country is generally permitted to define the conditions that have to be met by a person to legally cross its borders by its own laws, and to prevent persons from crossing its border when this happens in violation of those laws.

Some legal orders require presentation of passports and visas, or other identity documents to cross borders. To stay or work within a country's borders aliens (foreign persons) may need special immigration documents or permits that authorise them to do so. Having such documents (i.e., visa and passport) however does not automatically guarantee that the alien will be allowed to cross to the other side of the border.

Moving goods across a border often requires the payment of excise tax, often collected by customs officials. Animals (and occasionally humans) moving across borders may need to go into quarantine to prevent the spread of exotic or infectious diseases. Most countries prohibit carrying illegal drugs or endangered animals across their borders. Moving goods, animals or people illegally across a border, without declaring them, seeking permission, or deliberately evading official inspection

constitutes smuggling. In regions where smuggling, migration, and infiltration are a problem, many countries fortify borders with separation barriers and institute formal border control procedures. Some borders are only signposted. This is common in countries within the European Schengen Area and on rural sections of the Canada – United States border. Borders may even be completely unmarked, a common occurrence with remote or forested borders.

Hostile countries that are not at war may be separated by a militarized border. The most well-known of these is the former Berlin Wall. Furthermore, many hostile, militarized borders are separated by a buffer zone or demilitarized zone, such as the Korean Demilitarized Zone and the United Nations Buffer Zone in Cyprus, and may be separated by a buffer state. The most extreme borders are completely closed with no passage, such as the Blue Line that separates Israel and Lebanon.

Natural Borders

Natural borders are geographical features that present natural obstacles to communication and transport. Existing political borders are often a formalization of these historical, natural obstacles.

Some geographical features that often constitute natural borders are:

- Oceans: oceans create very costly natural borders. Very few nation states span more than one continent. Only very large and resource-rich states are able to sustain the costs of governance across oceans for longer periods of time.
- Rivers: some political borders have been formalized along natural borders formed by rivers. Some examples are; the Rio Grande border (Mexico-USA), the Rhine border (France-Germany), and the Mekong border (Thailand-Laos).
- Lakes: larger lakes create natural borders. One example is the natural border created by Lake Tanganyika (Congo-Burundi-Tanzania-Zambia).
- Forests: denser jungles or forests can create strong natural borders. One example of a natural forest border is the Amazon rain forest (Colombia-Venezuela-Guyana-Brazil-Bolivia-Peru).
- Mountain ranges: research on borders suggests that mountains have especially strong effects as natural borders. Many nations in Europe and Asia have had their political borders defined along mountain ranges.

Throughout history, technological advances have reduced the costs of transport and communication across these natural borders. This has reduced the significance of natural borders over time. As a result, political borders that have been formalized more recently — such as those in Africa or Americas — typically conform less to natural borders than very old borders — such as those in Europe or Asia — do. States whose borders conform to natural borders are, for similar reasons, more likely to be strong nation-states.

Maritime Borders

A maritime border is a division enclosing an area in the ocean where a nation has exclusive rights over the mineral and biological resources, encompassing maritime features, limits and zones. Maritime borders represent the jurisdictional borders of a maritime nation and are recognized by the United Nations Convention on the Law of the Sea.

Maritime borders exist in the context of territorial waters, contiguous zones, and exclusive economic zones; however, the terminology does not encompass lake or river boundaries, which are considered within the context of land boundaries.

Some maritime borders have remained indeterminate despite efforts to clarify them. This is explained by an array of factors, some of which illustrate regional problems.

Border Economics

The presence of borders often fosters certain economic features or anomalies. Wherever two jurisdictions come into contact, special economic opportunities arise for border trade. Smuggling provides a classic case; contrariwise, a border region may flourish on the provision of excise or of import–export services — legal or quasi-legal, corrupt or corruption-free. Different regulations on either side of a border may encourage services to position themselves at or near that border: thus the provision of pornography, of prostitution, of alcohol and/or of narcotics may cluster around borders, city limits, county lines, ports and airports. In a more planned and official context, Special Economic Zones (SEZs) often tend to cluster near borders or ports.

Even if the goods are not perceived to be undesirable, states will still seek to document and regulate the cross-border trade in order to collect tariffs and benefit from foreign currency exchange revenues. Thus, there is the concept unofficial trade in goods otherwise legal; for example, the cross-border trade in livestock by pastoralists in the Horn of Africa. Ethiopia sells an estimated $250 to $300 million of livestock to Somalia, Kenya and Djibouti every year unofficially, over 100 times the official estimate.

Human economic traffic across borders (apart from kidnapping), may involve mass commuting between workplaces and residential settlements. The removal of internal barriers to commerce, as in France after the French Revolution or in Europe since the 1940s, de-emphasises border-based economic activity and fosters free trade. Euroregions are similar official structures built around commuting across borders.

Politics

Political borders have a variety of meanings for those whom they affect. Many borders in the world have checkpoints where border control agents inspect those crossing the boundary.

In much of Europe, such controls were abolished by the Schengen Agreement and subsequent European Union legislation. Since the Treaty of Amsterdam, the competence to pass laws on crossing internal and external borders within the European Union and the associated Schengen States (Iceland, Norway, Switzerland, and Liechtenstein) lies exclusively within the jurisdiction of the European Union, except where states have used a specific right to opt-out (United Kingdom and Ireland, which maintain a common travel area amongst themselves).

The United States has notably increased measures taken in border control on the Canada–United States border and the United States–Mexico border during its War on Terrorism. One American writer has said that the 3,600 km (2,000 mile) US-Mexico border is probably "the world's longest boundary between a First World and Third World country."

Historic borders such as the Great Wall of China, the Maginot Line, and Hadrian's Wall have played a great many roles and been marked in different ways. While the stone walls, the Great Wall of China and the Roman Hadrian's Wall in Britain had military functions, the entirety of the Roman borders were very porous, which encouraged Roman economic activity with neighbours. On the other hand, a border like the Maginot Line was entirely military and was meant to prevent any access in what was to be World War II to France by its neighbor, Germany. Germany ended up going around the Maginot Line through Belgium just as it had done in World War I.

Cross-border Regions

Macro-regional integration initiatives, such as the European Union and NAFTA, have spurred the establishment of cross-border regions. These are initiatives driven by local or regional authorities, aimed at dealing with local border-transcending problems such as transport and environmental degradation. Many cross-border regions are also active in encouraging intercultural communication and dialogue as well as cross-border economic development strategies.

In Europe, the European Union provides financial support to cross-border regions via its Interreg programme. The Council of Europe has issued the Outline Convention on Transfrontier Co-operation, providing a legal framework for cross-border co-operation even though it is in practice rarely used by Euroregions.

Border Studies

There has been a renaissance in the study of borders during the past two decades, partially from creation of a counter narrative to notions of a borderless world that have been advanced as part of globalization theory. Examples of recent initiatives are the Border Regions in Transition network of scholars, the International Boundaries Research Unit at the University of Durham, the Association of Borderlands Studies based in North America, and the founding of smaller border research centres at

Nijmegen and Queen's University Belfast. Contemporary leading scholars in the field of border studies include Emmanuel Brunet Jailly at the University of Victoria, who is the Executive Secretary and Treasurer of the Association for Borderlands Studies, (Emmanuel Brunet Jailly, and Henk van Houtum and Martin van der Velde at Radboud University are the editors of the international scholarly *Journal of Borderlands Studies*), David Newman at Ben Gurion University (co-editor of the international journal *Geopolitics*). Other leading scholars include Paul Ganster at San Diego State University's Institute for the Regional Study of the Californias, Akihiro Iwashita at Hokkaido University, Oscar Martínez at the University of Arizona, Liam O'Dowd at Queen's University Belfast, Anssi Paasi at the University of Oulu, Anthony Payan at the University of Texas El Paso (Payan is the past President of the Association for Borderland Studies), James Scott at Karelian Institute, Joensuu University, David Shirk at the University of San Diego's Trans-Border Institute, Rick Van Schoik at Arizona State University's North American Centre for Transborder Studies, and Doris Wastl-Walter at the University of Bern.

Political Geography: A Basic Guide

Political Geography is the study of the spatial patterns of conflict and cooperation among political actors at all scales. The struggle of states for territory and resources has always been a key theme. Likewise, the impact of national identities in forming states, and the use of landscape and territorial symbols, or iconography provides insights into how geographical phenomena are deployed in struggles for power over earth space, whether terrestrial, oceanic or atmospheric.

State: defined as a bounded political unit with territory, population, and organized government possessing power and sovereignty (Use the mnemonic device "T-POPS"). This last term denotes supreme authority within the territory of the state, and recognition by other states that this authority is legitimate. The state is the primary unit of analysis in what is often referred to as 'international relations' or IR. In political geography, we study how and where states seek to consolidate or expand control over territory, people and natural resources This is done through diplomacy, and various forms of coercion, including trade embargoes, blockades and other military measures up to and including war.

The state is very old, with many historians and political scientists considering the various Greek, Roman and Chinese models as ancient examples, more than 2000 years ago. Some political geographers, such as Peter Taylor, have called these early states 'world empires,' because they dominated the worlds known to their subjects. In its modern form, the state evolved in Europe from about 1500, typically in core areas (eg Paris basin) dominated by a focused and determined monarchy. By about 1900, the state system had been spread through colonialism throughout the world, particularly in Latin America, Africa and Asia, regions that had been host to a wide

variety of political arrangements. State System: The modern state system is usually dated to the Peace of Westphalia in 1648, which ended a bloody European religious conflict called the Thirty Years War. From this point forward, the conflicts in Europe were no longer religiously based, but developed through clashes between monarchical governments over balance of power issues and overseas territories. After the French Revolution (1789), anti-monarchical republican forms of government added another element in the long struggle between Great Powers such as Britain and France. Gradually, over the next century and a half, the Ottoman, Russian and Hapsburg empires gave way to various forms of states. This reflected the social and economic changes due to the industrial revolution, which in turn led to the rise of socialism (discussed in theory section below) and nationalism, which is the ideology that privileges national identity over all other forms of human loyalty. Scholars say that nationalism as an ideology demands that individuals put their state or country first, even to the point of the (in)famous saying, 'My country right or wrong.'

Nation: defined by sociologist Benedict Anderson as an "imagined community," a nation is a group of people with a shared heritage who believe that they belong together, and who almost always develop political aspirations for special recognition, and perhaps a territorially-based state. These identities generally develop over long periods of historical time, typically centuries, though elites play a role in creating a sense of nationalism, which can develop within a generation. The Basque region of Spain and the Kurdish populations of Southwest Asia have often been identified as non-state nations, ie people without their own sovereign territorial unit, some of whom aspire to create such an entity. Many can be identified today, and these phenomena are a major factor in conflicts within states that refuse to recognize their claims.

Identities (ethnic and national): this concept relates to the deepest aspects of how humans think about and identify themselves, and refers to how individuals and other actors in society self-ascribe. Think of your own situation. Are you a Euro-, African-,Latino-or Asian American? Do you prefer to think of yourself as simply 'American'? If you are of Arab descent, do you call yourself 'Arab'? Or perhaps Jordanian or Yemeni? If you are from Taiwan, are you Chinese or Taiwanese? If born in Germany, German or European? If you are a religious person, perhaps your identity as a Christian (Protestant, Catholic or Orthodox), Hindu, Muslim, Buddhist or Jew is of the greatest importance. These decisions by people (whether as individuals, households or groups) have an impact on politics within states, and on transnational issues relating to conflict and cooperation.

Territory: Every state controls it own piece of terrestrial space, as well as the atmosphere above it, and (if it has a coastline), the waters out to twelve nautical miles, and the resources of its continental shelf. If an ethno-national group aspires to statehood, its political leaders will typically advance a territorial claim, and proceed to struggle

toward the realization of a new state. This idea of territoriality ties in with scale. Geographers Guntram Herb and David Kaplan refer to the manner in which scale and territory intermesh in the form of 'nested identities.' This means individuals and groups prioritizes a certain scale in self-ascription. The particular situation of the individual or group will impact the way s/he or it self-identifies. While nobody truly has total freedom in these decisions, different persons and groups will identify themselves by prioritizing a certain scale.

Territory is also a part of iconography. This is the use of key landscape symbols in the creation of a national identity. Think of poetry and songs that use the imagery of the physical geography of regions to inspire patriotic fervor ('America the Beautiful' is an excellent example).

Territorial Morphology: The shapes of states can have an impact on the ability of ruling governments to impose law and policy on state territory. While this is not a 'hard and fast' rule, states with territorial outliers are sometimes vulnerable to separatist political tendencies. We can use a term from physics to think about the state's ability to rule its territory effectively. Centripetal forces are those institutions, customs and symbols designed to keep a state together. Centrifugal forces, such as regional inequalities and ethnic division, are those conditions or forces that tend to tear a state apart. Compact states are those with a minimum of variation in distance between the centre and peripheral boundaries of the state. An atlas will reveal several good examples: Uruguay in South America, Zimbabwe in Southern Africa, Poland in Europe and Cambodia in Southeast Asia Generally, these states could be easier to rule, but the principle does not always hold if ethnic conflict and corrupt governance are in evidence. Zimbabwe, for example, is very unstable. Nevertheless, these states tend not to experience separatist movements.

Elongated states simply have a national territory that is long and narrow. Chile and Vietnam are examples. If the state has only one economic and political core region, as in Chile, then the state institutions will usually be able to impose power over distant regions more easily. If however, two core regions exist, as is the case with Vietnam (in the North around Hanoi, and in the South around Saigon or Ho Chi Minh city), then the state could very well experience political turbulence and division, as actually has occurred in Vietnamese history. Fragmented states are divided into multiple pieces of territory, separated by water bodies or other states. Island states such as Japan and Indonesia are good examples. Here, two states experience very different political outcomes. Japan has become a wealthy core economic state, with a largely homogeneous ethnic population (98-99% Japanese). It has little difficulty in ruling its territory. Indonesia is another matter entirely. This vast archipelago sweeps across the Indian Ocean to the Melanesian portion of the Pacific World, with over 13,000 islands! Inside its territory, many islands include ethnic groups that are hostile to each other, for religious and other reasons. Indonesia recently lost the territory of

East Timor, whose population had been fighting for independence since the 1970's. It became the world's newest state in 2002. Indonesia is a relatively poor peripheral country in the global economy, and this no doubt contributes to its political difficulties. Prorupt states can also experience political turbulence. Thailand and Burma (Myanmar) provide examples. These states have a relatively compact core region, with a long territorial extension. In both cases, the regions of extension contain ethnic groups that are not well integrated into the state.

In Thailand, a overwhelmingly Buddhist society, the provinces in the southern tip of the country, near Malaysia, are inhabited by Muslim peoples who feel little attachment to the state of which they are a part. This region has recently (as of 2004) experienced civil unrest. If ethnically distinct and economically poor people live in these prorupt regions, or any peripheral region in a state, then conditions are favorable for the development of separatism. Perforated states are a rarity; two examples will suffice. The state of Lesotho, landlocked and surrounded by South Africa, forms a perforation in that larger country, and is at the mercy of South Africa for its economic well-being, and political survival. In a previous era, up to 1989, the former German Democratic Republic (or East Germany) was perforated by the city formerly known as West Berlin. That city is now united as the capital of the unified Federal Republic of Germany.

Boundaries and Frontiers: The boundaries separating states are known as international boundaries, and they establish the territorial limits of the legal authority of states. These boundaries can be physical or cultural. From the standpoint of political geography, all state territories are separated from each other by legally recognized boundaries, though many of these are contested, sometimes leading to military conflict between states.

Physical or Natural boundaries typically entail the use of rivers and mountains. At first glance, these might seem to be useful due to their apparent clarity, visibility and distinctiveness on the earth's landscapes. But look more closely and we can see that mountains can be divided peak to peak or by ridgelines, which separate the headwaters of rivers. These two features do not always coincide. An example of a border dispute involving mountainous boundaries is that between India and China over two separate subregions in their Himalayan boundary: In the far western edge of their international border, Aksai Chin is between Indian Kashmir and the Chinese province of Xinjiang. To the east, we find another region between India's far northeastern region of Arunachal Pradesh, and the eastern edge of Tibet, which is ruled by China. These are not minor matters. China and India fought a brief war over these boundary disputes in 1962. Geopolitically, the two states are still antagonists. As for rivers, some of those used for boundaries can change course! An example is the Rio Grande, which forms the boundary between Mexico and the United States. This boundary had to be renegotiated between the two states, which have a history

of antagonism and conflict dating back to the wars of Texas independence (1836) and 1846-48. Fortunately, this history of conflict has changed drastically to one of greater cooperation in recent decades.

Artificial boundaries are delimited or imposed through the recognition of historical custom, by treaty or by the will of a stronger power. Antecedent boundaries existed on the cultural landscape prior to the emergence of the formal state system. Subsequent boundaries develop with the ethno-cultural divisions of a regional landscape. These are then adjusted through conflict and negotiation, changing with the relative strengths of the parties involved. Superimposed boundaries were typically drawn by colonial states. The best examples of these types of boundaries are the geometric lines drawn by European powers in Africa during the 19th century. The political consequences of these geographies would prove to be very difficult indeed for newly independent African states in the 20th century. Finally, relict boundaries are no longer politically recognized, but reflect previous political conditions. Examples include the 17th parallel division of Vietnam into northern and southern states, made relict by the collapse of the southern state in 1975. More recently, Germany's Cold War boundary between eastern and western entities ceased to exist with the collapse of the Communist East German state in 1989-90. These various types of boundaries are also discussed with maps in Map 11 of your Atlas.

Offshore or Maritime boundaries refer to international boundaries over water. The laws and customs of centuries have been encapsulated in the 1982 draft treaty of the United Nations Convention on the Law of the Sea (UNCLOS). The convention delimits territorial boundaries and rights over resources, using a series of four zones, with each successive zone representing diminishing control by a state. A territorial sea ranges up to 12 nautical miles (19 km) or 'nm' (a nautical mile is 1.15 statute miles), and within this zone states have sovereign rights, such as exclusive claim to live marine resources. Vessels from other countries have the right of innocent passage. A contiguous zone exists to 24 nm (38 km), in which coastal states can enforce customs, immigration and environmental laws, and enjoy the right of 'hot pursuit' of hostile aircraft and vessels. An exclusive economic zone exists up to 200 nm (370 km). Within this EEZ, the state has exclusive rights to explore and exploit all types of natural and marine resources in both waters and on and beneath the seabed. If the continental shelf continues beyond the coast, countries can exploit this up to 350 nm (560 km). Vessels have all rights of innocent passage. The high seas are beyond the EEZ's of states. They are open to all states, and no state has the right to interfere with others sailing, fishing, flying over or engaging in scientific research. Mineral resources are to be managed for the common benefit of humanity. These provisions of the Law of the Sea became a formal part of international law in 1994.

Frontiers are politically weak regions or places between two or more states or political units that are often ecologically marginal, but may be ascribed strategic

significance. Boundaries are weakly developed, poorly delimited or perhaps nonexistent. Tibet and Afghanistan provide historical examples. During the nineteenth century these arid, highland regions were sites of conflict between the British and Russian Empires. In what was called "The Great Game," both empires vied for influence with local rulers and peoples as the British sought to prevent Russian expansion toward India, which was the 'jewel in the Crown' of the British Empire. Today, Tibet is ruled by China, another powerful state, and Afghanistan is now dominated by the United States, and is struggling to rebuild a sense of nationhood from disparate regions after more than twenty years of civil war and foreign interference.

Geopolitics can be defined as the application of geographical knowledge and insights to problems of power and earth-space. Typically, a geopolitical approach to statecraft entails a strategic approach to a state's foreign policy that applies this geographical knowledge to maximize the power of a state relative to other states in world politics and the global economy. Even cultural influence, what is currently called 'soft power,' can be applied geopolitically. Remember that in interstate relations, power means the ability to coerce or convince other states to behave in a way which accords with one's own national interests. This also means the ability to set the agenda in intergovernmental organizations such as the United Nations, and to prevent other items or issues from being dealt with or discussed if these are not in accord with the state's interests.

Some of the most important thinkers in the nineteenth and early twentieth century highlight this kind of classical geopolitics. Friedrich Ratzel (1844-1904), influenced by the social Darwinism of the late nineteenth century, believed that states *behave like* organisms, and that they must grow in power at the expense of other states, or they will decline and eventually be eclipsed in power by other states. As a German university professor, he wrote in the context of Germany's rise as an industrial and military power before the First World War (1914-1918).

The Swedish academic Rudolf Kjellen (1864-1922) took this argument one step further, insisting that the state *is* an organism, which must expand or decline. His ideas influenced geopolitics in Nazi Germany (1933-1945), with Hitler's belief that Germany must expand eastward into the Heartland of Eurasia to seek resources, slave labour and greater *lebensraum* (living space) for the German people. This concept of a Eurasian World Island and Eastern European Heartland had come from the British politician and geographer Sir Halford Mackinder (1861-1947). Writing as the British Empire was at its height, he believed that land powers such as Russia would soon eclipse maritime powers such as Britain due to the superior resources and population base of the Eurasian Heartland. Thus the saying, "Who rules East Europe commands the Heartland: Who rules the Heartland commands the World-Island: Who rules the World-Island commands the World." Geopolitical thinking at the scale of the state

continues to influence the most powerful states in the global interstate system, although the specific ideas change with shifts in technology and society.

An excellent case study of geopolitics in action is provided by the Cold War, which began about 1947 and ended with the decline of the Soviet Union from 1989-1991. This struggle between the post-Second World War superpower, the United States and the Soviet Union, was characterized by the division of world politics into three spheres or 'worlds': a capitalist First World, led by the United States, a communist Second World led by the Soviet Union, and a multifaceted Third World of recently decolonized states in Africa, Asia and Latin America. The geographical form of the struggle was the division of Europe between a communist East and capitalist West, and the efforts of the United States and the Soviet Union to gain allies in the Third World. This entailed the use of these countries as proxies, in which the superpowers could indirectly fight each other for power in the global system. For example, Cuba was an important ally of the Soviet Union after the rise to power of Fidel Castro in 1959. For a brief period, China was also an ally of the Soviets, though these two powers eventually split by 1960. Many other states in the Third World were forced to choose sides in this conflict, which often had devastating results for the world's poorest countries, in which civil wars were intensified by the sale of weapons to rival armies by the US and Soviet Union. Sometimes the two powers intervened directly. The United States attempted to keep Vietnam divided into a non-communist South and communist North, but ultimately failed because many Vietnamese believed the American forces were foreign invaders rather than protectors. In an eerily similar conflict, the Soviet Union sent forces to Afghanistan in December 1979, to prop up a communist government that had taken power in that country. The United States and its Persian Gulf ally Saudi Arabia perceived this to be the prelude to a possible move on the oil-rich Gulf region and decided to support the Muslim *mujahadeen* rebels fighting to establish an Islamic state in the country. Among these rebels was a young guerilla fighter and financier of Saudi-Yemeni background named Osama bin Laden. While the Soviets eventually withdrew (in 1988) after suffering heavy casualties, Afghanistan became a failed state, with the fundamentalist Taliban regime eventually taking control of most of the country. They in turn provided the bases and logistical support for al-Qaeda.

The strategies and policies of this period were identified with certain practitioners of geopolitics. In the United States, the most famous were Henry Kissinger, Secretary of State during the Nixon (1969-1974) and Ford (1974-1976) Administrations and Zbigniew Brzezinski, National Security Adviser during the Carter Administration (1977-1981). Kissinger and Brzezinski were both convinced of the need to maintain the strategic 'credibility' of the US by preventing the Soviet Union from gaining influence in regions deemed strategically vital. This was accomplished through the use of covert military actions, spying, and destabilizing governments that were perceived to be too close to the Soviet Union, or too critical of the US. Both men

ultimately pursued a vision of geopolitics that would have been very familiar to Mackinder.

We can see from these examples that local geographical and socio-economic conditions were often set aside by the superpowers in the name of ideological competition. The resulting proxy wars of this period would prove to have destructive consequences for all concerned, particularly the Third World countries that were victimized by civil conflict, repression and superpower interference. Both superpowers engaged in these activities, though one can debate the relative merits of their respective socio-economic and political systems.

Globalization refers to the economic, political and cultural processes through which the certain states, regions and cities in the world are being tied together into a global system. These processes have accelerated due to the rapid advancements in information technology or IT during the last two decades or so. For example, computer technology allows for twenty-four hour financial trading in global cities of the Core economic regions. However, it is important to note that the spread of capitalism and its opportunities and instabilities is not uniform over space: most of the benefits accrue to transnational corporations and their shareholders in Core countries. Some of these TNC's, as they are known, are wealthier and more powerful than states in peripheral regions. In fact, the economic relations of core and periphery established during colonialism have not substantially changed, though some countries have managed to develop economically out of the periphery. Examples include the Four Asian Tiger economies: Korea (South), Taiwan, Hong Kong and Singapore.

Another key aspect of globalization is the phenomenon of 'coming together and coming apart.' Through investment and the spread of different cultural ideas, goods and peoples into many regions of the earth, some peoples and groups have reacted very negatively to the new conditions of multiculturalism and change. Thus, we have the phenomenon described by Canadian political scientist Benjamin Barber in his book, 'Jihad vs. McWorld', that various forms of fundamentalism are in struggle with the materialist capitalism of the global economy. The geographer Jean Gottman, writing in the 1950's, remarkably anticipated a key aspect of these processes through his analysis of how regions and peoples struggle to preserve their identities through iconography in the face of increasing instability in their lives due to capital flows, migrations and other shifts in economy, society and politics. This idea of iconography versus flows is playing out in various ways in different parts of the world. An example is the reassertion of religious identities such as Christian among political groups and region in the US, Islamic in various societies of North Africa and Southwest Asia, and Hindu within the context of Indian politics. To make matters more complex, transnational communities maintain a sense of identity through maintenance of social institutions in countries or regions where they are a minority. For example, the Chinese and Indian communities in the US and many other societies are noted for

a strong sense of ethnic and religious identity in many different geographical settings. Transnational issues include environmental degradation, narcotrafficking, terrorism, migration, and the spread of disease. The increasing emphasis on these issues by state elites is indicative of the stress that globalization has exerted on states, and the difficulties that states have in dealing with these issues. The flow of processed coca into the United States, terror attacks and complex issues related to the environment reflect these dilemmas.

Finally, these issues of geopolitics, and identity highlight the importance of studying peoples and places within a local context. Seeing the world through simplistic eyes leads us to make easy and false generalizations about certain categories of persons, and this can lead to tragic errors in politics and policy. Globalization(s) have led some geographers to a less state-based and more postmodern approach to studying the world. By this we mean taking regions and peoples on their own terms, acknowledging complexity and diversity, and seeking to encourage policies that acknowledge these realities.

Sovereignty, Territory & International Law

"No legal principle – not even sovereignty – should ever be allowed to shield genocide, crimes against humanity and mass human suffering."

The incentive of this paper is not to elucidate state sovereignty as an enemy, or hindrance, to international criminal law, nor vice versa, but to show that it is possible to construct a conceptual framework where the both exist in synchronization.

State sovereignty is not a static concept. This paper examines the notion of state sovereignty, first providing an outlook on the historical developments and the relationship between state sovereignty and international criminal law, then discussing the purpose of sovereignty and how the international community can be conceived as protecting sovereignty and its raison d'être in the light of international criminal justice.

The scope of absolute state sovereignty, authority over all matters as presented by Hobbes, cannot effectively correspond to existing global problems. The subsistence of the people living in the state is necessary for its existence and therefore the sovereign powers of the state should be limited so that the powers cannot be used to infringe the fundamental interests of the people. The protection of human rights can be derived from this rationale because for maintaining sovereignty the state must preserve the subjects that comprise it or are the very reason for its existence. If sovereignty of the state is understood to exist for the people, then gross human rights violations endanger the purpose of the state sovereignty – the interests of the people.

Yet in a situation where a state is unable or unwilling to carry out its duties for the people, the responsibility falls on the international community. Article 17 of the

Statute of the International Criminal Court determines that the state holds primary jurisdiction and only that when the state is "unwilling or unable genuinely to carry out the investigation or prosecution" the situation becomes admissible to the International Criminal Court. This arrangement is generally referred to as the complementarity principle. In order to efficiently guarantee the protection of the fundamental rights of the people, to prevent and punish violations, there must be enforcement mechanisms in place, thus the need for rules criminalising conduct detrimental to those rights, and eventually, to the notion of state sovereignty. However, those rules as set out in international criminal law do not arise nor exist in a vacuum, but necessitate continuous normative discourse between different regimes and paradigms of nation states and international community.

Sovereignty and International Law

Despite occasional claims for a fade-out of the Westphalian concept of State sovereignty, the international community does in fact continue to depend on it. The Marxist doctrine once predicted the fate of the concept, but developing countries, while adopting Marxist teachings in their criticism of the traditional international legal institutions, have tended to reinforce their sense of sovereignty in their dealings with the established international order.

1) International law has developed through increased co-operation among sovereign States in recent years as, for example, in the European Union, but it allows the State to assert sovereignty in a variety of ways: persistent objection to the formation of a customary rule of international law; nuclear threat in a world of general prohibition of the use of force; and above all, the unchanged concept of territorial sovereignty.
2) The very notion of the State has these essential components:
   (a) a permanent population,
   (b) a defined territory,
   (c) government,
   (d) capacity to enter into relations with other States" (Montevideo Convention on Rights and Duties of States of 1933, Article 1).

   This is evidenced in the actual state of the international community: sovereign States generally refrain from interfering in the domestic affairs of the others. Whatever political regime and social institutions a State may have is a matter for it within its own territorial limits.
3) In so far as international boundaries exist as a matter of fact, they may be disputed between the States concerned, as history abundantly shows. In the foreseeable future, therefore, boundary disputes, especially maritime boundary disputes, could inevitably arise or emerge if and when natural resources are involved in the boundary areas.

Towards a Framework for the 'Fairer' International Law on Territorial Disputes and Historical Critisicm Approach.

The notable decisions and awards by international judicial and arbitral bodies for formulating the relevant international law relating to territorial acquisition and loss are Islands of Palmas Arbitration, Clipperton Island Arbitration, Legal Status of Eastern Greenland Case, Minquiers and Ecrehos Case, Western Sahara, Case concerning the Land, Island and Maritime Frontier Dispute (El Salvador v. Honduras Case), Eritrea-Yemen Arbitration, Case concerning Kasikili / Sedudu Island, Case concerning Maritime Delimitation and Territorial Questions between Qatar and Bahrain, Land and Maritime Boundary between Cameroon and Nigeria, Sovereignty over Pulau Ligitan and Pulau Sipadan, Frontier Dispute (Benin / Niger), Sovereignty over Pedra Branca / Pulau Batu Puteh, Middle Rocks and South Ledge, and Territorial and Maritime Dispute between Nicaragua and Honduras in the Caribbean Sea.

This fact is significant given that the general understanding of territorial disputes in international law has been developed through decisions and awards by international judicial and arbitral bodies and these cases basically involve either colonising countries or colonised / newly independent countries. Consequently, and having regard to the fact that none of these cases parallels the situation in respect of the cases involve between colonising countries and colonised / newly independent countries. In other words, cases involve between former colonies and former imperial powers did not count as a major factor in formulating the relevant international law relating to territorial acquisition and loss. Thus, it is debatable whether the general rule of international law on territorial disputes can be applied, without modification, to the specific cases which cannot be clearly fitting into the terminologies and categories.

In the circumstances, it is imperative to clarify the relevant norms of international law and develop new norms to address the very nature of the territories in dispute. This will be the case for formulating a broad framework for the appreciation of the nature of territorial disputes.