Spatial data Models

"A **model** is a manageable comprehensible and schematic representation of a piece of reality."

Modelling in GIS- Map is the most familiar type of model, it is the simplification of reality. Data base is also a map. Analog model consist of map, whereas digital model consist of google earth.

Purpose of Modelling - To understand and solve problems. Modelling should be done keeping in mind the problem, accuracy, spatial domain - area, country, continents and temporal domain - time, years or point. User, who is going to use, normal people, GI user or organisation.

Function of modelling - to understand a special phenomenon, to represent a phenomenon, to measure or calculate a distance or to set up plans for disaster response unit. Can be used to predict future behaviour of a system.

MODELING IN GIS

The most familiar model is the *map*

Databases are also models

A collection of stored data representing real-world phenomena

vs

Analogue vs *digital* models

Digital models are more flexible and therefore more adaptable. They allow animations and simulations

Static vs *Dynamic* models

Static models represent a single state of affairs at a point in time (most maps & DB)

Dynamic models address developments or changes in the real world – i.e. process models

Different *computer representations* of the real world (e.g. GIS layers)

MODELLING

Geographic phenomena in the real world

Types of models

Static model - It represents single affair at the time. Traffic cannot be updated.

Dynamic model - it is a live version of the model, it changes with the change in the scenario.

Geographic phenomena - It's an event that takes place in time and space(place/area), they can be natural, artificial or both.

Properties - it can be named or described, geo referenced or can be assigned a time interval at which it took place.

Properties:

- Can be *named or described*
- Can be *geo-referenced*
- Can be *assigned a time interval* at which it is/was/will be present

Types- There are two types objects and field.

Object - It is the collect quantity, it does not cover the entire area/site. It can be studied in a collection as well as in a unit, such as a singular house and a cluster of the house.They can have a specific characteristics. Different type of object should not oc cupy the same location. Parameters to describe an object is location, shape, size, orientation.

Fields - They have a value in the entire study area, they have a specific value in a study area and can be derived by a mathematical equation.

Continues field - the value in such field changes smoothly such as temperature.

Discreate field - there is drastic change in the field such as the crop field or the building density.

TYPES: OBJECTS AND FIELDS

A (geographic) *object* is a geographic phenomenon that does not cover the total study area, the space in between objects is potentially empty or undetermined

- It occurs in *specific localities*

A (geographic) *field* is a geographic phenomenon in which every point can be assigned a value

- It manifests throughout the *whole study area*

Objects can be studied as a *collection* that we consider as a unit

These units may have specific characteristics

Different objects do not occupy the same location: *mutual non-overlap*

Always interesting to see objects at different levels of aggregation

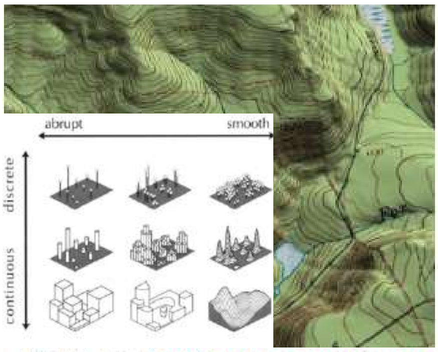
Sparsely populates the study area

Easily distinguished and named

Parameters to describe them:

- *Location*
- *Shape*
- *Size*
- *Orientation*

FIELDS



Fields can be *discrete* or *continuous*

- Have a value "*everywhere*" in the study area
- Can be defined as a mathematical function ***f*** that associates a specific value with any position in the study area
- If ***(x,y)*** is a location in the study a area, then *f(x,y)* expresses the value of ***f*** at the location ***(x,y)***

In a *continuous field*, the underlying function is assumed to be "mathematically smooth", meaning that the field *values* along any path through the study area *do not change abruptly*, but only gradually

In a *discrete field* on the other hand, *values do change abruptly*. Discrete fields are intermediate between continuous fields and geographic objects

DISCRETE FIELDS

- *Values change abruptly* from field to field
- Study space is divided into *mutually exclusive, bounded parts*, with all locations in one part having the same field value
- Are *intermediate* between continuous fields and geographic objects

Challenges -

Geographical scale - It is important to put scale, the problem will determine at which scale should model be made

Boundaries - boundaries consists of geographic objects but also discrete field, turning reality into a crisp boundary is a challenge.

Continuity - a continuous field like temperature is infinite, reality is infinite and computer if finite, so the challenge is to collect or turn infinite into finite. It applies to continues field and to lines. Data is collected to do this, such as Nominal (qualitative data) it consist of land type forest water etc, ordinal (hierarchical data) to differentiate between rods but we can't do mathematics on that, Interval data(quantitative data) it is numerical data.

Spatial interpolation - It allows for interpolation, it is based on principal called spatial autocorrelation.

CHALLENGES: GEOGRAPHIC SCALE

- Geographic phenomena can be studied at different levels of aggregation, at different *scales*
- The scale at which the phenomenon is studied determines the most appropriate way to represent it

Buildings

Building blocks

Neighborhood

Crisp boundaries can be determined at an almost arbitrary level of precision

More common in *man-made* phenomena

Fuzzy boundaries are not precise lines, but rather a transition area

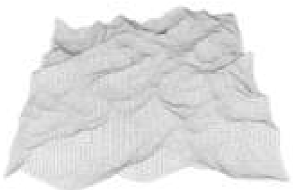
More common in *natural* phenomena

Poorly supported in GIS environment

CHALLENGES: CONTINUITY/INFINITUDE

Many geographic phenomena have intrinsic ***continuous and / or infinite characteristics*** that have to be represented with ***finite means*** (computer memory)

Continuous fields



Line objects

RESOURCES: DATA TYPES AND VALUES

Different types of values that we can use to represent geographic phenomena

Four different data values:

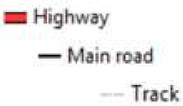
Nominal → **Qualitative data**

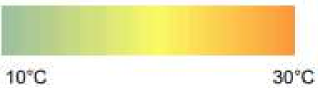
Ordinal → **Hierarchical data**

Interval
Ratio → **Quantitative data**

- *Nominal:* values that provide a name or identifier. Also called categorical.

- *Ordinal:* values that can be put in a natural sequence but do not allow any other type of computation





- *Interval :* numerical data that allow simple forms of computation like addition and subtraction. Interval data has no arithmetic zero value and therefore ratios are not possible to calculate

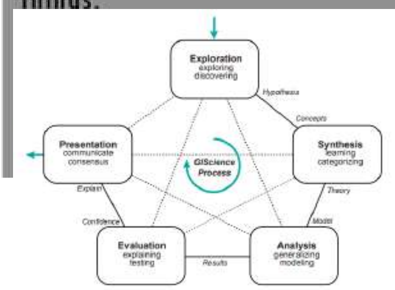
- *Ratio data values,* allow most, if not all, forms of arithmetic computation and have a natural zero value (e.g; distance, population)

Interpolation is based on a principle called ***spatial autocorrelation***

1. Tobler's first law

Phenomena with intrinsic continuous and/or infinite characteristics therefore have to be represented with finite means (computer memory) for computer manipulation, yet any finite representation scheme is open to errors and interpretation. To allow for this, fields are usually implemented with a tessellation approach, and objects with a (topological) vector approach. In this learning path we will first discuss tessellations (regular and irregular) and
vector-based representations and how these are applied to represent geographic fields and objects.

Everything is related to everything else, but near things are more related than distant things.

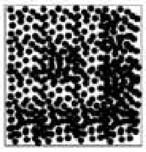


RESOURCES: INTERPOLATION

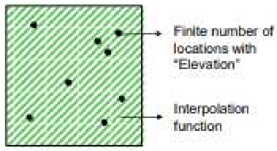
CONTINUOUS FIELDS - INTERPOLATION

Various geographic phenomena have the characteristics of continuous functions in space.

Options to represent these phenomena:



Option 1: store as many points as possible



Option 2: find a symbolic representation as a function

In GIS we combine both approaches: finite number of points + interpolation function

Data Models - two models to digitally represent geographic phenomena.

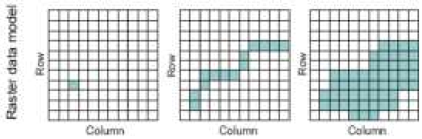
1. **Raster model** - It is the representation of the world as the surface is divided into regular grids and space, pixels are used. Better resolution gives better details.
   Tessellations - It is referred to cells that make up the complete study area like pixels. For each cell a value is assigned to characterise that part in space. They are in same size and same shape, easy to work with, they are faster than vector but they do not adopt to the phenomenon.
   Regular tessellation - Cells are or same size and shape, the fields value is assigned and is the same throughout. Square cells are commonly used.
   Irregular tessellations - the space is partitioned into mutually elusive cells that vary in size and shape they adapt to the spatial phenomena that they represent.

2. **Vector model** - reality is shown using points, lines and polygons. You can add points to the line to make it detail.
   Georeference - It is different than the raster, in vector each point has a location and that's why the algorithm is heavier here.

   Euclidean space - it is a model of space, location is given by using coordinates. In case of 2D we use (x,y) in case of 3D (xyz) is used. We can measure distances area, define topological space, there is **interior** and **boundary**

RESOURCES: SPATIAL DATA MODELS

Two models to digitally represent geographic phenomena:

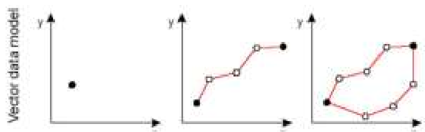


The ***raster*** model

A representation of the world as a surface divided into a regular grid of cells

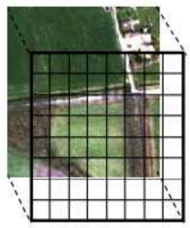
The ***vector*** model

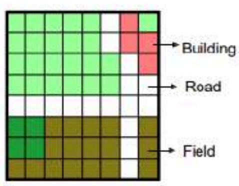
A representation of the world using points, lines, and polygons



THE RASTER MODEL

REGULAR TESSELLATIONS





A ***tessellation*** is a partitioning of space into mutually exclusive cells that together make up the complete study area

For each cell, a value is assigned to characterize that part of the space

In ***regular tessellations***, cells are of the same shape and size, and the field attribute value assigned to a cell is associated with the entire area occupied by the cell.

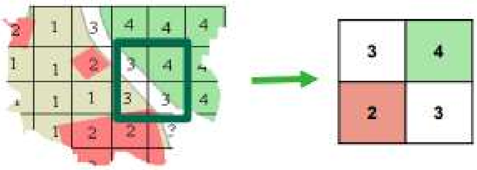
Square cell tessellation is commonly used, mainly because georeferencing of such a cell is straightforward

Topology - It deals with spatial relations that do not change under specific transformation, the permeant properties that remain the same overtime.
Interior and **boundary** is used to define topological space.

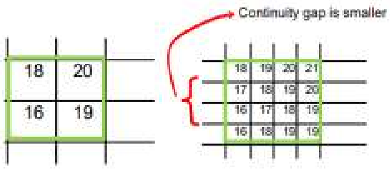
THE RASTER MODEL

The value for a cell is assumed to be valid for all locations within the cell.

Originates ***continuity gaps***



Make the cell size smaller, so as to make the 'continuity gaps' between the cells smaller



THE RASTER MODEL

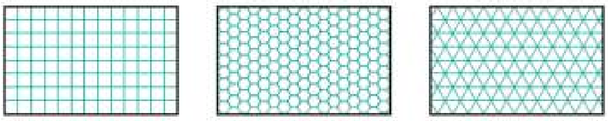
REGULAR TESSELLATIONS

Advantages:

- Structure simplicity
- We know how the space is partitioned. This leads to faster calculations and algorithms.

Disadvantages:

- They do not adapt to the spatial phenomenon we want to represent.
- Cell boundaries are both artificial and fixed: they may or may not coincide with the boundaries of the phenomenon of interest.

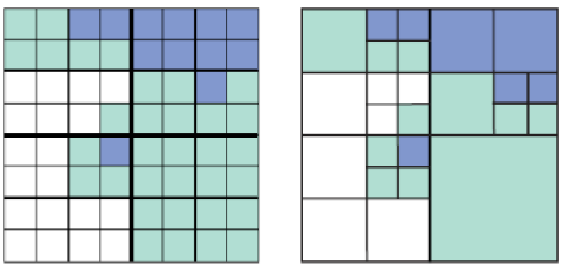


IRREGULAR TESSELLATIONS

In ***irregular tessellations*** the space is partitioned into mutually exclusive cells, but now the cells vary in size and shape → Cells *adapt* to the spatial phenomena they represent.

Region Quadtree → It splits up the area into four quadrants. This procedure stops when all the cells in a quadrant have the same field value.

When a conglomerate of cells has the same value, they are represented together provided their boundaries coincide with the predefined quadrant boundaries



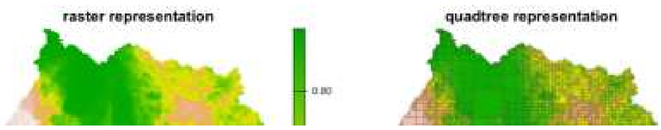
IRREGULAR TESSELLATIONS

Advantage:

More adaptive which typically leads to a reduction in the amount of computer memory needed to store the data

Disadvantage:

More complex. The partition is not regular anymore

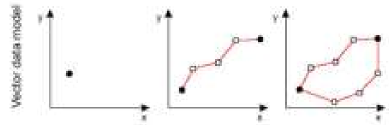


Temporal dimension -

Discrete time is composed of discrete time elements, in continues time no such discrete elements exists for every two moments in between.

THE VECTOR MODEL

A representation of the world using points, lines and polygons



A ***georeference*** is a coordinate pair from some geographic space, also known as a ***vector***

Rasters do not explicitly store georreferences of the phenomena. They provide a coordinate pair of the lower left corner and the resolution of the raster.

In ***vector representations*** a *georeference* is ***explicitly associated*** with the geographic phenomena.

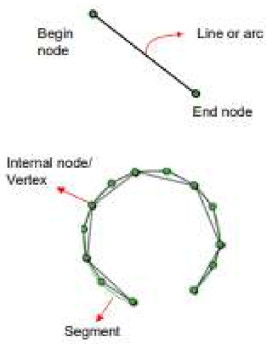
THE VECTOR MODEL

POINTS

Points are defined as ***single coordinate pairs (x,y)*** when we work in 2D or ***coordinate triplets (x,y,z)*** when we work in 3D

Points are used to represent objects that are shapeless, size-less and zero-dimensional at a certain scale and/or for certain purposes

LINES

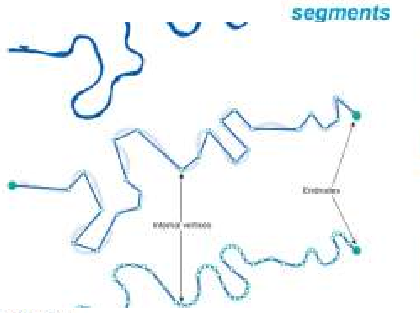


Used to represent one-dimensional objects (roads, railroads, canals, rivers…)

Line is defined by 2 ***end nodes*** and 0 to *n* internal nodes. An ***internal node or vertex*** is like a point that only serves to define the line

GISs store lines as a sequence of coordinates of its end nodes and vertices, assuming that all its line segments are straight.

Straight parts between 2 vertices are called ***segments***

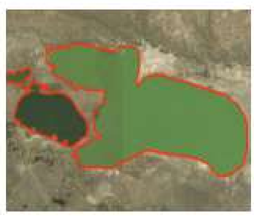


By increasing the number of internal vertices, we can improve the shape

The number of vertices determines the precision.

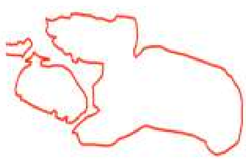
Scale is related to the spatial accuracy: coarser scale = lower number of internal vertices = higher degree of generalization.

POLYGONS



Used to represent ***areal phenomena***.

The usual technique is to apply a ***boundary model***: each areal feature is represented by some arc/node structure that determines a polygon as the area's boundary.



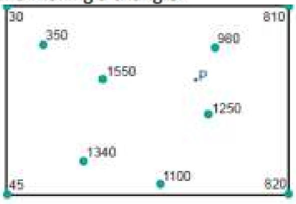
Area features of the same type are stored in a single data layer, represented by mutually non-overlapping polygons

Issues:
- finite approximation to an infinite reality
- fuzzy boundaries

TRIANGULATED IRREGULAR NETWORKS (TIN)

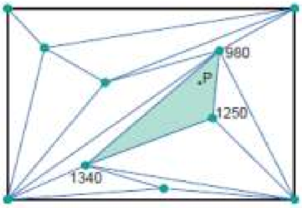
Triangulated irregular networks (TIN) - it is neither raster not vector, it is different type of approach, it is usually used to represent surface. Commonly used for continuous field, in 3d space 3 points are used. It is used by connecting 3 dots and making a triangle.

Commonly-used ***data model for continuous fields*** built from a set of locations for which we have a measurement (*x, y, z*)



In three-dimensional space, ***three points determine a unique plane***, as long as they are not positioned on the same line.

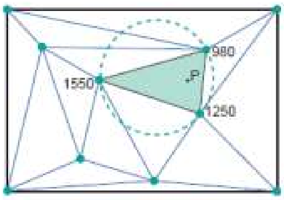
If we restrict the use of a plane to the area between its three anchor points, we obtain a ***triangular tessellation*** of the complete study area



Many tessellations are possible. Some better than others = smaller errors

A better approximation is achieved when the average distance from P to the three triangle anchors is smaller.

That is a ***Delaunay triangulation***, which in a sense is an optimal triangulation.

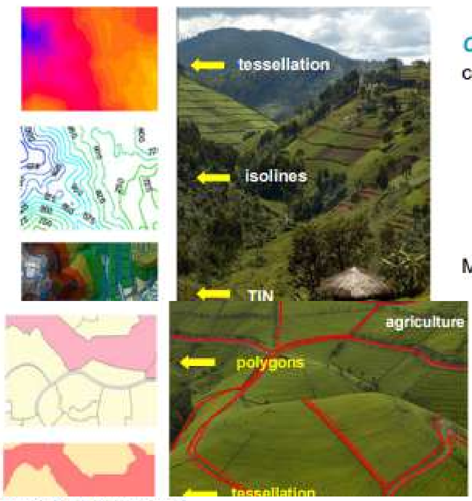


Delaunay triangulation

Properties:

1. the triangles are as equilateral as they can be
2. the circumcircle through the three anchor points does not contain any other anchor point

REPRESENTATIONS OF GEOGRAPHIC FIELDS



Continuous fields (*like elevation*) can be represented as:

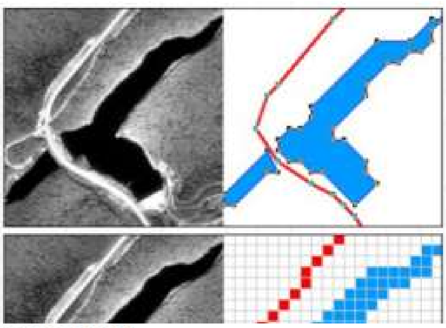
- Tessellation
- Isolines: linear feature that connects points with equal field values
- TIN

More common: tessellations

Discrete fields (like *landuse*) can be represented as:

- Polygons
- Tesselation

RASTER VS VECTOR

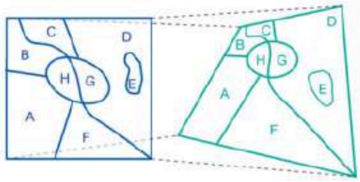


Line and point objects are more awkward to represent using rasters, as rasters are area- based

Objects are more naturally represented in vector

WHAT IS TOPOLOGY?

Topological	Non-topological
A point is at an end-point of an arc	Distance between two points
An arc is a simple arc (the arc does not cross over itself)	Bearing of one point from another point
A point is on the boundary of a region	Length of an arc
A point is in the interior of a region	Perimeter of an area
An area is open (excludes all its boundary)	Area of an area
An area is simple (has no holes)	



Why is topology useful:

- Ensuring ***geometric correctness*** of the data
  - Detecting and correcting digitizing errors
- Carrying out some types of ***spatial analysis*** (selections, network analysis):
  - Find all points that are inside a polygon (cities – admin unit)
  - Find all the polygons adjacent to a line (parcels – river)
  - Network analysis

TOPOLOGICAL DATA MODEL

Topological relationships are built ***from simple elements into more complex elements***:

- Nodes define line segments
- Line segments connect to define lines
- Lines define polygons

Topological ***simplices***

0-simplex　　1-simplex　　2-simplex

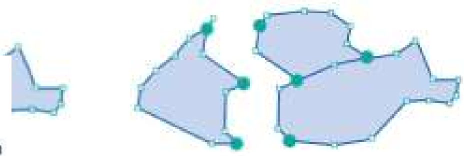
TOPOLOGICAL DATA MODEL

AREA REPRESENTATIONS

POLYGONS

A simple representation of area features would be to list for each polygon the lines that describes its boundary. This is called a ***polygon-by-polygon*** representation. Each line in the list would be a sequence that starts with a node and ends with another one.

This approach generates ***data redundancy:*** same line will be stored twice, once for each polygon



WHAT IS TOPOLOGY?

A mathematical discipline that studies geometrical properties that are unaffected by the continuous change of shape or size of figures.

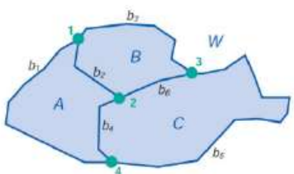
Topology deals with spatial relations that do not change under specific transformations

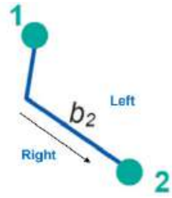
POLYGONS

The ***boundary model*** is an improved representation

It stores parts of a polygon's boundary as non-looping arcs and indicates which polygon is on the left and which is on the right of each arc.

line	from	to	left	right	vertexlist
b_1	4	1	W	A	...
b_2	1	2	B	A	...
b_3	1	3	W	B	...
b_4	2	4	C	A	...
b_5	3	4	W	C	...
b_6	3	2	C	B	...





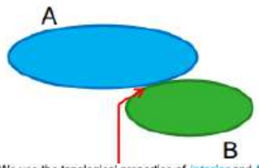
- We can determine the left and the right polygon, because the line segment has a direction.
- The direction of the line segment is from the "From Node" to the "To Node"

TOPOLOGICAL RELATIONSHIPS
MATHEMATICAL PROPERTIES OF THE GEOMETRIC SPACE

- The space is a three-dimensional ***Euclidean space*** where for every point we can determine its three-dimensional coordinates as a triple (x,y,z) of real numbers
- The space is a ***metric space***, which means that we can always compute the distance between two points according to a given distance function. Such a function is also known as a *metric*.

- The space is a ***topological space.*** For every point in the space we can find a neighbourhood around it that fully belongs to that space as well.
- ***Interior*** and ***boundary*** are properties of spatial features that ***remain invariant*** under topological mappings. When transformed, the interior and the boundary of a feature remains unbroken and intact.

2. Tessellation
Tessellations can be regular or irregular. We will start with "Regular Tessellation" followed by "Irregular Tessellation".
3. Regular Tessellation
Tessellations can be regular or irregular. We will now discuss the "Irregular Tessellation".
4. Irregular Tessellation
Tessellations do not explicitly store geo-references of the phenomena they represent. In vector representations, georeferences are explicitly associated with the geographic phenomena. Examples of vector representation are discussed in the next concepts.
5. Vector Representation
Examples of vector representations are discussed in the next concepts. To start, we will examine the TIN representation for geographic fields, which is a hybrid between tessellations and vector representations.
6. Triangulated Irregular Networks
The next vector representation to be discussed is the "Point representation".
7. Point representation
In vector representations, lines are used to represent one-dimensional objects.
8. Line representation
In the next concept we will discuss Area representations. Areas can be represented with and without topology. The Topological area data model will be discussed in the next Learning path (Spatial Data Modelling - Topology and Time).



We use the topological properties of *interior* and *boundary* to define relationships between spatial features.

Region A →


Interior is the largest set of points of A for which we can construct a disc-like environment around it that also falls completely inside A

Boundary is the set of those points belonging to A that do not belong to the interior of A (i.e. one cannot construct a disc-like environment around such points that still belongs to A completely)

When the two regions meet this is defined as the boundary of A intersects the boundary of B

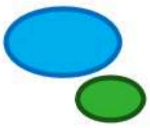
The mathematical definition of meets:

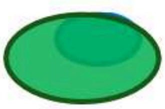
A meets B =

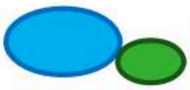
interior (A) ∩ interior (B) = Ø ∧
interior (A) ∩ boundary (B) = Ø ∧
boundary (A) ∩ interior (B) = Ø ∧
boundary (A) ∩ boundary(B) ≠ Ø

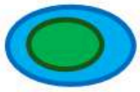
NOT EMPTY

TOPOLOGICAL RELATIONSHIPS
SET THEORY


Blue disjoint green

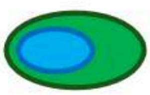

Blue covered by green


Blue meets green

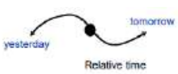

Blue contains green


Blue equal green

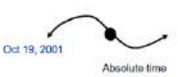

Blue covers green


Blue inside green


Blue overlap green



Relative time



Absolute time

Valid time (world time) or *transition time* (database time)

Time granularity: precision of a time value (year, month, day, second)

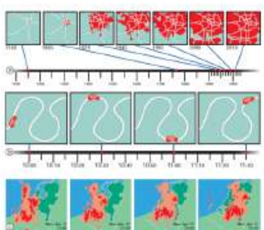
Absolute (fixed time) or *relative* time.

SPATIO-TEMPORAL ANALYSIS

Change of spatial and thematic attributes over time

Different approaches:

a. Location fixed, phenomenon varies over time
b. Phenomenon fixed, location varies over times
c. Both vary over time

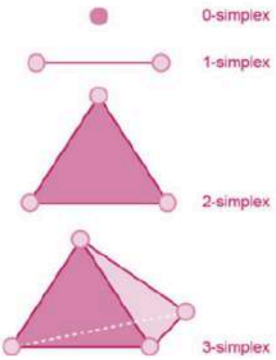


TOPOLOGICAL SIMPLICES

Features within the topological space that are ***easy to handle and that can be used as representations of geographic objects.***

These features are called simplices as they are ***the simplest geometric shapes of some dimension:***

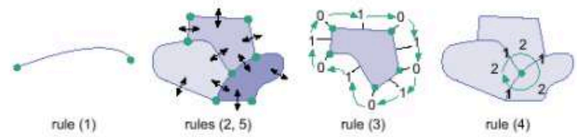
- ***point (0-simplex),***
- ***line segment (1-simplex),***
- ***triangle (2-simplex),***
- ***and tetrahedron (3-simplex).***



UNIVERSITY OF TWENTE.

RULES OF TOPOLOGICAL CONSISTENCY

1. Every 1-simplex ('arc') must be bounded by two 0-simplices ('nodes', namely its begin and end node)
2. Every 1-simplex borders two 2-simplices ('polygons', namely its 'left' and 'right' polygons)
3. Every 2-simplex has a closed boundary consisting of an alternating (and cyclic) sequence of 0- and 1-simplices.
4. Around every 0-simplex exists an alternating (and cyclic) sequence of 1- and 2-simplices.
5. 1-simplices only intersect at their (bounding) nodes.

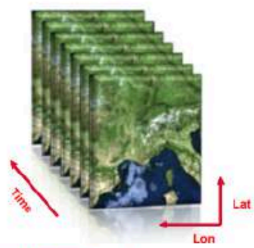


TEMPORAL DIMENSION

Geographic phenomena also change over ***time*** and are thus ***dynamic***

Time is ***continuous***. Needs to be discretized to be modeled in GIS

GISs still offer ***limited support*** to the temporal component.



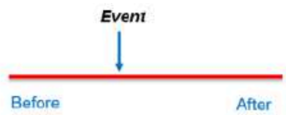
Spatio-temporal data models are ways of organizing representations of space and time in a GIS.

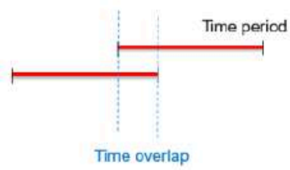
Snapshot state technique represents a single moment in time of an ongoing natural or man-made process.

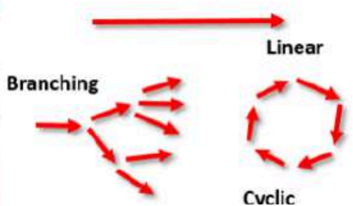
Discrete time is composed of discrete time elements (days, years).

In ***continuous*** time no such discrete elements exist (for every two moments there is another moment in between)

We can also structure time by ***events*** (points in time) or ***periods*** (time intervals)







Time can be considered to be ***linear, branching or cyclic***

RS-Applications, orbits, radiations, resolutions

Applications of Earth observation - observing, monitoring and classifying all nearby parts of the earth system such as, atmosphere, lithosphere, hydrosphere, cryosphere, biosphere and anthroposphere.
Atmosphere - around us, Lithosphere - under our feet/land, Hydrosphere- ocean, lakes and waterbody, Cryosphere- Ice and glacier, Biosphere - vegetation and humans, Anthroposphere - humane made structure and effects.
Geo- stationary orbit- stares at the same position of earth. Images acquired once every hour, low resolution. Geo stationary look at earth at one region, they move around the earth tat the same speed as the earth is rotation. It stares at the same position at the earth, It has low resolution, it is used for weather. Altitude - 35,786Km Use - weather, communication, navigation. Examples - MeteoSat, GOES, Himawaeri, FengYun
Low Earth Orbit (Polar orbit)- It goes from pole to pole, mostly inclined toward the pole. Passes over whole globe, same local time at each overpass, image acquired between twice parlayed once every few days. Medium to high spatial resolution.
Examples- ESA sentinels, NASA landstats , ISS, telecom

GEO-STATIONARY ORBIT

Meteostat is an e

- Altitude: 35,786 km
- Characteristics:
  - continually "stares" at same portion of Earth
  - images acquired once every hour or so
  - low spatial resolution
- Instrument types:
  - weather
  - communication, navigation
- Examples:
  - MeteoSat, GOES, Himawari, FengYun

LOW EARTH (POLAR) ORBIT

Used for earth sand space observation. Every time it sees the different position of the earth so it covers the entire globe

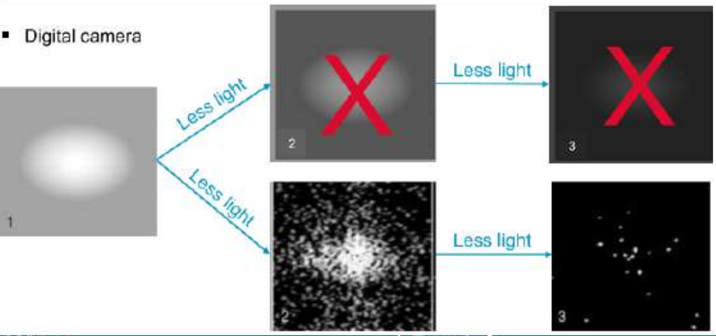
- Altitude: 200-1200 km
- Characteristics:
  - Passes over whole* globe
  - In sun-synchronous orbit, same local time at each overpass
  - Images acquired between twice per day and once every few days
  - Medium to high spatial resolution
- Instrument types:
  - EO, Deep space observation, ISS, telecom
- Examples:
  - ESA Sentinels, NASA Landsats, Cubesats

EMR

Satellite sees electromagnetic radiation.
Electromagnetic radiation - It travel at the speed of light, it travels in a wave. We see both wave and if the light is decrease we see photons, we use both wane and the particle(photon) approach,
wave approach is used when we talk about reflection, refraction and emission and particle approach is used when we talk about absorption, scattering, emission.
Source of EM - Every single object that has temperature higher than 0 kelvin (T>0 K) emits
radiation.

EMR AS A PARTICLE

- Digital camera



Wave	Particle
Reflection	Absorption
Refraction	Scattering
Emission	Emission

- When the wavelength(λ) is small the frequency(ν) is large.
- The higher the temperature the shorter the wavelength.
- Increase in the temperature the intensity increases.
- Speed of light is constant, higher the wavelength smaller the frequency and vice versa.
- EMR moves through space at the speed of light.
- Energy of photon is related/depends to its frequency.
- Colour blue has higher energy since more electros emitted hence more scattering.
- short wavelength implies a high frequency, while long wavelengths are equivalent to low frequencies. Blue light has a higher frequency than red light.
- Only radiation with correct wavelength causes ejection of electron.
- Lower the resolution the bad the picture, resolution starts with radiation.

Blackbody Radiation - Both the intensity and wavelength both depends on its temperature. It absorbs all incident radiations. **Planks law is used**. The black body does not chance its temperature, so outgoing flux would be the same as the incoming flux, if heat goes in then the heat goes out.
Wein's law - It stated that the wavelength of the peak intensity is equal to certain constant value (b) divided by temperature. $\lambda_{max} = \frac{b}{T}$
Where, λ = wavelength, b = constant, t= temperature.

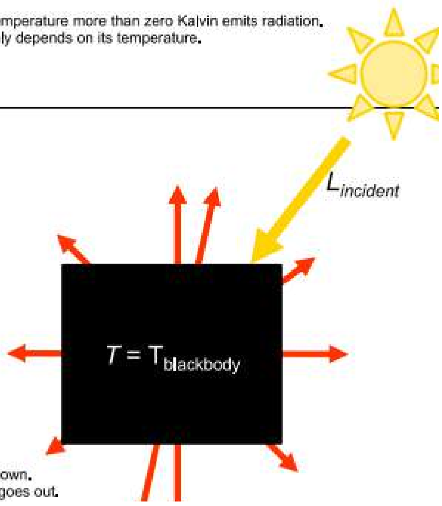
Emission by earth and sun - sun is much more hotter than the earth, which mean sun emits a lot more radiation than the earth. The emission curves pr the sun and earth barely overlaps.

Sources of radiations are every where, every object that has temperature more than zero Kalvin emits radiation.
Black body emits radiation where the intensity and radiation only depends on its temperature.

BLACKBODY RADIATION

- Every object with T>0 K emits radiation
- Blackbody:
  - Absorbs all incident radiation
  - Emitted radiation depends only on its temperature
- Planck's law: $L_{BB} = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}}-1}$
- In thermal equilibrium (no change in T):

$L_{BB} = L_{incident}$



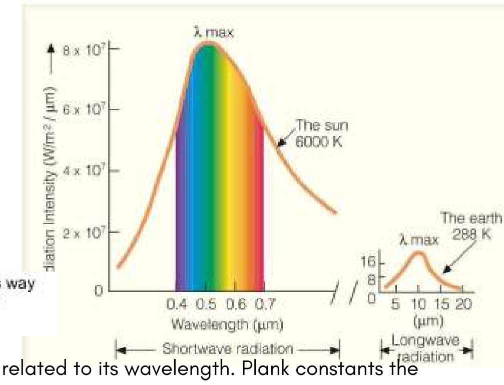
By knowing the temperature of black body spectral flux is known.
Black bodies receives radiation as well, if heat goes in heat goes out.

EMISSION BY SUN AND EARTH

- Sun is much hotter than Earth -> λ_{max} shorter and M much greater
- Emission curves of Sun and Earth barely overlap

The emission curves of the sun and the earth barely overlaps

Sun emits more radiation than earth since it is way higher in temperature than the earth. The sun emits million time more.



Max plank - The energy of a photon is related to its wavelength. Plank constants the frequency/speed of light. Energy of a photon is related to its wavelength to its frequency Q is the energy of one photon, planks constant times the frequency, speed of light dived by its wavelength, energy of photon is dependents on its frequency higher the energy higher the frequency. Gamma rays are higher in energy and then it goes down followed by uv and so on.

PLANCK'S LAW OF RADIATION

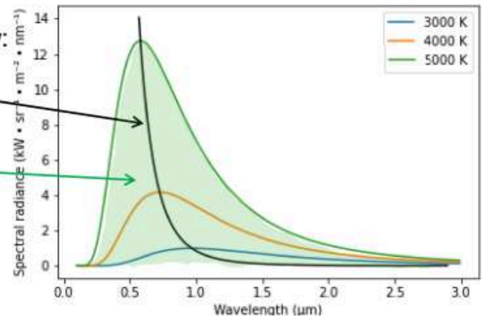
- Emission of radiation is governed by Planck's law:

$$L_{BB}(\lambda, T) = \frac{2hc^2}{\lambda^5} \cdot (e^{hc/\lambda kT} - 1)^{-1} \text{ [Wm}^{-2}\mu\text{m}^{-1}]$$

- Derivations from Planck's law:
- Wien's law: $\lambda_{max} = \frac{b}{T}$
- Stefan-Boltzmann equation:

$M = \sigma T^4$

Area under the curve (green thing) is calculated by Stefan equation



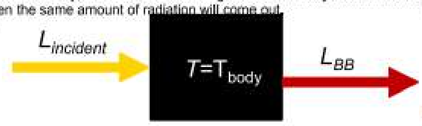
UNIVERSITY OF TWENTE.

Properties of EMR - Maxwell defined EMR as coupled electric and magnetic fields traveling through space at the speed of light. Speed of light is always constant.
When the wavelength(λ) is small the frequency(ν) is large. Frequency and wavelength of EM
radiation area related to the speed of light. **The higher the temperature the shorter the wavelength**. EMR can occur naturally or can be man-made.
Real objects - they are not like blackbodies they reflect or transmit. T is constant. It might be reflective like a mirror in this case some of the radiation is reflected and then some emitted, some bodies are also transparent which means the radiation goes through it, it's not emitted but transmitted.

REAL OBJECTS

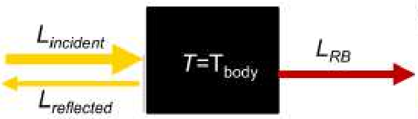
This represent a black body, the incident radiation goes in if the body does not chan temperature then the same amount of radiation will come out.

- Real objects are not blackbodies, they
  - reflect and/or
  - transmit



- In thermal equilibrium (T constant):

$$L_{in} = L_E + L_T + L_R$$



This represent a Real body, it might be reflected like a mirror meaning if the incident goes in some of the light is reflected, and then the rest is emitted.

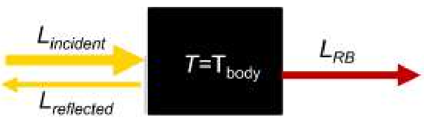
some of the real body are also transparent that means the radiation goes through them, they are not emitted but transmitted

UNIVERSITY OF TWENTE.

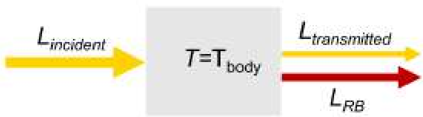
- Real objects are not blackbodies, they
  - reflect and/or
  - transmit

- In thermal equilibrium (T constant):

$$L_{in} = L_E + L_T + L_R$$



- It follows: $L_{BB} > L_{RB}$
- Emissivity $\varepsilon_{RB} = \frac{L_{RB}}{L_{BB}}$



UNIVERSITY OF TWENTE.

2.1 Waves and photons

EM wave

EM radiation can be modelled in two ways: by waves, or by radiant particles calle photons. The first publications on the wave theory date back to the 17th century. According to the wave theory, light travels in a straight line (unless there are externa influences) with its physical properties changing in a wave-like fashion. Light waves have two oscillating components: an electric field and a magnetic field. We refer, therefore, in this context to electromagnetic waves. The two components interact—an instance of a positive electric field coincides with a moment of negative magnetic field (Figure 2.1). The wave behaviour of light is common to all forms of EM radiation. All EM waves travels at the speed of light, which is approximately equal to 2.998×10^8 m s^{-1}. This is fast, but the distances in space are literally astronomical: it takes eight minutes for the sunlight to reach the Earth, thus when we see, a sunrise, for example, the light particles actually left the Sun that much earlier. Because they travel in a straight line, we use the notion of light rays in optics.

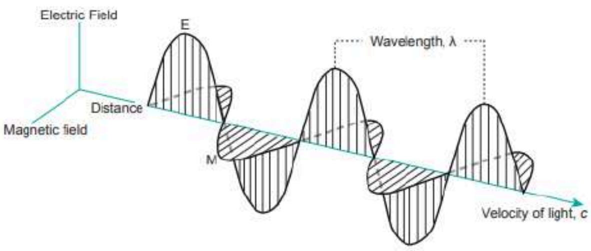
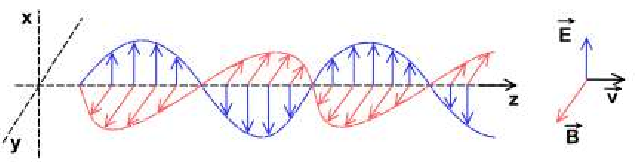


Figure 2.1
The two oscillating components of EM radiation: an electric and a magnetic field.

A sine wave can be described as:

- Maxwell's theory: coupled electric and magnetic fields travelling through spa



electromagnetic radiation was defined by maxwell, electromagnetic radiations moves through space it needs nothing move, it moves through speed of light.

2.2 Sources of EM radiation

All matter with a temperature above absolute zero emits EM radiation because of molecular agitation. *Planck's law of radiation* describes the amount of emitted radiation per unit of solid angle in terms of the wavelength and the object's temperature:

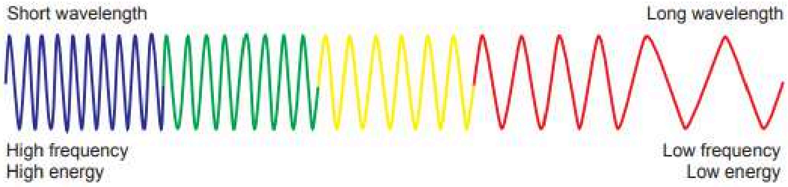
$$L(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}, \tag{2.4}$$

where h is the Planck's constant, $k \approx 1.38 \times 10^{-23}$ J K^{-1} is the Boltzmann constant, λ is the wavelength (m), c is the speed of light and T is the absolute temperature (K). $L(\lambda, T)$ is called the spectral radiance.

the wave. The reciprocal of the period is called the *frequency* of the wave. Thus, the frequency ν is the number of cycles of the wave that occur in one second. We usually measure frequency in hertz (1 Hz = 1 cycle s^{-1}). Since the speed of light c is constant, the relationship between wavelength and frequency is:

$$c = \lambda \times \nu. \tag{2.2}$$

Obviously, a short wavelength implies a high frequency, while long wavelengths are equivalent to low frequencies. Blue light has a higher frequency than red light (Figure 2.3).



Although wave theory provides a good explanation for many EM radiation phenomena, for some purposes we can better rely on particle theory, which explains EM radiation in terms of photons. We take this approach when quantifying the radiation detected by a multispectral sensor (see Section 2.6). The amount of energy carried by a photon of a specific wavelength is:

$$Q = h \times \nu = h \times \frac{c}{\lambda}, \tag{2.3}$$

where Q is the energy of a photon measured in joules (J) and h is Planck's constant ($h \approx 6.626 \times 10^{-34}$ J s).

The energy carried by a single photon of light is just sufficient to excite a single molecule of a photosensitive cell of the human eye, thus contributing to vision. It follows from Equation 2.3 that long-wavelength radiation has a low level of energy while short-wavelength radiation has a high level. Blue light has more energy than red light (Figure 2.3). EM radiation beyond violet light is progressively more dangerous to our body as its frequency increases. UV radiation can already be harmful to our eyes, so we wear sunglasses to protect them. An important consequence of Formula 2.3 for RS is that it is more difficult to detect radiation of longer wavelengths than radiation of shorter wavelengths.

- Maxwell's theory: coupled electric and magnetic fields travelling through space

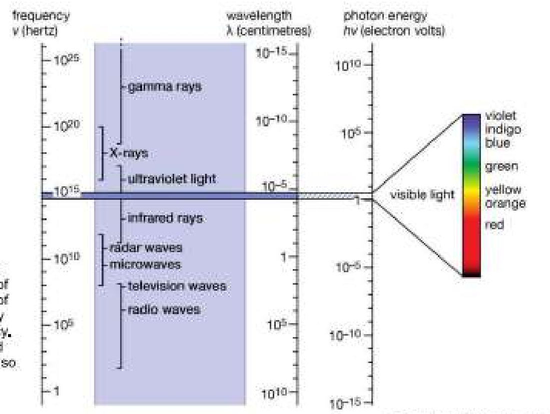
- Key characteristic of waves: $c = \lambda\nu$ (c speed of propagation, here: speed of light)
- It follows: wavelength λ is inversely related to frequency ν

speed of light is always constant, higher the wavelength smaller the frequency and vice versa

ENERGY OF EMR

- If wavelength λ is small, frequency ν is large
- Max Planck: energy of a photon $Q = h\nu = \frac{hc}{\lambda}$

Energy of a photon is related to its wavelength to its frequency
Q is the energy of one photon, planks constant times the frequency, speed of light dived by its wavelength, energy of photon is dependents on its frequency higher the energy higher the frequency. Gamma rays are higher in energy and then it goes down followed by uv and so on.



UNIVERSITY OF TWENTE.

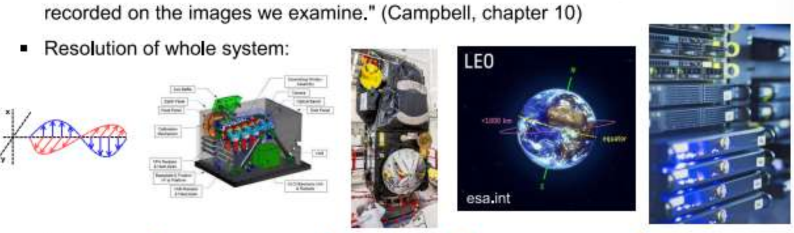
By frequency
By wavelength
By photon energy

1.Ultraviolet, visible, infrared, microwave and radar. (opposite of 2nd because higher the frequency lower the wavelength)
2. Radar, microwave, infrared, visible, ultraviolet.
3. Ultraviolet, visible, infrared, microwave and radar. (because it energy of photos as same as wht)

- **Resolution** - Degree to which two features can be distinguished. It starts with radiation, then the instrument characteristics also affect the resolution on the platform on which the instrument is mounted and in the orbit then lastly the computer. The resolution domains that are relevant for earth observation are - **spatial, spectral , temporal and radiometric.**
- **Spectral resolution** - distance between consecutive steps. UV and NIR range 0.1 - 100s of nm**.**
- **Spatial resolution - It** is defined by pixels and pixels size. Temporal resolution - frequency of observation, In case of LEO - orbit inclination and scan and in case of GEO it depends on the repeat time.
- **Radiometric resolution** - it is define by bits, more bits mean more shades in colour hence the image is more clear.
- These resolution bits are all interconnected, there are always trade-offs when the resolution is built. The atmosphere, the slant, illumination can affect the resolution.

WHAT DO WE MEAN BY RESOLUTION?

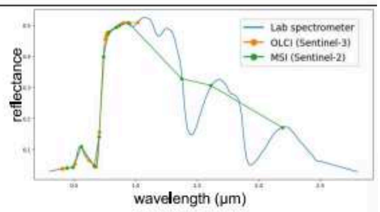
- Degree to which two features can be distinguished
- "Knowledge of image resolution is a prerequisite for understanding the information recorded on the images we examine." (Campbell, chapter 10)
- Resolution of whole system:

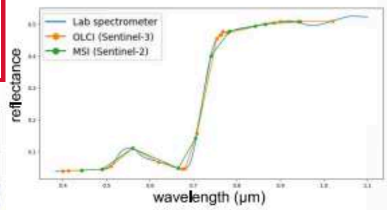


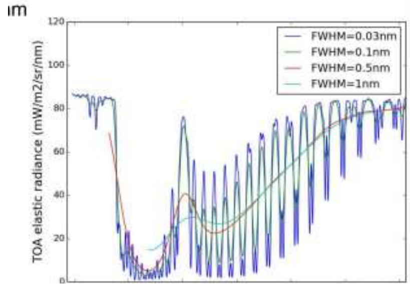
radiation + instrument + platform + orbit + computer

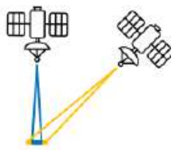
SPECTRAL RESOLUTION

- Distance between consecutive steps
- In UV-visible-NIR range, 0.1 - 100s of nm
- Examples:
  - Lab (2 nm (vis) or 100 nm (NIR))
  - MSI on Sentinel-2 (23-500 nm)
  - OLCI on Sentinel-3 (10-100 nm)
- For atmospheric composition: <0.1 nm!



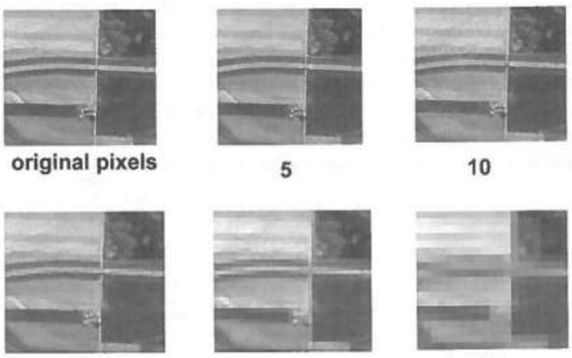






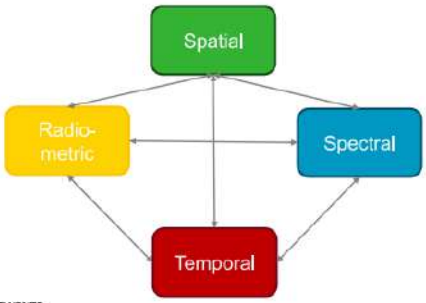
SPATIAL RESOLUTION

- Footprint extent
- Definition dependent on shape



UNIVERSITY OF TWENTE. Campbell, Chap.10

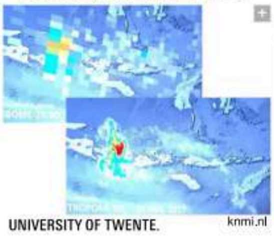
Interplay

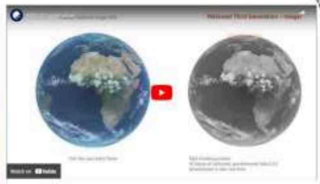


UNIVERSITY OF TWENTE.

TEMPORAL RESOLUTION

- Frequency of observation
  - LEO: orbit inclination and scan width (swath)
  - GEO: repeat time (GEO)
- Accuracy (time averages)

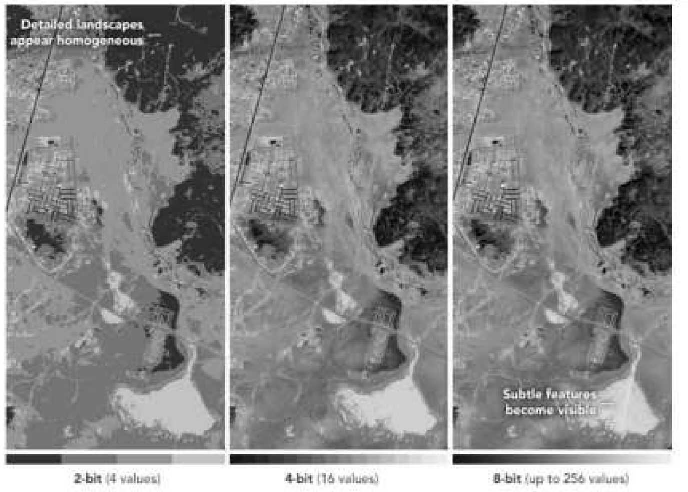




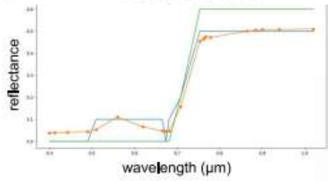
UNIVERSITY OF TWENTE. knmi.nl

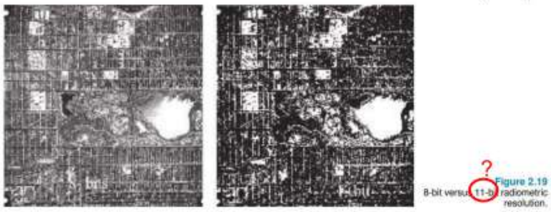
https://aerospace.csis.org/aerospace101/earth-orbit-101/
https://www.eumetsat.int/mtg-flexible-combined-imager-fci

RADIOMETRIC RESOLUTION



- Discrete observation <-> continuous signal
- Number of bits used to record (or save) signal
- More discrete bits -> closer to real signal

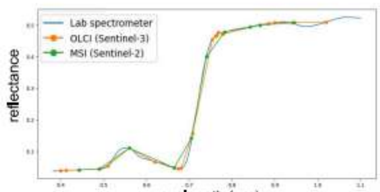




UNIVERSITY OF TWENTE. Core textbook, chapter 2.

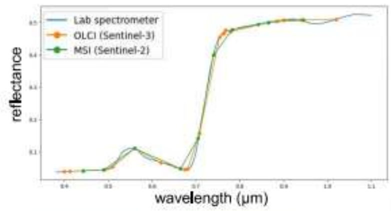
RESOLUTION TRADE-OFFS

- Trade offs between spatial & spectral
  - OLCI spectral resolution
  - MSI spatial resolution
  - and temporal!



RESOLUTION TRADE-OFFS

- Trade offs between spatial & spectral
  - OLCI spectral resolution
  - MSI spatial resolution
  - and temporal!
- Confounding factors: atmosphere, illumination
  - Topic of atmospheric correction (lecture next week!)

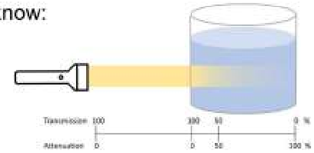


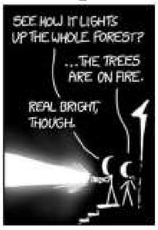
UNIVERSITY OF TWENTE.

RS- Radiative transfer in atmosphere

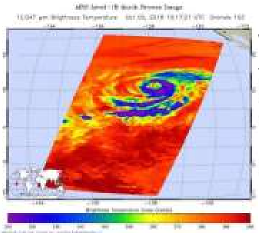
APPLICATIONS OF RADIATIVE TRANSFER

- You need to understand RT when you want to know:
  - how far your signal will travel
  - how strong your source needs to be
  - what you're looking at (remote sensing)
  - what your radiation might do...

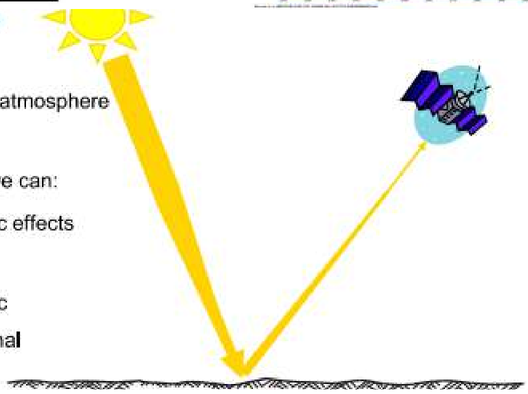




xkcd



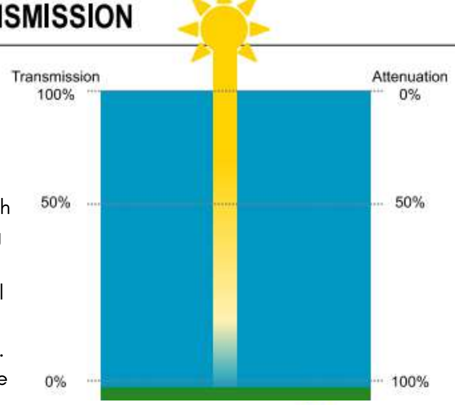
- Interpretation of satellite observations
- Radiation is affected by atmosphere and surface
- If we understand how, we can:
  - Remove atmospheric effects from signal
  - Retrieve atmospheric information from signal



ATTENUATION AND TRANSMISSION

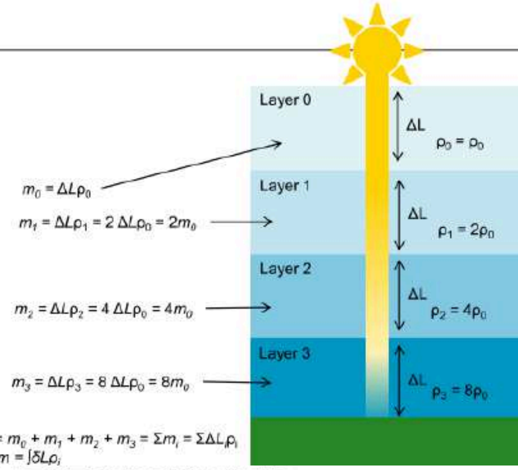
- Radiation is either attenuated or transmitted by the atmosphere
- Radiation can be attenuated by scattering or absorption

Relative air mass- slant path through the atmosphere is longer, normalising the slant airmass is called relative air mass. Slant paths have longer optical air mass. When the relative air mass increases the transmission decreases. With the increase in solar zenith angle the relative air mass increases



ATTENUATION

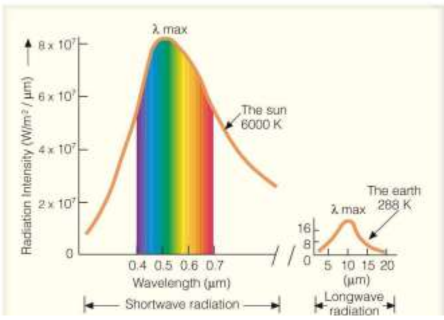
- Degree of attenuation (optical air mass m) depends on:
  - path length through medium (L) and
  - density of medium (ρ)
- But the atmosphere is layered...



UNIVERSITY OF TWENTE.

EMISSION BY SUN AND EARTH

- Sun is much hotter than Earth -> λ_{max} shorter and M much greater
- Emission curves of Sun and Earth barely overlap



SO HOW TO CALCULATE TRANSMISSION?

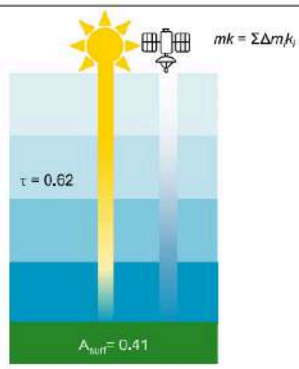
- Bouguer's law: transmission $\tau = e^{-mk}$
  - m = absolute air mass
  - k = extinction coefficient (probability)
- For medium with N layers: $mk = \sum_{i=0}^{i=N} m_i k_i$
- Hence: $\tau = e^{-\sum_{i=0}^{i=N} m_i k_i} = \prod_{i=0}^{i=N} e^{-m_i k_i}$
- So transmissions are multiplicative!

$$\tau_{\uparrow} = e^{-\int_{z=BOA}^{z=TOA} m(z)k(z)dz}$$

$$\tau_{\downarrow} = e^{-\int_{z=TOA}^{z=BOA} m(z)k(z)dz}$$

$$\tau = \tau_{\uparrow} \cdot \tau_{\downarrow} = 0.62 \cdot 0.62 = 0.38$$

$$R = A_{surf}\tau_{\downarrow}\tau_{\uparrow} = 0.41 \cdot 0.62 \cdot 0.62 = 0.16$$

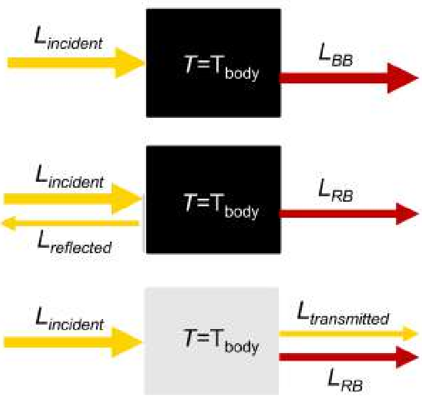


REAL OBJECTS

- Real objects are not blackbodies, they
  - reflect and/or
  - transmit
- In thermal equilibrium (T constant):

$$L_{in} = L_E + L_T + L_R$$

- Emissivity $\varepsilon_{RB} = \frac{L_{RB}}{L_{BB}}$



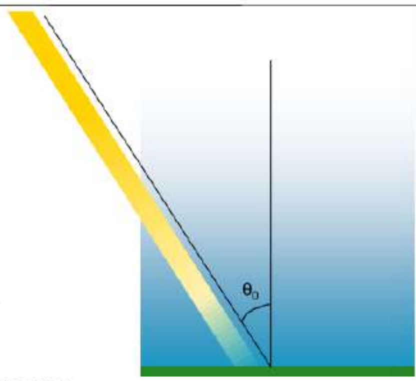
UNIVERSITY OF TWENTE

RELATIVE AIR MASS

- Slant paths are longer -> larger optical air mass
- Normalization using relative air mass, m_r

$$m_r = \frac{L_S \cdot \rho}{L \cdot \rho} = \frac{m_s}{m}$$

- Radiation from Sun to Target: $m_r = \cos\theta_0$
- So: $m = m_r \Sigma\Delta L\rho_i$



CALCULATE TRANSMISSIONS (2)

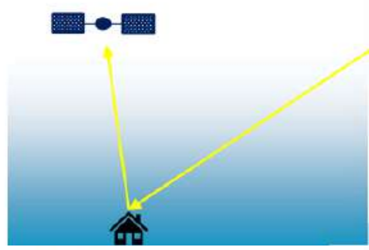
A medium has four constituents with optical transmissivities:

$\tau_1 = 0.99$, $\tau_2 = 0.98$, $\tau_3 = 0.97$, $\tau_4 = 0.96$

What is the total transmission of the medium?

Bouguer's law: transmission $\tau = e^{-mk}$

With increasing solar zenith angle, the relative air mass increases, hence the transmission...



Satellites do not measure cloud cover or soil moisture content or even temperature

Satellite instruments detect upwelling (reflected, scattered, emitted) radiation

It's up to the scientist to interpret the radiation patterns

Satellites do not measure cloud cover, soil moisture or even temperature the satellite detects **reflected, scattered, emitted radiations**. For understanding it attenuation and transmission is required.

Attenuation and Transmission - Radiation is either attenuated or transmitted by the atmosphere. Radiation can be attenuated by scattering or absorption. Radiations comes from the sun its either attenuated or transmitted by the atmosphere, attenuated means its decreasing and transmission means that it is going through.

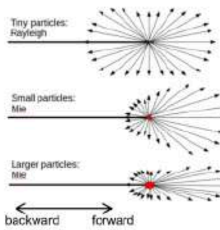
- At the top of the atmosphere the sun rays are transmitted and then below the atmosphere the sun rays are attenuated.
- Degrees of attenuation is also called optical air mass (m). It depends on the path length of the medium and on the density of the medium.
- Path length id how far the radiation travels if the path length and the density of the particles increases then the attenuation increases.
- Atmosphere is layered and each layer has its own density as the density decreases.

Transmission - **Bouguer's law is used to calculated transmission**. Transmissions are multiplicative, meaning total attenuation from bottom of atmosphere to the top of the atmosphere and vice versa the 60 percent of the radiation coming from the sun is taken by the satellite and 40 percent is absorbed by the ground. To calculate total transmission multiply all the transmissions.

SCATTERING

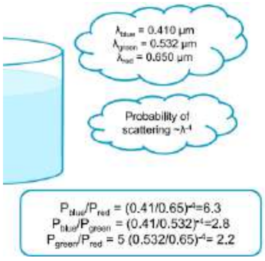
- Interaction of radiation with large objects (>> wavelength)
  - Reflection
  - Absorption

- Interaction of radiation with small objects (≤ wavelength)
- Absorption + (partial) re-emission
- Re-emission in all directions: scattering



RAYLEIGH SCATTERING

- Scattering is continuous, but depends on wavelength:
  - Size parameter $\alpha = \pi D/\lambda$
- Rayleigh scatter: $\alpha \ll 1$
- Probability of Rayleigh scatter $\sim\lambda^{-4}$
- Rayleigh scatter explains why we perceive the sky as blue and sunsets as red



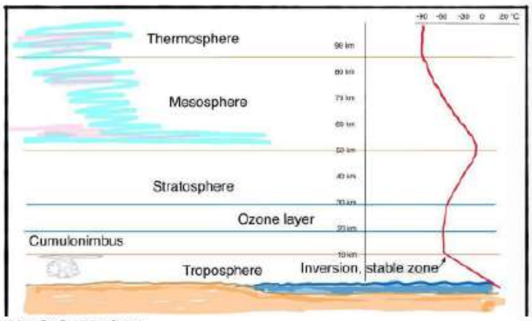
Scattering - It is one of the process of seeing attenuation in the atmosphere. It is interaction of radiation with large objects. It is a continues process it depends of wavelength(λ) and the size parameter. The probability of scattering depends upon λ^{-4}

Rayleigh scattering - it contains of tiny particles and they scared throughout like in example **sunsets or the sky as blue,** it highly depends on the wavelength. The longer the wavelength less the Rayleigh scatter.

Mie scattering - it contains larger to medium size particles meaning there is more secreting in the forward direction.

ATTENUATION IN THE ATMOSPHERE

- Atmosphere is layered:
- Layers separated by temperature profile
- Density decreases exponentially with altitude
- Most gases located in lowest layer



ATMOSPHERIC COMPOSITION

- Nitrogen, oxygen, and argon make up 99.96% of the dry atmosphere
- Carbon dioxide (CO_2) adds 0.04%
- Water vapour is variable, mostly in troposphere

- ppm = parts per million (10^6)
- ppb = parts per billion (10^9)

Constituent gas	Content % Volume or ppmv
Nitrogen (N_2)	78.084 %
Oxygen (O_2)	20.946 %
Argon (Ar)	0.934 %
Carbon Dioxide (CO_2)	0.0416 % or 416 ppm
Other noble gases (Ne, He, Kr, Xe)	0.0024 % or 24 ppm
Hydrogen (H_2)	0.00006 % or 6 ppm
Methane (CH_4)	1.8 ppm
Nitrous oxide (NO_x)	0.27 ppm
Ozone (O_3)	0 to 0.04 ppm
Sulfur dioxide (SO_2)	0.01 ppm
Nitrogen dioxide NO_2	0.001 ppm
Ammonia (NH_4)	0.02 ppm
Carbon Monoxide (CO)	0.09 ppm
Nitric Oxide (NO)	0.005 ppm
Hydrogen sulfide (H_2S)	0.002 ppm
Nitric acid vapor	traces

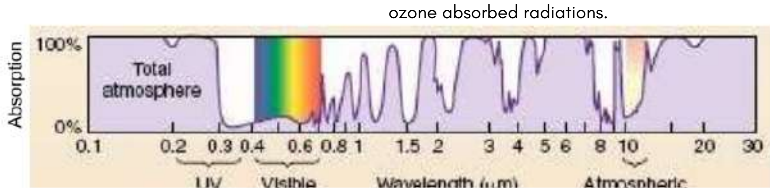
Permanent gases

Variable gases

ABSORPTION IN THE ATMOSPHERE

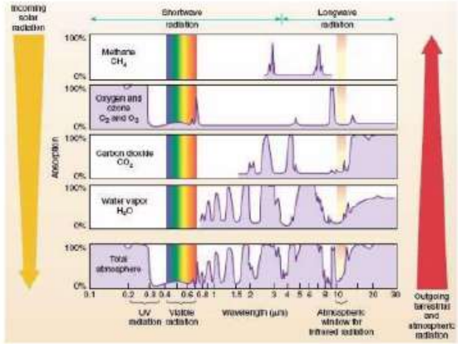
- Question: In which wavelength regions is the atmosphere (mostly) transparent to radiation?

Absorption - Atmosphere is separated by their temperature profile by layers, zero percentage means the radiation goes through the surface and zero percent means it does not goes through, in visible region radiation reaches us. Water vapour and carbon dioxide, oxygen and ozone absorbed radiations.

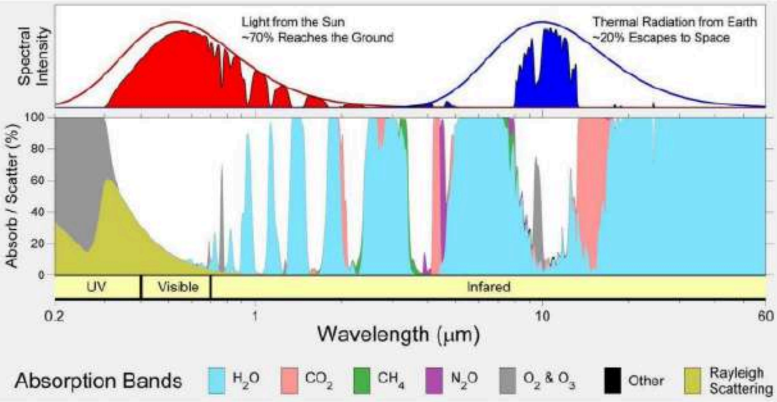


ABSORPTION IN THE ATMOSPHERE

- Main absorbers in the atmosphere: O_3, H_2O, CO_2
- Most of the radiation emitted by Earth is absorbed
- In thermal equilibrium: energy received = energy emitted
- Atmospheric windows allow Earth to cool down!



GREENHOUSE EFFECT /ABSORPTION



https://twitter.com/rarohde/status/1196761516051238912

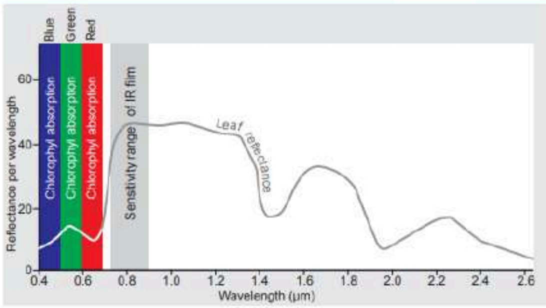


Figure 2.13 An idealized spectral reflectance curve of healthy vegetation.

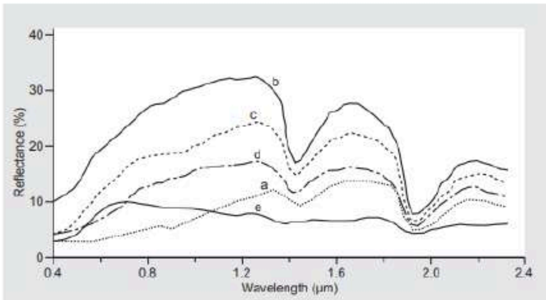


Figure 2.14 Spectral reflectance of five mineral soils: (a) organic dominated, (b) minimally altered, (c) iron altered, (d) organic affected and (e) iron dominated (from [71]).

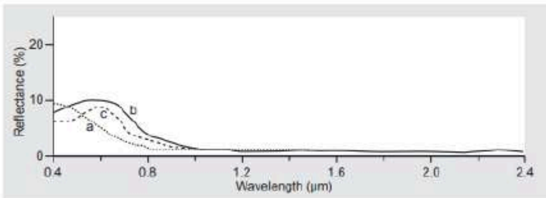


Figure 2.15 Typical effects of chlorophyll and sediments on water reflectance: (a) ocean water, (b) turbid water, (c) water with chlorophyll (from [71]).

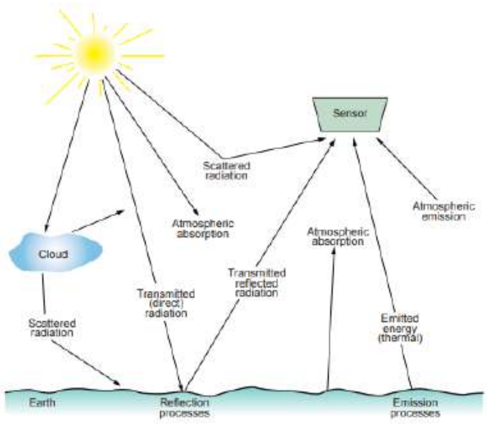


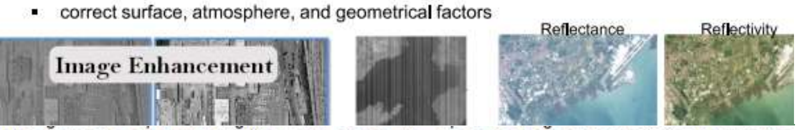
Figure 2.6 Interactions of EM radiation with the atmosphere and the Earth's surface.

Atmospheric Correction

Atmospheric correction refers to the removal of atmospheric effects in order to obtain the pure radiometric signal from the surface. We apply atmospheric correction if we want to obtain information about the surface of the earth. In that case, atmospheric effects can be considered as a disturbance of the measurements, for which we need to correct.

APPLICATIONS OF ATMOSPHERIC CORRECTION

- You need to do an AC to:
  - improve visualization (image enhancement)
  - correct sensor inaccuracies
  - correct surface, atmosphere, and geometrical factors



The larger the solar and viewing zenith angles, the longer the path of the radiation through the atmosphere and the larger the air mass. The relative optical air mass, m_r accounts for this. It is given by the ratio of the air mass encountered by the solar radiation to the vertical air mass:

$$m_r = \frac{\int_0^\infty \rho \cdot ds}{\int_0^\infty \rho \cdot dz} = \frac{m_s}{m} \qquad \{1\}$$

Where z is the vertical distance to the surface (or altitude), m the absolute air mass, and m_s the air mass observed by the satellite instrument.

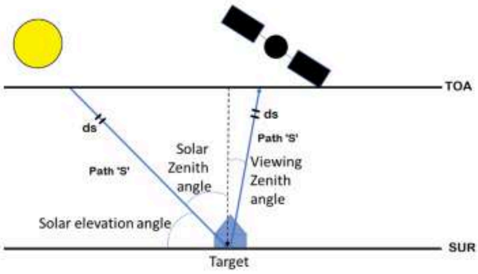


Figure 3: Geometry Sun-Target-Sensor and angle definitions.

Notice that the beam of light passes through the atmosphere twice, Sun-Target and Target-Sensor, so the radiation is attenuated twice.

For a clear atmosphere (no aerosols) and weak absorption (away from strong ozone or water vapour absorption bands) and solar zenith angles less than 70°, the relative air mass can be approximated as $m_r = {}^{1}/_{\cos(\theta_z)}$ where θ_z is the solar zenith angle. The same relation holds for the viewing angle.

Each component of the atmosphere (gas or aerosol) has its own specific ability to attenuate radiation of a specific wavelength, as explained later in this note. We can express this ability by means of a so-called extinction coefficient k. The attenuation of radiation, a, passing through a homogeneous medium can be evaluated by Bouguer's law[3].

$$a = 1 - I/I_o = 1 - e^{(-k \cdot m)} \qquad \{2\}$$

- I= radiation after the medium [W m^{-2} µm^{-1}]
- I_0= radiation reaching the medium. [W m^{-2} µm^{-1}]
- k = extinction or attenuation coefficient. [m^{-1}]
- m = optical depth or absolute air mass (defined above). [m]
- $m{\cdot}k$= extinction optical thickness [dimensionless].

IRRADIANCE, RADIANCE AND REFLECTANCE

- Irradiance I_0:
  - Description of source strength
  - flux per unit area (in Wm^{-2})
- Radiance I:
  - Measured by satellite instrument
  - flux per unit area and solid angle (in Wm^{-2}sr^{-1})
- Reflectance R:
  - Property of surface, atmosphere, geometry
  - $R = \dfrac{\pi I}{I_0 \cos\theta_0}$
  - NOT equivalent to reflectivity!

Attenuation and transmission

When radiation travels through the atmosphere, it is either transmitted or attenuated. Attenuation occurs when radiation interacts with matter (molecules or particles), affecting the radiation. Transmission is the opposite: there is no interaction and no effect on the radiation. A beam of light passing through vacuum cannot interact with matter and is not attenuated, therefore all radiation will be transmitted, like above the blue atmosphere of Fig. 1. In contrast, if the beam passes through a dense medium (for example the atmosphere in Fig. 1), the radiation interacts with matter and is attenuated.

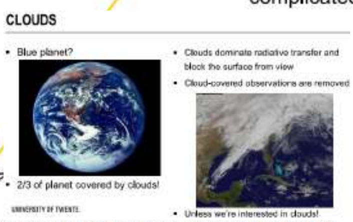
Radiation can be attenuated in two ways:

(1) by **scattering**: The direction of the radiation is changed, but its wavelength does not change[2].

(2) by **absorption**: The energy of the photon is taken up by a molecule or particle, causing excitation of the particle (change in electronic, vibrational, or rotational state). This energy quickly dissipates in the atmosphere, leading to an increase in temperature.

In the next section, we will discuss scattering and absorption in further detail. But first, we focus on the sum of the two, the attenuation. The degree of attenuation depends on two factors:

1. The path length through the medium. The longer the path, the more attenuation occurs (Fig. 1).
2. The density (ρ) of medium in the path. The higher the density, the higher the probability of interactions with molecules or particles, and the higher the chance for attenuation.



The attenuation due to Rayleigh scattering varies approximately as a function

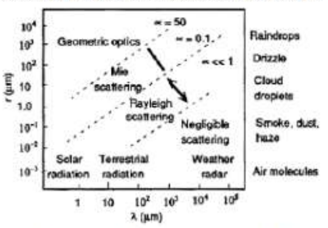


Figure 7: Different regions of scattering as a function of particle size (radius in this case) and wavelength. An indication of the size of particles normally found in the atmosphere appears to the right of the

Scattering changes the direction of the radiation, but the direction is not random (Fig. 6). When the radiation continues (more or less) in the same direction after being scattered, we use the term **forward scattering**. If the radiation continues (more or less) in the opposite direction after being scattered, we use the term **backward scattering.**

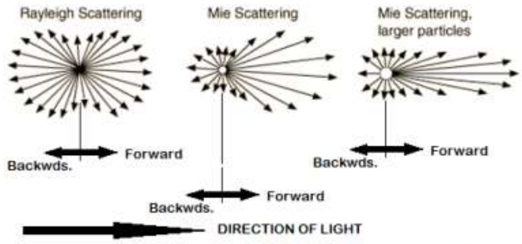


Figure 6: Scattering on small, spherical particles: for Rayleigh scattering, the probability of radiation to be scattered in the backward and forward directions are equal. Mie scattering on larger particles yields ever more forward scattering. In all cases, in this figure, the incoming radiation travels from left to right.

Rayleigh scattering causes roughly the same amount of radiation to be scattered backward and forward, whereas for Mie scattering, more forward scattering occurs. Consequently, the amount of diffuse radiation reaching the ground (or a satellite instrument) will be very different depending on the size of the particles present and the wavelength of radiation.

Aerosols

Aerosols is a term that has come to mean particles suspended in the atmosphere with sizes between a few nanometers and a few micrometers, with the exception of cloud droplets.

Aerosols can be roughly divided into primary and secondary aerosols. The most important primary aerosols are mineral dust, sea salt, and soot (smoke). In contrast to primary aerosols, which are emitted directly, secondary aerosols form by chemical and physical reactions in the atmosphere. There is a suite of gases, such as sulphur dioxide, nitrogen oxides, and organic compounds, that condense into small droplets and can remain airborne for many hours to days. Smog is an excellent example of secondary aerosols. Secondary aerosols are generally much smaller than primary aerosols, roughly 0.1-0.3 µm versus 1-10 µm. The wavelength of UV radiation is about 0.3 µm and that of near-infrared radiation 1-2 µm. Do you notice something? Secondary aerosols are smaller than the wavelengths we observe whereas primary aerosols are larger.

Aerosols can be quantified in:

- Number of particles per cm^3: the so-called number concentration.
- Aerosol optical depth (AOD) or thickness (AOT), describing the attenuation due to aerosols in the atmosphere.
- Visibility: A measure for the effect of aerosols on the attenuation of light. It is defined as the distance required to attenuate green light at 550 nm to 2% of the original intensity.

AEROSOLS

Small particles (0.1-10 µm) suspended in the atmosphere

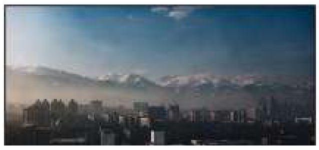
Scatter and absorb radiation

Residence time: hours-days-months

Separation into primary and secondary aerosols

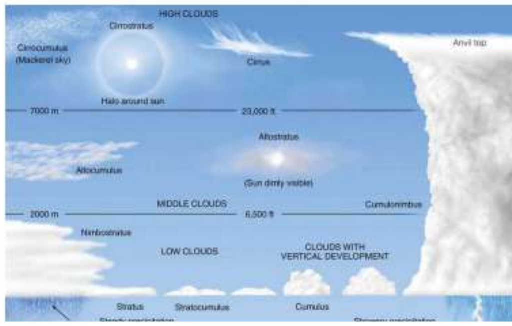
Secondary aerosols

- Form from gases in atmosphere
- Small (0.1-0.5 µm), round droplets
- Smog, vog, biogenic aerosols



CLOUDS

- Nearly all clouds in troposphere
- No absorption of visible radiation
- Warm clouds: water droplets >10 µm
- Cold clouds: ice particles
- Radiative transfer complicated!



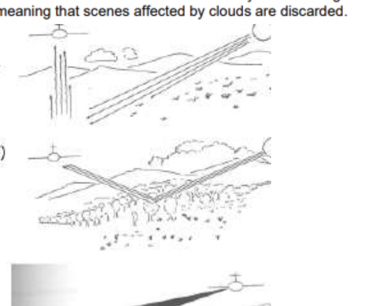
Clouds

Clouds are, technically speaking, a collection of very large aerosols. For most practical purposes in (satellite) remote sensing, however, clouds and aerosols are treated separately. Clouds consist of droplets of (nearly) pure water on the order of 20 µm in diameter. Water does not absorb visible light, but cloud droplets scatter radiation from the UV to the infrared. Clouds are very bright, reflecting on average about 80% of visible radiation. Because clouds are a dense collection of droplets, radiation scattered into the cloud is often scattered several times within the cloud before being scattered out again (multiple scattering).

The fact that clouds generally dominate the spectral signal of a scene, their physical complexity, and their variability, makes it difficult to apply a correction to measurements affected by clouds. Images are therefore generally first cleared of clouds, meaning that scenes affected by clouds are discarded.

ACTIVE & PASSIVE

- Passive: receiver only
  - Reflected/scattered radiation from Sun, surface, atmosphere (UV-MW)
  - Emitted radiation from surface, atmosphere (IR and longer)
- Active: sender and receiver
  - RaDAR (MW and longer)
  - LiDAR (visible)



Atmospheric composition

It is now obvious that a correct description of the interactions between radiation and the atmosphere require knowledge and understanding of the structure and composition of the atmosphere.

The dry atmosphere (no water vapour) is composed of 78% nitrogen, 21% oxygen, 1% of Argon and many other compounds including ozone, water vapour, carbon dioxide, ammonia and methane. The main gas composition of a dry atmosphere is listed in Table 1. The atmosphere consists of permanent gases, such as N_2, O_2, and CO_2, which are generally unreactive and well mixed in the atmosphere (i.e., spread out homogeneously). Non-permanent or reactive gases include CH_4, NO_2 and NH_3. These gases are rapidly removed from the atmosphere by chemical reactions and are variable in time and space. For gases contributing far less than 1% to the volume of the atmosphere, we use the unit **ppm** (parts per million), or even **ppb** (parts per billion). This means that, according to Table 1, for every million molecules of air there are about 1.7 molecules of methane in the atmosphere.

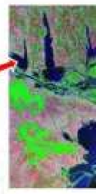
In the following sub-sections, we briefly introduce water vapour and ozone, because they are often required in atmospheric correction methods. Subsequently, we will discuss aerosols and clouds, which also play an important role.

Table 1: Average composition of a dry atmosphere (no water vapour included). Source: National Weather Servicehttps://www.weather.gov/jetstream/atmos_intro

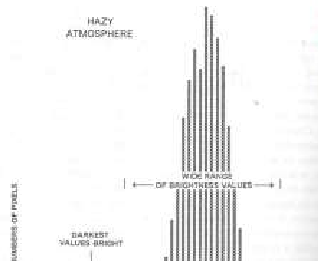
Gas	Symbol	Content
Nitrogen	N_2	78.08%
Oxygen	O_2	20.95%
Argon	Ar	0.93%
Carbon dioxide	CO_2	0.04%
Neon	Ne	18.182 ppm
Helium	He	5.24 ppm
Methane	CH_4	1.70 ppm
Krypton	Kr	1.14 ppm
Hydrogen	H_2	0.53 ppm
Nitrous oxide	N_2O	0.31 ppm
Carbon monoxide	CO	0.10 ppm
Xenon	Xe	0.09 ppm
Ozone	O_3	0.07 ppm
Nitrogen dioxide	NO_2	0.02 ppm
Iodine	I_2	0.01 ppm
Ammonia	NH_3	trace

RELATIVE AC (1)

- Dark object subtraction
- Assume reflectance of dark target due to atmosphere only
- Subtract this value from all reflectances
  - Each channel separately!
- Simple, "cheap" method
- Universal applicability
  - if dark areas are present
- Approximation
- No absolute physical quantity

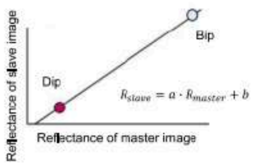


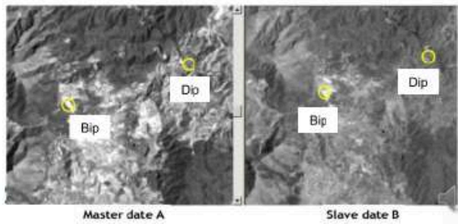
https://www.s... ylmos/image...



RELATIVE AC (2)

- Dark and bright reflective-invariant areas
- In a time series, assume that atmosphere is only cause of changes
- Calibrate slave images to master





- Simple
- Approximate; no absolute physical quantity

Figures from G. Parodi (https://www.slideshare.net/parodign/atmospheric-correction-albuferaweb)

UNIVERSITY OF TWENTE.

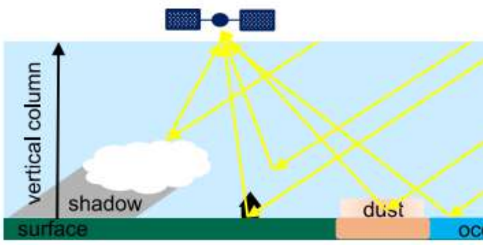
ABSOLUTE AC

- Known target
- Adjustment to radiances detected by calibrated instrument
- Simple, absolute physical quantities, not exact

- Modelling of atmosphere
- Radiative Transfer Model (RTM)
- Exact
  - If input parameters are known
- Expensive

INTERACTION OF RADIATION WITH THE ATMOSPHERE: LIGHT PATHS

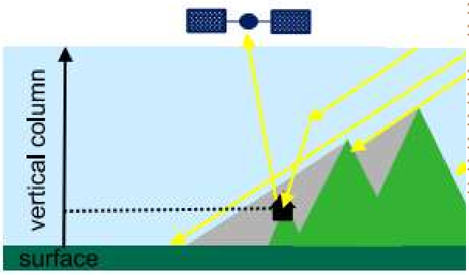
- Attenuation detected by the satellite is the attenuation along the light path… or rather along all light paths
- What effects do clouds, aerosols, surface, atmosphere have on the light paths?
- attenuation
- change of direction



What is the influence of topography on light paths?

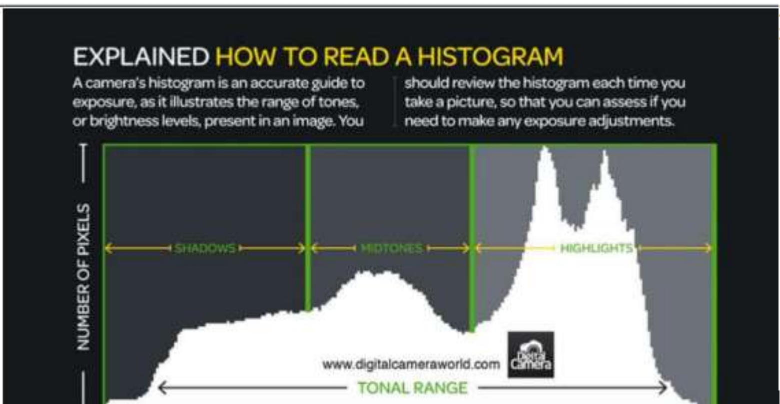
Shadows
- Only indirect radiation on target
- Less information

Less Rayleigh scatter
- Atmosphere most dense near surface
- Adapt AC



HISTOGRAMS

-



In that case, atmospheric effects can be considered as a disturbance of the measurements, for which we need to correct.

1. Kirchhoff's law, means that in thermal equilibrium (object's temperature does not change) absorption and emission are equal.
2. A reflective invariant area- An area of which surface properties do not change over time.
3. A relative atmospheric correction method cannot be used to calculate absolute ground radiance.
4. An absolute atmospheric correction method results in a radiance or reflectance atmospherically corrected image with correct physical units.
5. The absolute correction method usually requires the use of a radiative transfer model.
6. Different atmospheric corrections need to be applied in the visible range and the thermal infrared (TIR) range. This is because atmosphere scatters more radiation in the visible than in the TIR range, sun is the only significant source of radiation in the visible range.
7. Image striping - To correct for image striping, the histograms of the detectors are matched to a selected reference detector, image striping effects may change with time.
8. Components of the atmosphere affect electromagnetic radiation in different ways: absorption, scattering, and emission. Carbon dioxide, Water vapour, ozone are responsible for most absorption in the UV, visible, and infrared wavelength range (0.2 – 20 µm).
9. Oxygen, nitrogen, aerosols three of those compounds cause most scattering of radiation in the same wavelength range. Nitrogen and oxygen are the most abundant gases in the atmosphere and contribute most to scattering. Aerosols come in third in most cases.
10. Red colour sunlight has the highest probability of reaching Earth's surface without being scattered by the atmosphere.
11. About 17percent h is the probability that blue light is scattered by the atmosphere.
12. Invariant pixel correction - The radiances from a single band of two images are linearly correlated.
13. Clouds do not absorb visible radiation, Clouds scatter all colours of visible light with (nearly) equal probability that's why they appear white.
14. The process of scattering is a continuous function of wavelength .
15. The process of absorption is discrete function of wavelength.
16. Scattering occurs mainly at shorter wavelengths.
17. Absorption occurs at specific wavelengths.
18. A particular gas absorbs radiation within one or more wavelength ranges.
19. If the viewing angle of the instrument increases (becomes more slant), the light path length increases and the visible radiation detected by a satellite instrument decreases. An atmospheric correction takes the viewing angle of the instrument into account and corrects for atmospheric absorption and topography. Radiation scattered by the atmosphere into the instrument's field of view may cause blurring of an image.
20. The temperature of the atmosphere increases exponentially with altitude.
21. The pressure of the atmosphere decreases exponentially with altitude.
22. Pressure and temperature are NOT highly correlated in the atmosphere.
23. Starting from the Earth's surface, atmospheric temperature first decreases with altitude, then increases, and decreases again.
24. Oxygen is well-mixed (homogeneously distributed) throughout the atmosphere.
25. The water vapour content of the atmosphere is highest close to the Earth's surface.

Absorption

Attenuation includes both absorption and scattering, as shown in Fig. 4. Gases absorb radiation in a very specific part of the electromagnetic spectrum, related to their molecular and electronic structure. Thus, gases can be recognized by their unique spectral signature. This characteristic **absorption spectrum** often contains spectrally narrow absorption features. This is the basis for the retrieval of atmospheric gas concentrations by absorption spectroscopy, and is used in astrophysics to measure the composition of stars and planets.

Scattering

The process of elastic scattering does not affect the wavelength of the radiation, but may change its direction. The solar beam that reaches the top of the atmosphere has a very specific direction[5]. Scattering changes the unidirectional solar beam into more diffuse radiation. It occurs over a wide spectral range: unlike absorption, it is not limited to narrow wavelengths. Scattering is, nevertheless, wavelength dependent. The wavelength dependence of scattering is determined by the size parameter, which is the relative diameter of the particle (D) compared to the wavelength (λ) of radiation, and is defined as: $\alpha = \pi D/\lambda$.

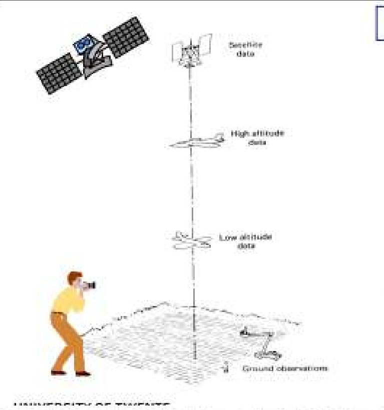
RS Data acquisition

Platform - They are used to carry sensors, platforms can be ground based, airborne and spaceborne.
Platforms can be small(nanosat) can carry one instrument, the larger the sensor more the instruments.

Orbits - Platforms can be put into orbits, Orbits are classified into two categories low earth and hight earth. Every rotation around the earth is called an orbit.
Landsat Mission - first earth observation satellite, 1972 was made by US .

PLATFORMS CARRY SENSORS!!

Sensing from 1 meter to 36,000 km height



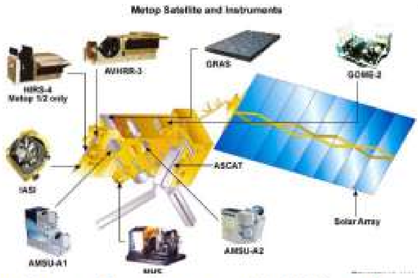
Sensors are mounted on Platforms:
- Ground based
- Airborne
- **Spaceborne**

Here focus on space-borne sensors

SPACE BORNE REMOTE SENSING: MISSIONS

Mission characteristics:
- Platform
- Orbit
- Sensor(s)

Larger satellites carry multiple sensor and instruments



Low earth orbit (LEO) - The altitude is low for this type of satellite, range 180-2000km such as space shuttle, spy satellites. Used for navigation satellite.
Medium Earth orbit (MGeo) - The altitude is between LEO and HEO, range - 2000-3750km, used for navigation.
High Earth Orbit(HEO) Altitude is more than 35750 km. Example - Geostationary satellite.

Geo stationary orbit - It is in high earth orbit, it is at the static position to the earth, its positions at the equator, it rotates with the earth, used for weather. Every image is from the same area, it is parallel to the equator. It acquires image every 15 minute, for fast changing phenomena like weather it is used. It rotates with the earth meaning it focus on the same part of earth, depending on the altitude the image will get coarse. It is used for global analysis. It has spatial resolution.

ORBIT CHARACTERISTIC: ALTITUDE

- **Low earth orbits (LEO)**
  - Altitude 180-2000 km; e.g. spy satellites, space shuttle and most satellites for EO
- Medium Earth Orbit (MEO)
  - Altitude 2000-35750 km; e.g. Navigation
- High Earth Orbit (HEO)
  - Altitude >35750 km; e.g. geostationary

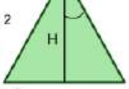
Polar orbiting satellites are closer than 35750 km of altitude

GEO STATIONARY ORBITS

- Relative position to Earth is fixed
- Hemispherical distorted view
- Easy geo-referencing
- Coarse resolution due to distance. (Getting better by reducing IFOV)
- Meteorological weather systems use a combination of polar (HR) and geostationary (LR) satellites.

ORBIT AND FIELD OF VIEW (FOV) SATELLITE

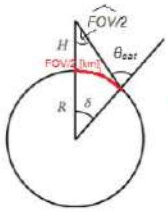
- Satellite field of view (FOV) and height determine how much a satellite can see.
- For quantitative remote sensing there is limitation of satellite zenith angle θ_{sat} ($\theta_{sat\_max} = 60°$)
- Thus for geostationary satellite maximum FOV = 11,800 km
- For narrow FOV/flat terrain: FOV [km] = tan(FOV/2) x H x 2



Period - it is referred as time it takes for a satellite to complete one rotation.
Repeat cycle - time taken by the satellite to complete two successive identical orbit
Revisit time - time when a satellite visits the same area.

For wide FOV compensate for earth curvature : FOV [km] =2 R [km] (δ° π/180°).

For AVHRR with $\theta_{sat\_max} = 60°$ and H = 840km the FOV = 2227 km



Field of view - is the area that is acquired by the satellite. It angle of the view can be increased, it should not increase more than 60 degrees.
Inclination angle - It is the angle between orbital plane and equator. For geo stationary the inclination angle is 0 and for polar it is between 80 to 100 degrees. It determines which parts of the angle to take measurements from.

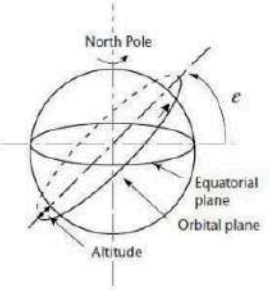
ORBIT CHARACTERISTIC: ALTITUDE VERSUS VELOCITY

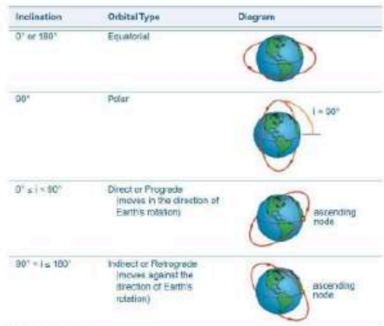
- Satellites travel fast, at a height of 35786 km the forward movement of the satellite, from west to east direction and in equatorial plane is in sync with the earth rotation
- The higher the satellite the lower the velocity
- SunSync EO satellites orbiting at about 700 km travel at 7 km/s (approx).

Implications?

ORBIT CHAR: INCLINATION ANGLE

- Angle between orbital plane and equator
- Determines which latitudes are measured, e.g. Jason-3 at 66° and GPM at 65°
- Polar orbit inclination between 80° and 100°. Implications?



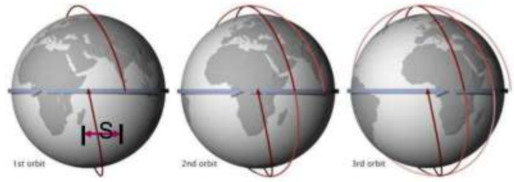


ORBIT CHAR: INCLINATION ANGLE

- Special polar orbit: Sun Synchronous Orbit (SSO) , inclination about 98,5°
- A nearly polar orbit around the Earth, in which the satellite passes over any given point of the planet's surface at the same local mean solar time (mostly between 09:30 to 10:30 am)

ORBIT CHAR: ORBIT STEP

- Orbit step (S) is next overpass at equator upon completion of another orbit (distance)
- Orbit step, together with FOV – zenith angle, determine how long it takes to get global coverage

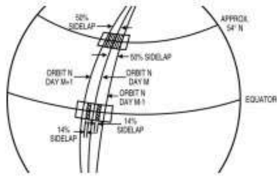


- In SSO, at 800 km, swath with of 2800 km, slightly more than 14 orbits: daily global coverage

ORBIT CHAR: PERIOD & REPEAT CYCLE

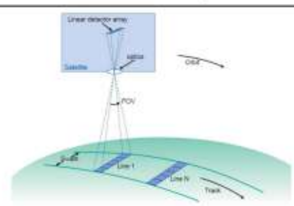
- Period: Time to complete one full orbit (360 degrees rotation)
- Repeat cycle: Time between two successive identical orbits
- Revisit time: Time between two subsequent images of the same area

https://www.youtube.com/watch?v=y_jM_BxQGvE

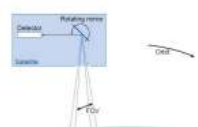


SENSOR CLASSIFICATION (HOW THEY OPERATE)

- Cameras (many 'pixels' at a time)
  - CCD (detector array)
    - Line (camera or push broom)
    - Array (camera)



- Scanners (single 'pixel' at a time)
  - Whiskbroom (ind. detectors)
  - Multiple sensors/bands



Sensors operation - Cameras takes many pixels at a time, CCD (detector array) Camera, line camera, Array camera. Scanners that take single pictures at a time. Fewer pixels are measures pixels by pixels are measured. Whiskbroom it's like a pendulum it moves and continually takes pictures. Multiple sensors/bands. Camera is preferred more than scanners.

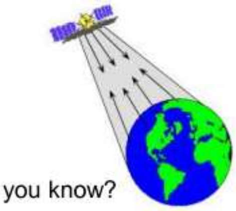
SENSOR CLASSIFICATION BASE ON SOURCE

Passive Sensors

- Source of radiation is external
  - Solar reflected radiation (SW)
  - Emission from Earth (LW)
  - Emission from Atmosphere (LW)

Active Sensors

- Emits radiation towards a target. Partly is reflected to the sensor
  - Radar
  - Lidar

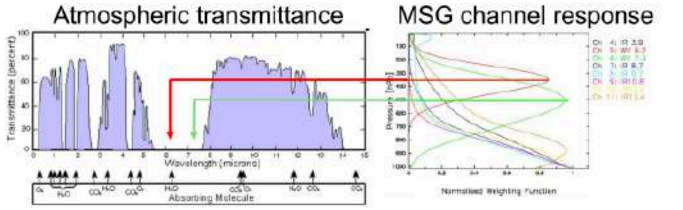


Any other you know?

Passive sensors - use solar reflected radiation, emission from the earth and emission from the atmosphere.
Active sensors - Has its own power, emits radiation towards a target. Partly is reflected to the sensors. Like radar and lidar.
Source of sensors - for measuring on the earth surface what is in the atmosphere should be known, so for that we need multiple sensors on the platform.

SENSOR CLASSIFICATION (WHAT SOURCE) VIP

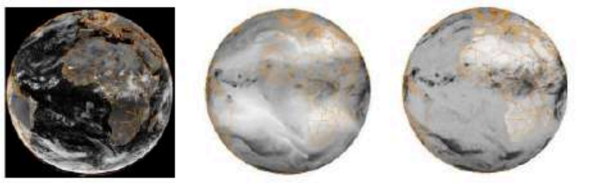
1. Surface sensing with no contribution from atmosphere (passive). We are interested in the surface
2. Atmospheric sensing, when there is no contribution from the surface (passive). We are interested in the atmosphere.
3. Sensing when there is no contribution from surface emission or atmosphere (active), from 2 cm – 10 m wavelength. Active RS do not require other source.



Weighting function describes the width and the weight of layer of the atmosphere from which the radiation measured by a particular satellite channel was emitted (atmospheric sensing or surface sensing)

SENSOR / INSTRUMENT – IMAGE CHARACTERISTICS VIP

Resolutions/Characteristics:	Example MSG:
▪ Spectral	11 spectral channel and 1 HRV
▪ Spatial	3 km sub-satellite and 1 km sub satellite
▪ Temporal	every 15 minutes
▪ Radiometric	10 bits

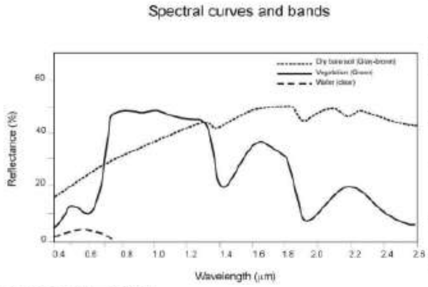


MSG Vis 0.8 WV 06.2 TIR 10.8 micron

SPECTRAL CHARACTERISTIC

Answers to…

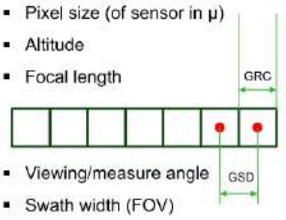
- How many bands?
- How narrow/wide?
- Location/range of bands in EM spectrum



SPATIAL CHARACTERISTICS - RESOLUTION

Camera

- Pixel size (of sensor in μ)
- Altitude
- Focal length



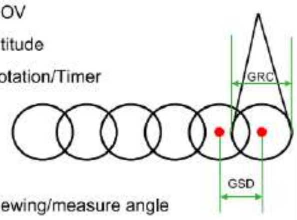
- Viewing/measure angle
- Swath width (FOV)
- Ground Resolution Cell GRC

= Ground Sampling Distance GSD

UNIVERSITY OF TWENTE.

Scanner

- IFOV
- Altitude
- Rotation/Timer



- Viewing/measure angle
- Swath width (FOV)
- GRC <> GSD

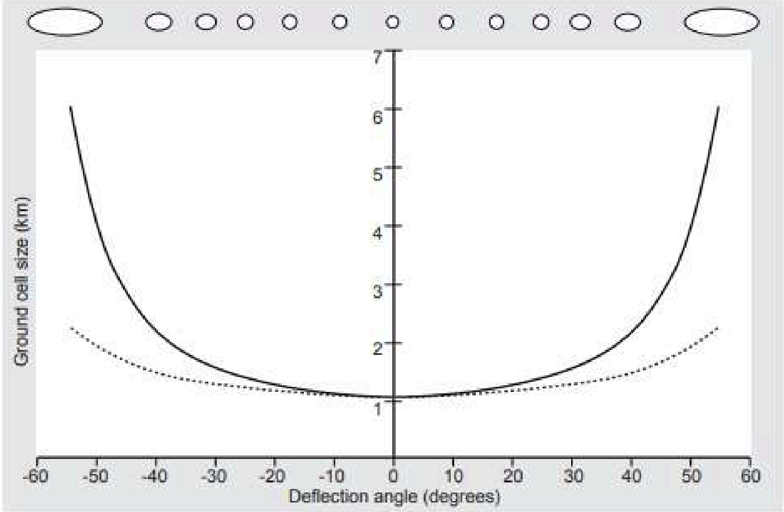
Spectral characteristics - a combinations of factors like bands, narrow or wide of spectral sensitivities.
Spatial characteristics - Camera and the scanner have different spatial characterises, in camera the ground size is square and in scanner the ground resolution size is circular. Camera is superior to the scanner. It has pixel size.
Radiometric characteristics - it is a continues scale that has to be stored and hence bits are used (1-8bits), the higher the bit the better the radiometric characteristics.
Temporal characteristics - Higher the altitude of the satellites higher the period, vice versa. Satellite collects the data and send it to the ground stations (receiving station) and they process it, ground stations are provided multiple areas on the path.

Geometric aspects

At a particular instant, the detector of an across-track scanner observes an elliptical area on the ground, the ground resolution cell of the scanner. At nadir, the cell is circular, of diameter D. D depends on the IFOV, β, of the detector and the flying height.

$$D = \beta H \tag{4.2}$$

A scanner with $\beta = 2.5$ mrad operated at $H = 4000$ m would have, therefore, a ground resolution of 10 m at nadir. Towards the edge of a swath, the ground resolution cell becomes elongated and bigger (Figure 4.29).



The width of the area that is covered by one sweep of the mirror, the *swath width*, depends on the FOV of the scanner. AVHRR has a very wide FOV of 110°; easy geometry was not a concern in the AVHRR design. Landast-7 has an FOV of only 15°, hence geometrically more homogeneous images result.

RADIOMETRIC CHARACTERISTICS

A sensor measures radiance in a range (dynamic range).
Too little = no measurement
Too much = saturation
Dynamic range: [Minimum, Maximum] intensity watt/m²/sr

Within the range the sensor has a precision… 1, 2.. Decimals
With the dynamic range and the precision, a quantization is defined. The higher the more precise.

AD conversion and quantization
Radiometric resolution [min, max] DN, #bit (~precision)
Effective radiometric resolution is combination of both

UNIVERSITY OF TWENTE.

34.1 <=Range <= 42.0
I need to measure 79 values.
My instrument is unable to read more than with 0.1°C precision.

The real measurement is 39.62 °C

1 bit= 2^1= 2
2 bits= 2^2= 4
3 bits= 2^3= 8
4 bits= 2^4= 16
5 bits= 2^5= 32
6 bits= 2^6= 64
7 bits= 2^7= 128
8 bits= 2^8= 256
…
16 bits= 2^{16}= 65536

Quantization to 7
Sensor records 39.6 because is unable to measure 39.62

-61 <=Range <= 55
I need to measure 116 values because my instrument can read with a precision of 1°C.

The real measurement is 25.7 °C

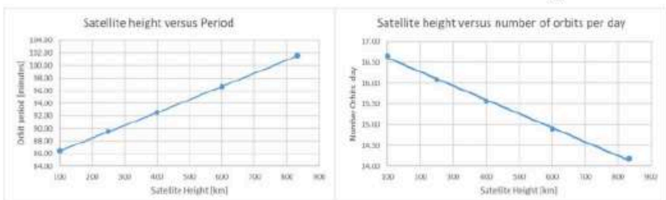
1 bit= 2^1= 2
2 bits= 2^2= 4
3 bits= 2^3= 8
4 bits= 2^4= 16
5 bits= 2^5= 32
6 bits= 2^6= 64
7 bits= 2^7= 128
8 bits= 2^8= 256
…
16 bits= 2^{16}= 65536

Quantization to 7
Sensor records 26 because is unable to measure 25.7

TEMPORAL CHARACTERISTICS

- Repeat cycle; time between two identical overpasses
- Revisit time; time between two acquisitions of the same area

Period and number of orbits as function of satellite height



(REAL TIME) DATA PROVISION

From satellite observation to availability "on the ground" requires ground stations, eventually by a Data Relay System and further data pre-processing by satellite data provider

GIS Data acquisition

Primary data - it is directly captured from the site. Such as, terrestrial data, field surveys, ariel survey and satellite remote sensing.
Indirect Spatial data -

Geoportal play a key role in access of spatial data.
Dutch notational geoportal is a geoportal available in the Netherlands.

Volunteered geographic information (VGI) - it is a crowed source data meaning people put information in it, it is free to use. Like, Wikipedia, OpenStreetMap, Instagram.

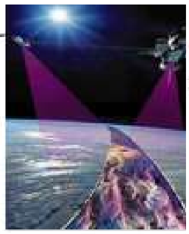
Digitizing is tracing features from maps photographs, they are present dots, lines and polygons. Techniques are depended on the input you want to put, if its complex then manual digitizing is used whereas for simpler we can use automatic.
Types -

1. On tablet digitizing - it is discontinued.
2. On-screen digitizing it is manual or semi-automatic.
3. Automatic scan digitizing.

Vector file is needed before digitizing. Geo package stores points, lines and polygon data in a relational database as tables.
Modes in digitizing -
Point mode - you have to manually digitize all the coordinate, its time consuming.
Stream mode - the coordinates will be put according to the curve.

Scanners - for high quality scan drum and flatbed scanner is used, for large documents fed scanner is used. Resolution depends on digitizing if its printed map then 200-300 dpi is sufficient for aerial photograph more or equal to 800 epi.
Vectorization - Transformation from raster to vector lines.

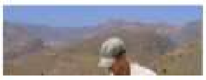
DIRECT SPATIAL DATA ACQUISITION



- Data directly captured from the environment
- Primary data

Terrestrial survey

Field surveys



Aerial surveys and satellite remote sensing

INDIRECT SPATIAL DATA ACQUISITION

- Data derived from existing sources
- Data that have been collected for other purposes
- Secondary data

Geoportals

Digitizing

GEOPORTALS

- Spatial data can be obtained via a Geoportal. They enable the discovery and delivery of data through a service-based architecture known as Spatial Data Infrastructure[1].

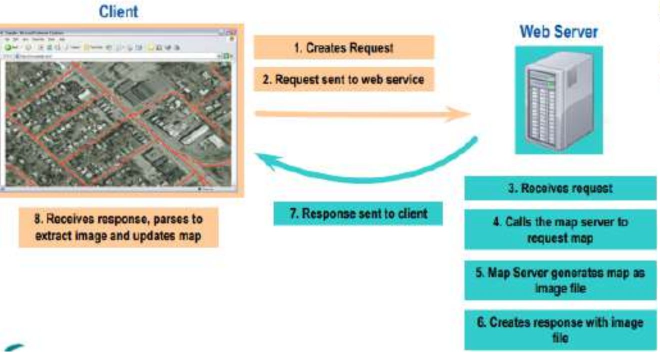
GIS WEB SERVICES

- A Web Map Service (WMS) is a web server that provides access to raster data, a Web Feature Service (WFS) is a web server that provides access to vector (GML) data, and a Web Coverage Service (WCS) is a web server that provides access to imagery.

- These web service standards for spatial data were developed by the Open Geospatial Consortium (OGC). Some of the most recognized services in this context are WMS, WFS, and WCS

https://www.ogc.org

WEB MAP SERVICE OPERATION



VOLUNTEERED GEOGRAPHIC INFORMATION (VGI)
CROWDSOURCED DATA COLLECTION

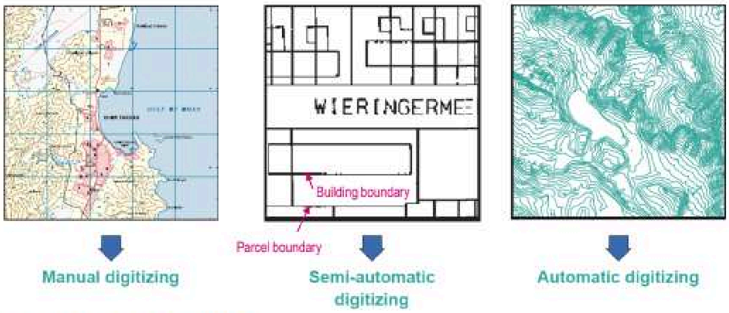
- Citizens can play a complementary role by providing geo-referenced information known as Volunteered Geographic Information (VGI).
- Products such Wikimapia and OpenStreetMap (OSM) can be considered as VGI. Citizens can directly participate in the collection of spatial data. It is made by people like you and me.

A typical example of VGI is when users add their locations and information about themselves on a map. Therefore, sites such as TripAdvisor, Flickr, Twitter and Instagram can also be considered as VGI.

Adding a building to OpenStreetMap

SELECTING A DIGITIZING METHOD

- Manual digitizing > complex documents / interpretations from satellite imagery or aerial photographs
- Semi-automatic digitizing > simple documents that require some interpretation
- Automatic digitizing > simple documents with one type of information



VECTOR DATA FORMAT
SHAPEFILE

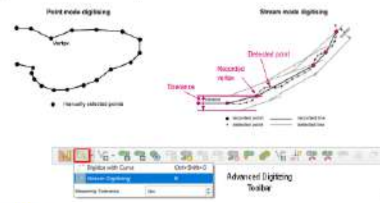
- A Shapefile is vector data structure for storing spatial and attribute data
  - It contain only one geometry type (either point, line or polygon)
  - It consists of a minimum of three files: *shapefile.shp*, *shapefile.shx* and *shapefile.dbf*
  - It is a non-topological data structure (polygon-by-polygon)

VECTOR DATA FORMAT
GEOPACKAGE

- A geopackage stores point, line, and polygon data in a relational database as tables.

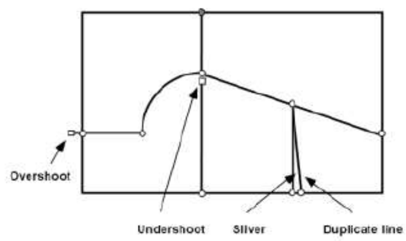
DIGITIZING MODES
POINT MODE AND STREAM MODE

- When you trace a line feature you can do this in point mode or stream mode. In point mode you digitize each point of the line. In stream mode you follow the line with your cursor and the vertices are automatically recorded depending on a preset tolerance from a straight line.



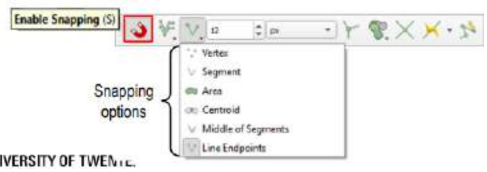
DIGITIZING TIPS

- Use snapping to avoid overshoots and undershoots
- Avoid duplicate lines (adjoining polygons)



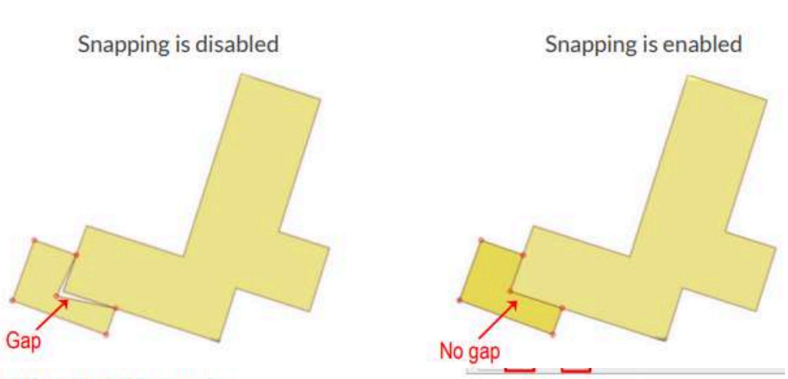
DIGITIZING TIPS
DIGITIZING AREAS

- In certain cases it is more convenient to digitize area objects as lines and thereafter use an automatic procedure to construct the polygons.



IIVERSITY OF TWENTE.

DIGITIZING TIPS
CREATING ADJOINING POLYGONS IN QGIS



DIGITIZING TIPS
CREATING ADJOINING POLYGONS IN QGIS

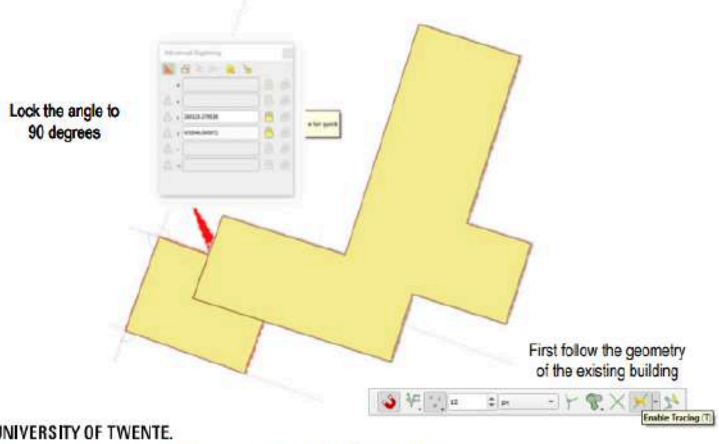
- With Automatic Tracing, you no longer need to manually place all the vertices during digitization. The resulting geometry exactly follow an existing geometry. It prevents the creation of gaps.

DIGITIZING TIPS
CREATING ADJOINING POLYGONS IN QGIS

- You can also use the Avoid Overlap on Active Layer tool to digitize adjoining polygons. The geometry of the existing polygon is used to complete the second polygon.

DIGITIZING TIPS
CREATING RECTANGULAR POLYGONS IN QGIS

- On the Advanced Digitizing panel you can lock angles while drawing. This option can be used to create rectangular polygons.



UNIVERSITY OF TWENTE.

ENTER ATTRIBUTE DATA IN QGIS
USING THE ATTRIBUTE TABLE

- It is a good practice to enter and edit attribute information immediately after digitizing a feature. Sometimes you need to add remaining attribute information afterwards.

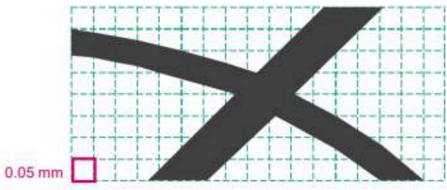
USING THE FIELD CALCULATOR

- The field calculator allows you to calculate attribute values on the basis of existing attributes values or defined functions, for instance, to calculate the area of geometry features. The results can be written to a new **field**.

SCANNING
MODES AND RESOLUTION

Scanning resolution:

- millimeters (e.g. 0.05 mm)
- microns (e.g. 50 µm)
- dots per inch (e.g. 500 dpi)
  nr. of pixels per inch (2.54 cm)



Original with scanner resolution

Scanning modes:

- Bitmap (black=1 white=0)
- Grey-scale (256 grey values per pixel)
- Color (Red, Green, Blue)

Scan: grey scale pixel values

SCANNING
SELECTING SCANNER RESOLUTION

Manual digitizing:

- Printed map: 200-300 dpi
- Aerial photographs: ≥ 800 dpi

Semi-automatic / automatic digitizing:

- Printed map: 300-600 dpi

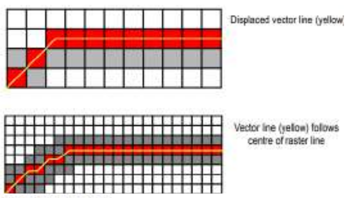
(depending on the thickness of the thinnest lines)

Map scanned with 100 dpi (enlarged 4 times)

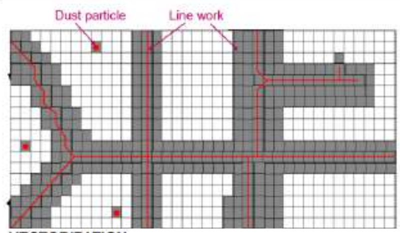
Map scanned with 300 dpi (enlarged 4 times)

UNIVERSITY OF TWENTE.

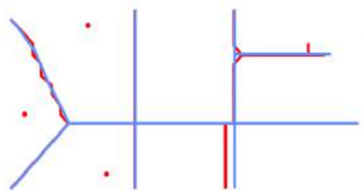
VECTORIZATION
DISPLACEMENT OF VECTOR LINES



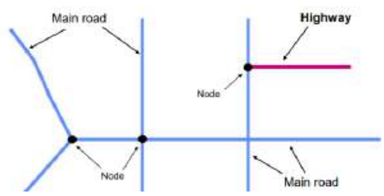
VECTORIZATION
PROCESS: SKELETONIZING



VECTORIZATION
PROCESS: POST-PROCESSING



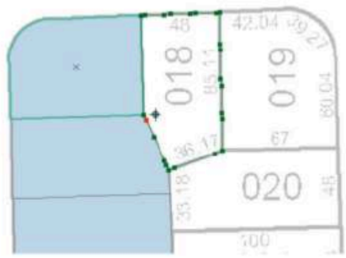
VECTORIZATION
PROCESS: FEATURE FORMING



SEMI-AUTOMATIC DIGITIZING
RASTER TRACING

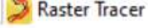
- Semi-automatic digitizing, also known as raster tracing, involves the process of vectorizing line and polygon features through a semi-automated approach, where line features are traced partially using automation.

Raster tracing used to create line features.



Raster tracing used to create polygon features.

SEMI-AUTOMATIC DIGITIZING
RASTER TRACER

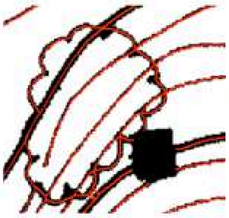


- In QGIS, there is a plugin called Raster Tracer. The Raster Tracer can be used for semi-automatic digitizing of an underlying raster layer.

AUTOMATIC DIGITIZING

- Automatic digitizing is the process of vectorizing line and polygon features through a completely automated approach, wherein line fea are traced entirely using automation.

Automatics vectorization of contour lines

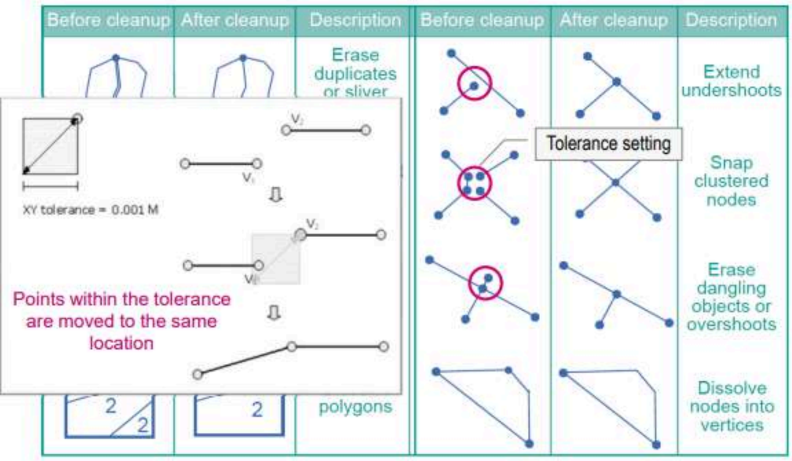


overlapping and intersecting features

SHAPE RECOGNITION

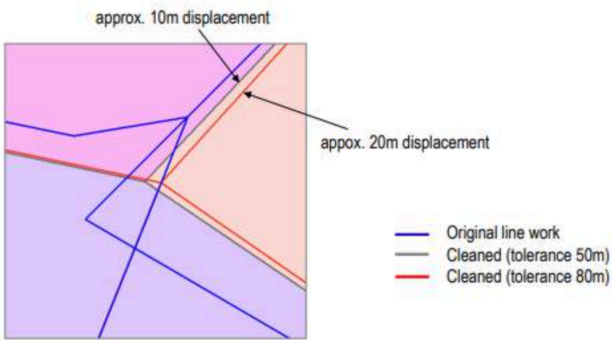
- You can also capture vector features of a certain shape, such as a building or storage tank, via the shape recognition tools.

CLEAN-UP OPERATIONS FOR VECTOR DATA



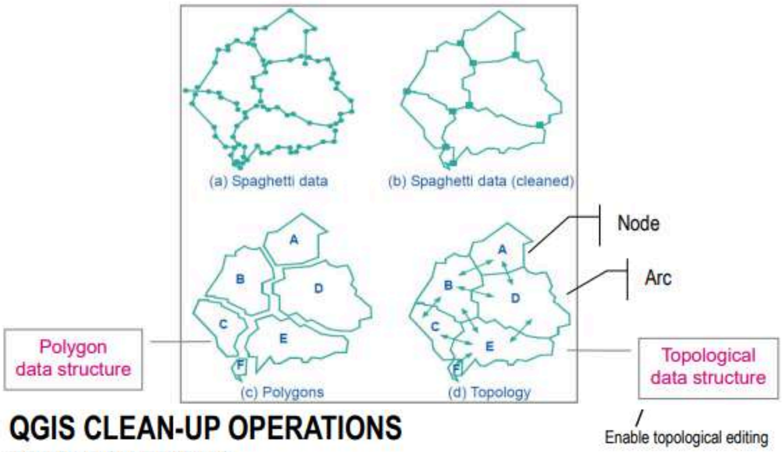
TOLERANCE SETTING

- Settings such as a tolerance for the cleaning of overshooots, undershoots or node clusters have to be set carefully to avoid that lines displace, do not connect or connect where they should not connect.



CLEAN-UP OPERATIONS FOR AREAS

- The topological data structure can aid in data editing area features.



QGIS CLEAN-UP OPERATIONS
TOPOLOGY EDITING

- Topological editing in QGIS can be used to edit the shared edges of two adjacent polygons

Standard editing

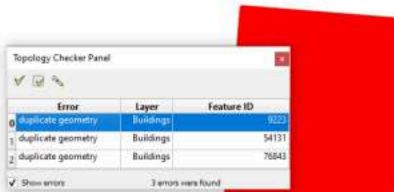
Topology editing

QGIS CLEAN-UP OPERATIONS
TOPOLOGY EDITING

- The Topology Checker plugin can be used to check the geometry validity by evaluating a set of topology rules. The identified errors require manual editing.

1. Define topology rules

2. Errors found are colored red



QGIS CLEAN-UP OPERATIONS
TOPOLOGY EDITING

- The Geometry Checker plugin can also be used to validate the geometry of data layers. Any identified errors can be automatically corrected.

Spatial Reference Systems

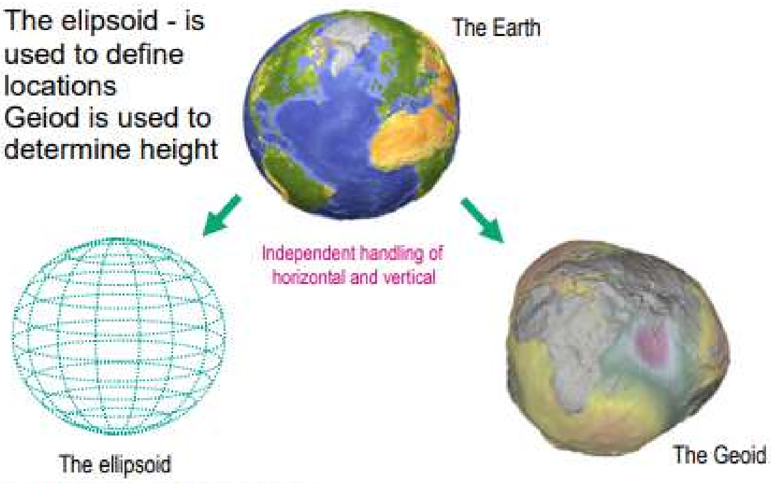
The ellipsoid - It is used to define location. (h) is the height by ellipsoid. It is defined by pole and equatorial plane. Major axis and minor axis are used.

The geoid - it is used to determine heights. The shape of geoid is only affected by the gravity. The sea level is observed over the years and then its average is considered as mean sea level. In the Netherlands, mean sea level is established in Amsterdam and the heights are levelled according to that.

Geodetic level - mean sea level is used via the help of levelling instrument, height difference is taken in order to compute the levels the height difference is called as orthometric height. Orthometric height(H) is the hight above the sea level. For each country there is a certain sea level.

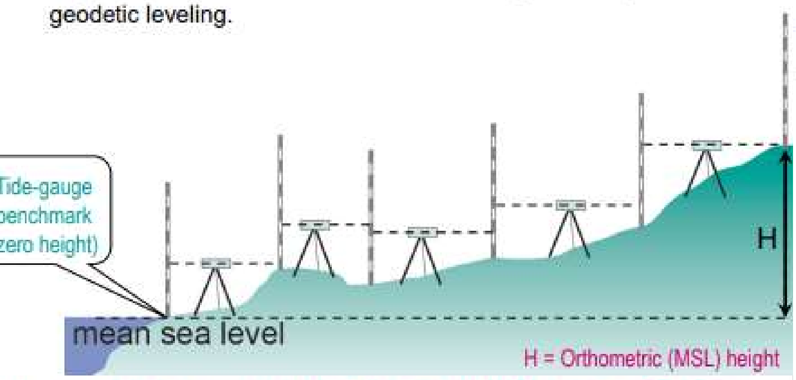
REFERENCE SURFACES FOR MAPPING

The elipsoid - is used to define locations
Geiod is used to determine height



GEODETIC LEVELLING

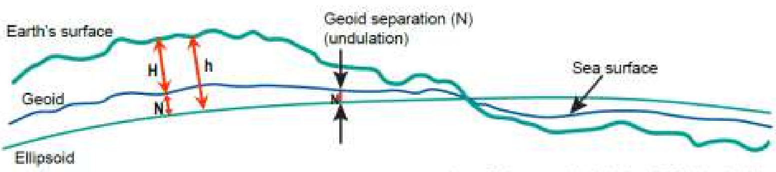
- Starting from Mean Sea Level (MSL) points, the orthometric heights (H of points on the Earth can be measured using a technique known as geodetic leveling.



H_{NAP} = 107.5m (h_{WGS84}) - 43.6m (N) - 26.4m (ITC building) = 37.5m

Ellipsoid height | Geoid undulation(N) | Building height | Orthometric (MSL) height

RELATION BETWEEN GEOID AND ELLIPSOID

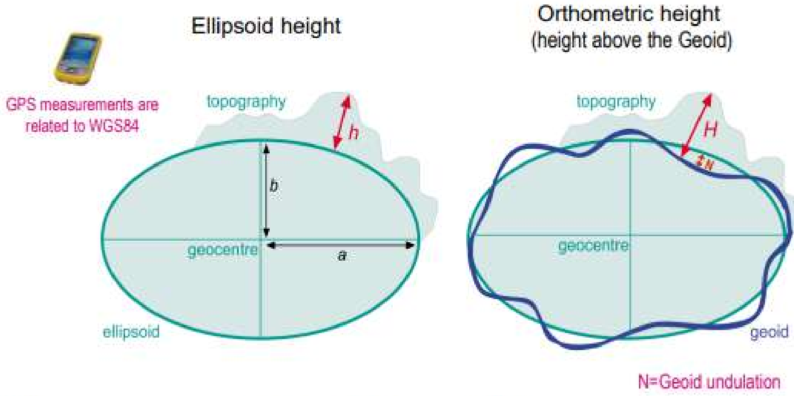


H = Orthometric height (MSL height)
h = Ellipsoidal height
N = Geoid undulation (separation)

The Geoid will rise above or below the ellipsoid. The differences (geoid undulation (N)) globally vary between +/- 110 meters.

The surface of the Earth is far from uniform. Its oceans can be treated as reasonably uniform, but the surface or topography of its land masses exhibits large vertical variations between mountains and valleys. These variations make it impossible to approximate the shape of the Earth with any reasonably simple mathematical model. Consequently, two main reference surfaces have been established to approximate the shape of the Earth : one is called the *Geoid*, the other the *ellipsoid*; see Figure 3.1.

ELLIPSOID HEIGHT VERSUS ORTHOMETRIC HEIGHT



Datum - mean sea level a country uses is known as height datum.
Geoid undulation (N) - The difference between mean sea level and ellipsoid height. The difference is between +/- 110 meters.

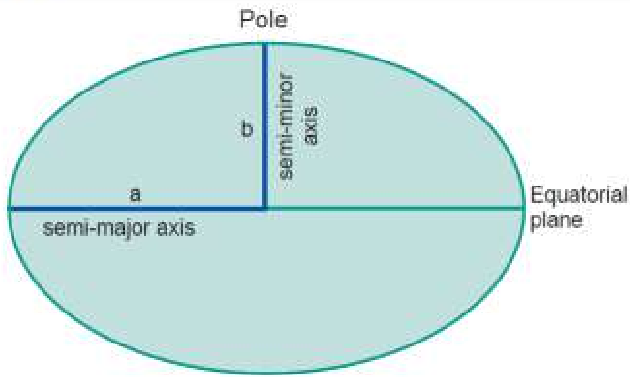
Horizontal datum systems(geodetic datum system)- Locally defined ellipsoid are used to define accurate maps of a country or a region as they fit mush better to geoid, globally defined ellipsoid are on a global level like in a GPS.

Geographic coordinate system - latitude angle and longitude angle define coordinated. The angle is taken considering the equator. Ellipsoid height is also taken into consideration.
WGS84 is globally used datum system.
Local datum system - Survey points like a network is used to define local datum system. Local systems are stable.
International datum system - GPS stations make a network that define global system. Since the plates move the stations are relocated hence the measurement is taken quite frequently.

TRENDS IN MAPPING: GLOBAL VERTICAL DATUMS

- Global height datums (e.g. EGM2008, GGM02) can be determined with centimetres accuracy by satellites (e.g. GOCE, GRACE) that measure the earth's gravity.

THE ELLIPSOID



Typical values of the parameters for an ellipsoid:

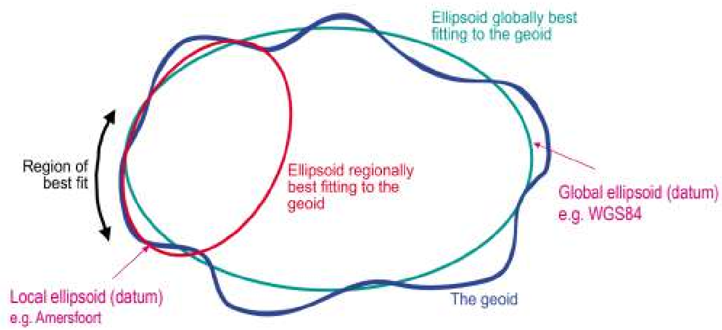
a = 6378137.0m b = 6356752.31m

f = 1/298.26 e = 0.0818187

Flattening: $f = (a-b)/a$ Eccentricity: $e^2 = (a^2 - b^2)/a^2$

HORIZONTAL DATUM SYSTEMS

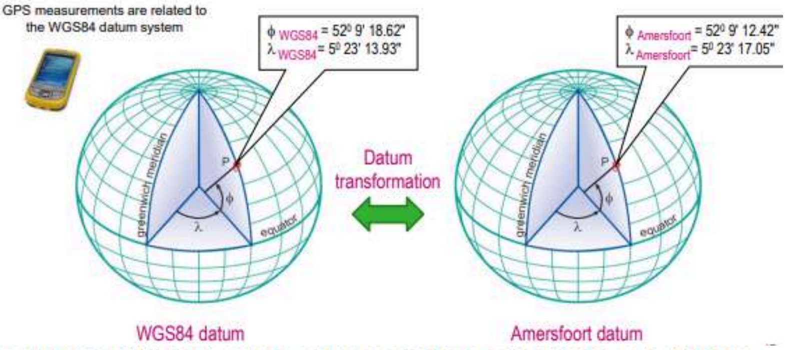
- There are many different ellipsoids defined. Each of them may have a different position and orientation to fit a particular region on earth (often a country). They are called horizontal datums (or geodetic datums).



Geo graphic coordinated are used to define location on the ellipsoic
Latitude and longitude are used to define location.

GEOGRAPHIC COORDINATE SYSTEMS

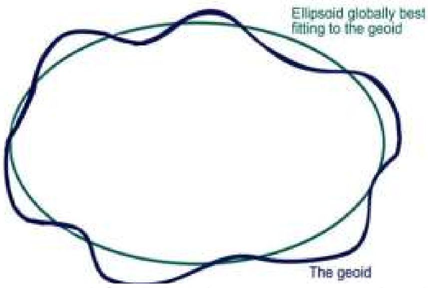
- The differences in geographic coordinates of one location can be up to several hundreds of meters (several seconds) depending on the used datum system.



TRENDS IN MAPPING: GLOBAL HORIZONTAL DATUMS

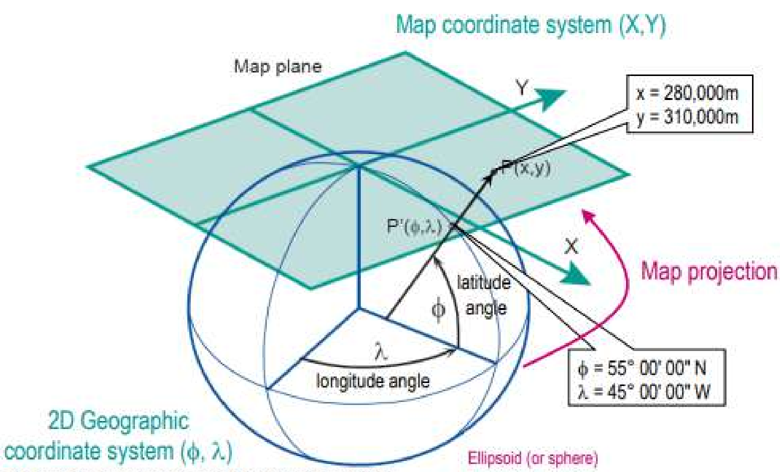
- Global ellipsoids and datums to approximate the earth-as-a-whole -with the aid of satellites- are becoming more in use (WGS84, ITRF2008, ETRS89, and others).
- Many countries use global datum systems to align and refine their local datum systems, making them suitable for data exchange and various geospatial applications.

Countries who have their own system used global to refine theirs system.



Map projection - once the datum system is defined, then the map projection takes place. Transform the geographic coordinates to the map. Each projection have their own distortion characteristics.
Scale distortions on maps - There are always distortions since earth is curvilinear and hence taking it from 3D to 2D there are distortion, there is more distortions near the poles but not so much near to the equator.
How to select the projection - cylindrical are used to mapping the world, conical are used for mid latitudes since it does not have a lot of distortion, azimuthal projection is used to map near the poles and region, Netherlands use azimuthal. For rectangular the best projection used is cylindrical.
EPSG Code - it defines map coordinates system for Dutch it is 28992.
Mapping equation - each projection has its own mapping equation. It is used for going from one coordinate system to other coordinated system.
Control points are used to relate one layer to the other.

MAP PROJECTION PRINCIPLE



The equations for the Mercator projection (spherical assumption)

Forward mapping equation:

$$x = \underbrace{R}_{\text{Earth radius}}(\underbrace{\lambda}_{\text{Longitude angle}} - \underbrace{\lambda_0}_{\text{Central Meridian}})$$

$$y = R(ln(tan(\frac{\pi}{4} + \frac{\overbrace{\phi}^{\text{Latitude angle}}}{2})))$$

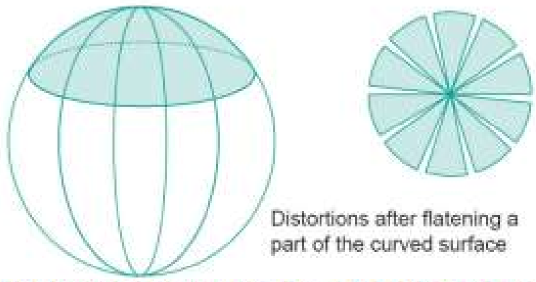
Inverse mapping equation:

$$\phi = \frac{\pi}{2} - 2\arctan(e^{\frac{-\overbrace{y}^{\text{Y-coordinate}}}{R}})$$

$$\lambda = \frac{\overbrace{x}^{\text{X-coordinate}}}{R} + \lambda_0$$

SCALE DISTORTIONS ON MAPS

- Any map projection is associated with scale distortions. The amount and which kind of distortions a map will have depends largely - next to size of the area being mapped - on the type of the map projection that has been selected.



Distortions after flatening a part of the curved surface

DISTORTION PROPERTIES OF MAP PROJECTIONS

- Conformal

Angles (with short sides) and shapes (of small areas) are shown correctly on the map.

- Equivalent (or equal-area)

Areas are correctly represented on the map.

- Equidistant

Distances from 1 or 2 points or along certain lines are correctly represented on the map.

SELECTION OF A SUITABLE DISTORTION PROPERTY

- Conformal property

Maps which require measuring angles (e.g. aeronautical charts or topographic maps)

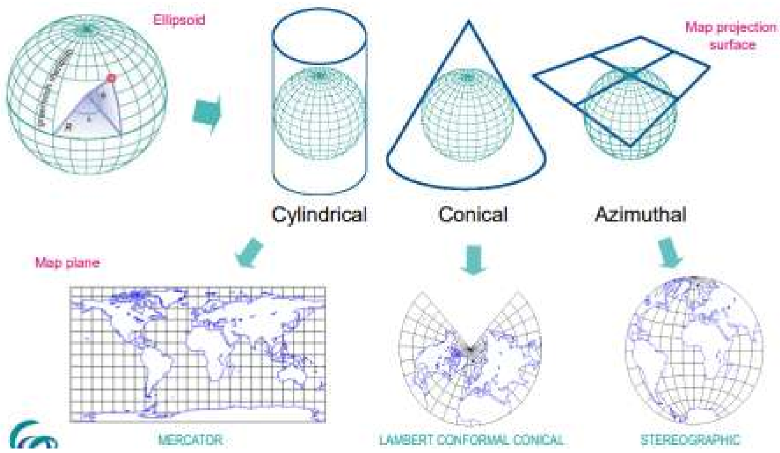
- Equivalent (or equal-area) property

Maps which require measuring areas (e.g. thematic or distribution maps)

- Equidistant property

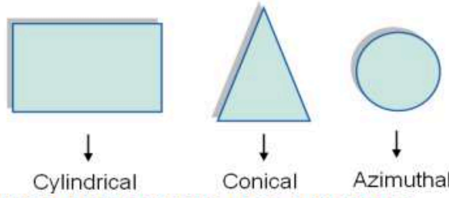
Maps which require reasonable area and angle distortions (e.g. thematic or presentation maps)

CLASSES OF MAP PROJECTIONS



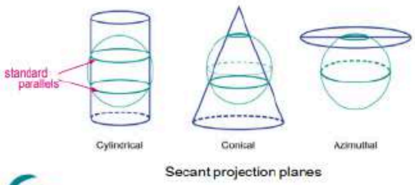
SELECTION OF A SUITABLE PROJECTION CLASS

- Normal cylindrical projections are typically used to map the World in its entirety. Conical projections are often used to map the different continents and mid-latitudes, whereas the normal azimuthal projection may be used to map the polar areas.
- Also consider the shape of the area to be mapped:

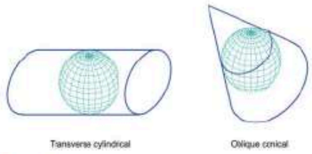


CLASSIFICATION OF MAP PROJECTIONS

- Class
  - **Cylindrical**
  - **Conical**
  - **Azimuthal**
- Secant or tangent projection plane

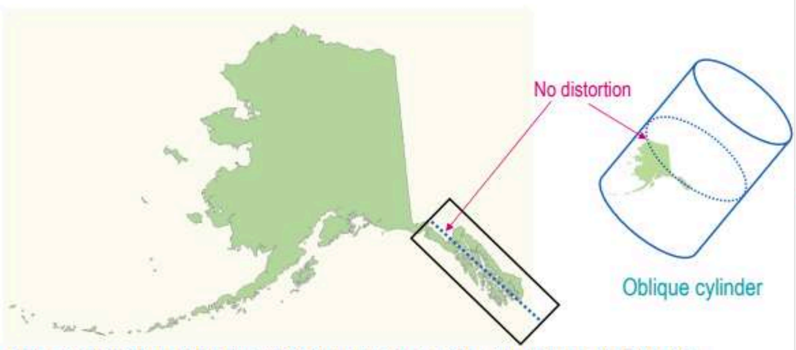


- Property
  - **Equivalent (or equal-area)**
  - **Equidistant**
  - **Conformal**
  - ***Minimum-error***
- Aspect (orientation)
  - **Normal**
  - **Oblique**
  - **Transverse**

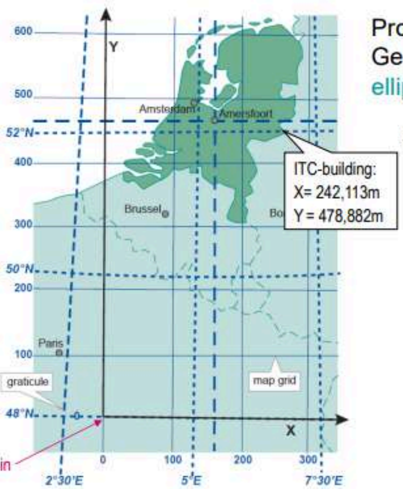


SELECTION OF A SUITABLE ASPECT

- The position (and orientation) of the projection plane is optimal when the projection plane is located along the main axis of the area to be mapped or when the projection centre coincides with centre of the area.

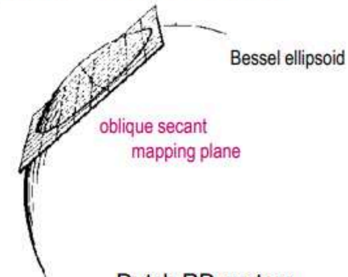


MAP COORDINATE SYSTEM OF THE NETHERLANDS



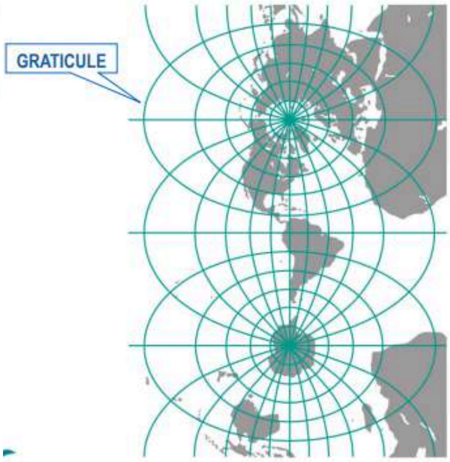
Projection: Stereographic
Geodetic datum: Amersfoort (Bessel ellipsoid)

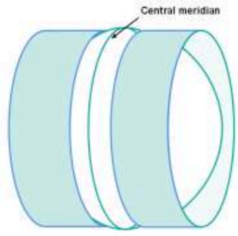
System used in the Netherlands



Dutch RD system
(RijksDriehoekstelsel)

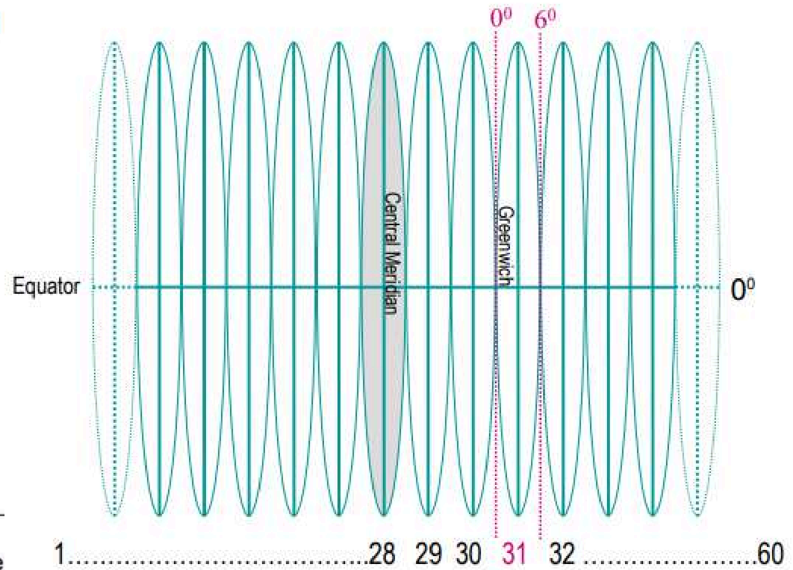
TRANSVERSE MERCATOR PROJECTION



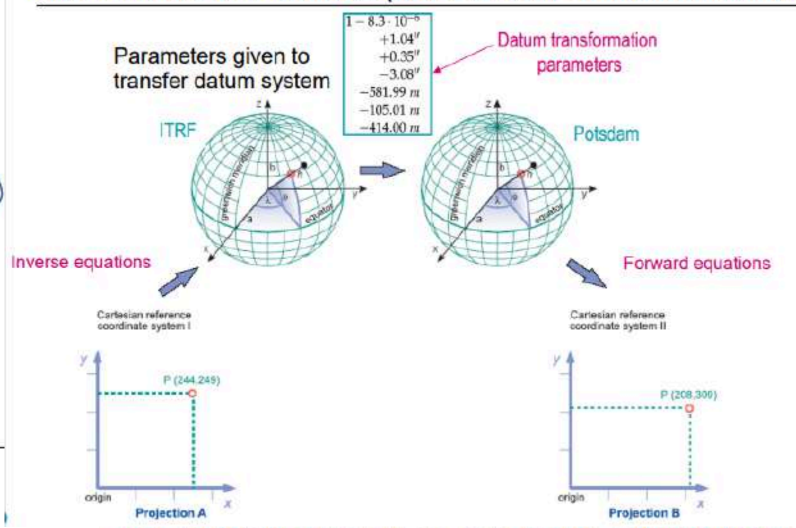


Conformal Cylindrical (transverse secant) projection

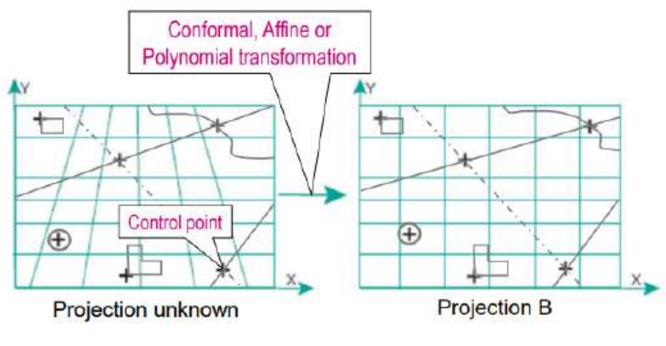
UTM GRID ZONE NUMBERING SYSTEM



MAP PROJECTION CHANGE (INCL. DATUM TRANSFORMATION



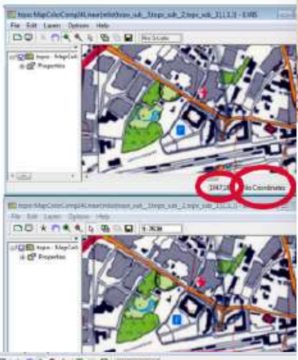
PROJECTION CHANGE VIA 2D CARTESIAN TRANSFORMATIO



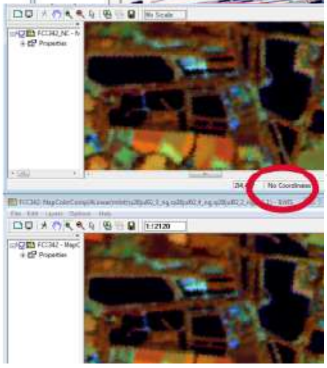
The unknown coordinate system is related to a known coordinate system based on a set of known points

Image re-projection and resampling

GEOREFERENCING TOPOGRAPHIC MAPS: WHAT'S DIFFERE



- These are both scanned maps
- Both images are identical in geometry
- Upper: not georeferenced
  - Pixel size, location is unknown
- Lower: with georeference
  - Pixel size and location is known
- Upper only having image coordinates.
- Lower having image coordinates, Cartesian coordinates and global coordinates.
- Georeferencing topographic maps is easy: the coordinate grid is seen on the scanned image!!



These are FCC images. (It can be any)

Upper, not georeferenced and lower is georeferenced.

Georeferencing images is similar to georeferencing maps, but…

FCC's are composed of different bands of identical geometry: the individual bands must be georeferenced. There are methods to do all together.

Georeference of images is more difficult: clear Ground Control Points (GCP) in the image need to be found. No grid available!!,… no easy way to collect GCP with assigned coordinates.

GR= POSITIONING A PIXEL IN A COORDINATE SYSTEM

In a GIS/RS package the user needs to impose of a referenced geometry.

- Images/maps without a georeference cannot be used for any kind of geometrical information in a GIS/RS system:
  - No pixel size or scale possible… no location.
  - Most GIS operations are not operational.
  - Many RS operations are not possible.
  - *However, radiometric information in the raster is not affected.*
- **Georeferencing** is the process executed over an image/map to assign to each pixel a real world coordinate.
  - After georeferencing, GIS and RS operations are all possible
- In other words: georeferencing is the process of fixing a raster data set (2D or 3D) in a space with cartographic coordinates.

POINTS TO FIX A SHAPE IN SPACE.

In a 2D plane, 3 points (minimum)

- After fixing 3 points (GCP's), the "solid" mantle does not move.
- The GCP points must be located in a plane: (X,Y)
- To avoid errors: the bigger the number of points, the better.

In 3D, we require to fix the height of some points.

- The GCP points are then located in the space: (X, Y, Z)
- What can bring many X,Y,Z points in a GIS?: i.e. A DEM

1. The georeferencing process does not changes the geometry of an image (values, pixel size and or orientation)
2. The geo correction or geocoding process changes the geometry of an image (values, pixel size and or orientation) Depends on the mathematical model used in the georeferenced.
3. To control that the georeferencing error is within the accepted values we seek more (sometimes much more) than three points to do georeferencing.
4. Sigma (RMSE), given as control quality check of the georeferencing, strictly means It gives the overall error of only the points selected as GCP's.
5. Landsat image - This image may have a rotation, a displacement and a linear scaling as distortion. As such an affine model will be the most adequate choice.
6. Modis 1 km image - polynomial model.
7. Topographic map - A topographic map does not have distortions as it is orthographically corrected. The error/distortion that comes out of the scanning process to make it digital could be solved with a Conformal model.
8. Map grid (a grid crosses the map area) - t is best to use for the registration or georeferencing of a scanned topographic map.

WHEN HEIGHTS (THE "Z" COORDINATE) ARE IMPORTANT?

Map

Medium resolution image

High resolution image

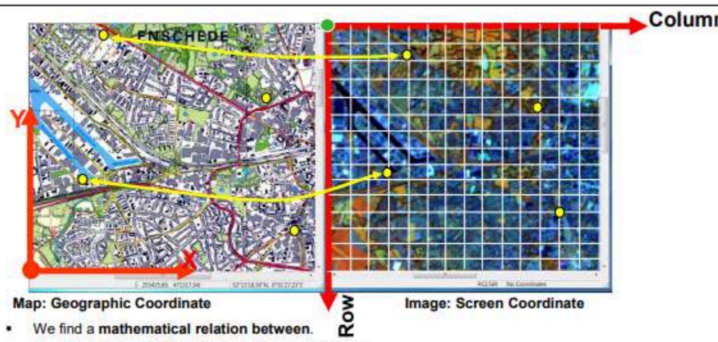
- When heights are neglected due to poor resolution of imagery or lack of interest to our study: PERFORM a 2D (x,y) georeferencing.
- If heights are of interest to our study, look for a high resolution imagery and PERFORM a 3D (x, y, z) georeferencing (limited in this course)

GEOMETRIC ASPECTS OF IMAGE DATA

- 2D approaches: Solves X, Y (medium and coarse satellite imagery)
  - Heights of objects are negligible
  - Possible operations:
    - Georeferencing: Allows location of points
    - Geocoding (resampling): allows location and reorientation of points
- 3D approaches: Solves X, Y, Z (Aerial Photos and very high resolution satellite imagery)
  - Heights of objects are essential
  - Deals with relief displacement
  - Possible operations after a 3D Georeferencing:
    - Monoplotting
    - Orthoimage production
    - Stereoplotting

Challenge:
- You need to have clear concepts of location and orientation?
- Location (absolute position)
- Orientation (relative to)

GEOREFERENCING IN 2D



Map: Geographic Coordinate

Image: Screen Coordinate

- We find a **mathematical relation between**.
  - *Pixel coordinates* (Row, Column) inherent of the image.
  - *Map coordinates* (a coordinate system for the Earth point where the pixel is x,y)
- The distortion of the image as seen by the satellite is NOT CORRECTED, but the geographic coordinate of every pixel will be acceptable.

INTRO TO GEOREFERENCE (2D)

- Georeferencing solves 2 problems:
  - Assigns coordinates to pixels, via a model that solves the coordinate distortion.
  - Distortions in a 2D images are produced by several causes but are treated altogether!!
- After Georeferencing it is possible to:
  - Measure in the image
  - Spatially combine vector and raster data (ANY GIS operation is possible
  - Compare and fuse images (software dependent)
  - But to produce images in a certain map projection you need GEOCODING!!! (1 step more, but easy)

CAUSES OF GEOMETRIC DISTORTION

- The perspective of the sensor optics (oblique viewing)
- The curvature and rotation of the earth.
- The terrain relief: Relief Displacement
- Others

- The motion of the scanning system.
- The motion and instability of the platform.
- The platform attitude, altitude and velocity.

What is the challenge in georeferencing?

The challenge (or problem to solve) in georeferencing is that every pixel represents a growing size area on Earth as the position moves from nadir (vertical) to off-nadir (sideways). The variable size caused by the perspective of the view, has to be solved by the mathematical relation between the real and then image coordinate set of equations.

The IFOV (instantaneous field of view) of a sensor is the constant (fixed) conic solid angle through which a sensor "sees" and "samples" one pixel on the Earth. As this sensor scans from nadir to off-nadir, the area of the Earth seen by this sensor grows, see Figure 1. (You can do this experiment with a paper cone in your eye looking straight down and then sideways and asking someone to delineate in the ground the areas you see in both cases, see Figure 2).

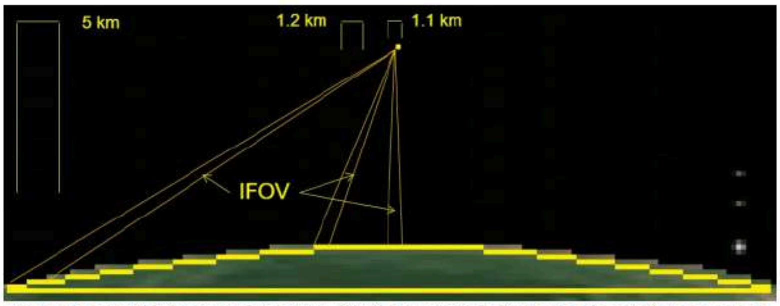
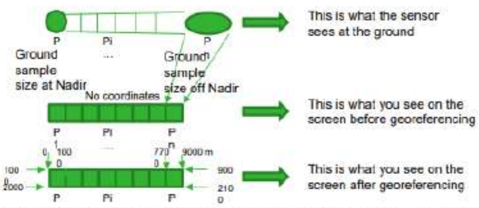
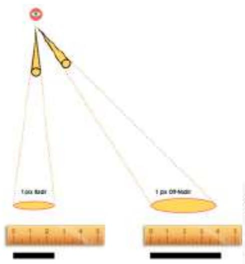
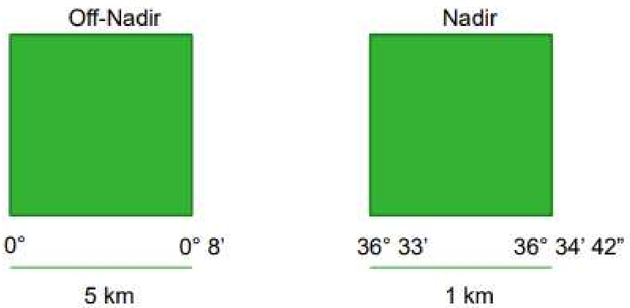


Figure 1: The sensor is at the center top of the figure. IFOV is constant. Despite that the size of one pixel in the screen is the same, they represents different distances on Earth. The distance depends on how far the pixel is with respect to Nadir position.



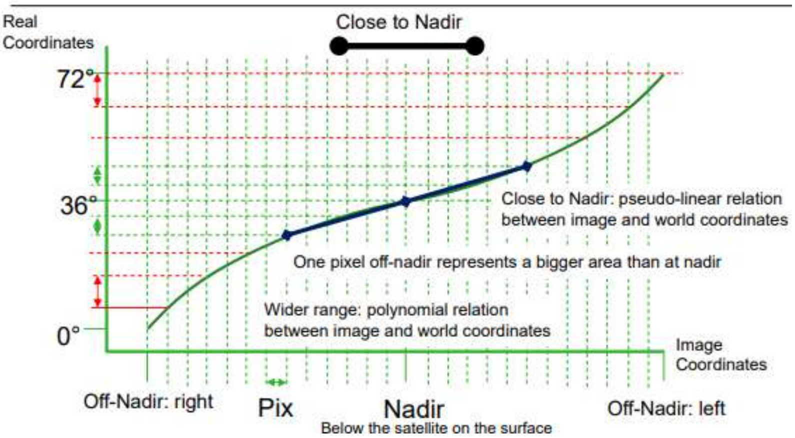


WHAT ARE THE COORDINATES OF A SINGLE PIXEL?



For images taken under an angle, the relation between image coordinates and real coordinates is not linear!!

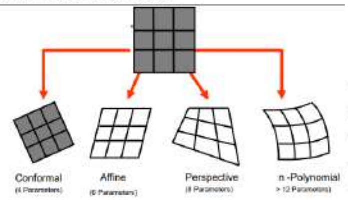
PLOTTING IMAGE VS REAL COORDINATES



WHAT FOR MEDIUM TO HIGH RESOLUTION IMAGES?

- Moderate to Very small swath:
  - Landsat 185 km
  - IRS 146 km
  - SPOT 117 km
  - Ikonos 11 km
- An almost Linear distortion stands for nadir and off-nadir pixels
- *The relation between image and real coordinates can be handled with an affine (or eventually polynomial degree 1) transformation.*

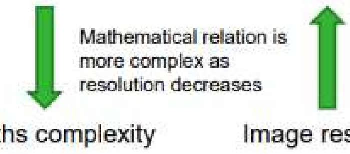
GEOREFERENCE MODELS



SO... WHAT IS GEOREFERENCING?

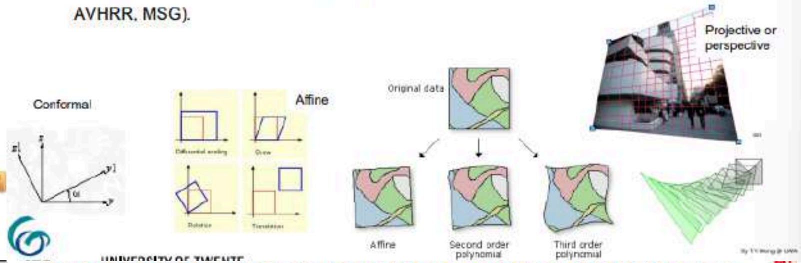
- We need to find **the mathematical relation** that starting from the IMAGE COORDINATES, gives the MAP COORDINATES.
- NOTICE that this mathematical relation must solve the distortion model, but it does not affect the pixel values read by the sensor, only gives good coordinates!!

Planar distortion:	Local areas (maps or images)	i.e. cadastral scanned map.
Linear distortion	High resolution images	i.e. Landsat, Spot, etc.
Complex distortion	Moderate to low resolution images	i.e. MODIS, AVHRR, MSG, etc



THE DISTORTION TYPE DECIDES THE MATHEMATICAL RELATION

- Image is evenly distorted → a linear transformation (conformal, affine) is applied (small FOV images. i.e. Landsat).
- Image is unevenly distorted → higher order transformations (large FOV images. i.e. NOAA AVHRR, MSG).



CONFORMAL TRANSFORMATION OF COORDINATES

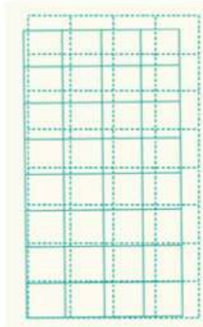
AFTER RICHARD KNIPPERS

2D Linear (or first-order) transformation that contains:

- rotation,
- a uniform scale change,
- followed by a translation.

$$X' = s\,X\cos(\alpha) - s\,Y\sin(\alpha) + x_0$$
$$Y' = s\,X\sin(\alpha) + s\,Y\cos(\alpha) + y_0$$

- The rotation is defined by one rotation angle (α), and the scale change by one scale factor (s). The translation is defined by two origin shift parameters (x_o,y_c)



In the equations, (X,Y) represent the image coordinates and (X', Y') the Calculated world coordinates

LINEAR OR AFFINE TRANSFORMATION OF COORDINATES

AFTER RICHARD KNIPPERS

2D Linear (or first-order) transformation that contains:

- rotation,
- a scale change in x- and y-direction,
- followed by a translation.

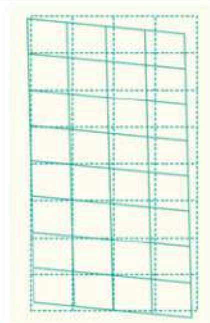
The transformation function is expressed with 6 parameters:

one rotation angle (a),

two scale factors, in "x" (s_x) and in "y" (s_y), and two origin shifts (x_o, y_o).

$$X' = s_x\,X\cos(\alpha) - s_y\,Y\sin(\alpha) + x_0$$
$$Y' = s_x\,X\sin(\alpha) + s_y\,Y\cos(\alpha) + y_0$$

You always need three coordinates to solve this equation.

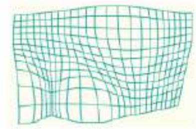


In the equations, (X,Y) represent the image coordinates and (X', Y') the Calculated world coordinates

POLYNOMIAL TRANSFORMATION OF COORDINATES

AFTER RICHARD KNIPPERS

- It is a 2-D non-linear transformation containing
- *A translation,*
- *a rotation and a*
- *variable scale change.*
- The transformation function can have an infinite number of terms. Normally no more than order 3.

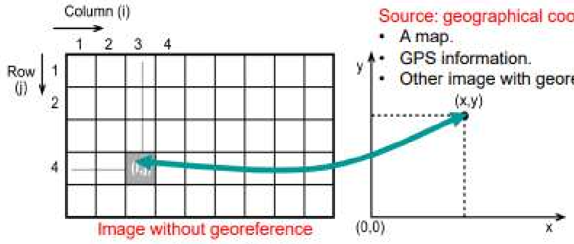


Distortion in an aerial photograph. The coordinates can be approximated with a high order polynomial equation.

$$X' = x_0 + a_1X + a_2Y + a_3XY + a_4X^2 + a_5Y^2 + a_6X^2Y + a_7XY^2 + a_8X^3 \ldots$$
$$Y' = y_0 + b_1X + b_2Y + b_3XY + b_4X^2 + b_5Y^2 + b_6X^2Y + b_7XY^2 + b_8X^3 \ldots$$

In the equations, (X,Y) represent the image coordinates and (X', Y') the Calculated world coordinates

2D GEOREFERENCE: LEAST SQUARES ADJUSTMENT



Source: geographical coordinates
- A map.
- GPS information.
- Other image with georeference

Data to be collected

Image		Source	
i	j	x	y
254	68	958	155
149	22	936	151
40	132	916	176
26	269	923	206
193	228	954	189

In general
x = a + b(i) + c(j)
y = d + e(i) + f(j)

Since (i,j) is known for each point, the coordinate can be calculated

i,j represents row and columns.

UNIVERSITY OF TWENTE.

Add three 0 to change it to Km

With 3 points the image is georeferenced but with 4 points you can calculate the errors.

2D APPROACHES: GEOREFERENCING: QUALITY CONTROL

	Measured				Modeled		Difference	
GCP	i	j	x	y	x_c	y_c	d_x	d_y
1	254	68	958	155	958.552	154.935	0.552	-0.065
2	149	22	936	151	934.576	150.401	-1.424	-0.599
3	40	132	916	176	917.732	177.087	1.732	1.087
4	26	269	923	206	921.835	204.966	-1.165	-1.034
5	193	228	954	189	954.146	189.459	0.146	0.459

x-xc=dx
y-yc=dy

$$m_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} dx_i^2} \qquad m_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n} dy_i^2} \qquad m_{total}(RMSE) = \sqrt{m_x^2 + m_y^2}$$

These equations are done by the software

Mx and my calculate the errors in x and y
Mx = total

- After entering a limited number of GCPs the transformation model can be calculated.
- Entering more control points allows the estimation of errors!! This estimation is essential
- Errors are measured at the GCP's and not at any place in the image!!
- The accuracy of the georeference can be *characterized* based on this error, but strictly speaking, it applies only to the GCP's. Therefore, it cannot be ensured that the same error applies for all points in the map!!

PROCEDURE: IT IS A SEQUENCE [TO READ]

Select appropriate transformation **based on the image characteristics**: (conformal, affine or polynomial)

Decide on a required accuracy for the georeferencing (RMSE).

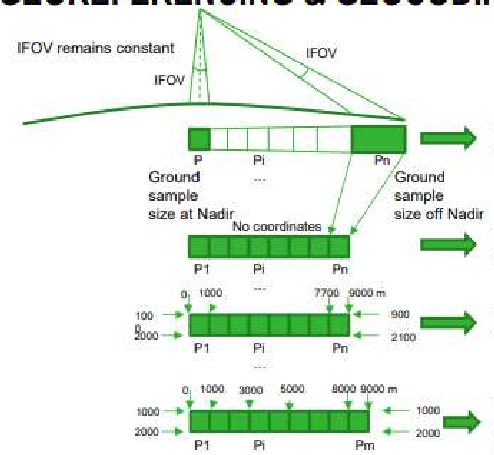
Process to calculate the transformation parameters

- ➢ Select ground control points
  - o Sufficient to solve the transformation equations and derive an error estimate. They should be accurate and reliable
  - o Well distributed all over the image, covering inside & outside the work area
- ➢ Compute transformation (automatic)

Assess GCP's residual errors and overall RMSE

- ➢ If RMSE does not match requirements:
  - o Review all GCPs. Correct, adapt, change or reject.
  - o Very eventually: Review selected transformation

GEOREFERENCING & GEOCODING AT PIXEL LEVEL



This is what the sensor sees at the ground

This is what you see on the screen before georeferencing

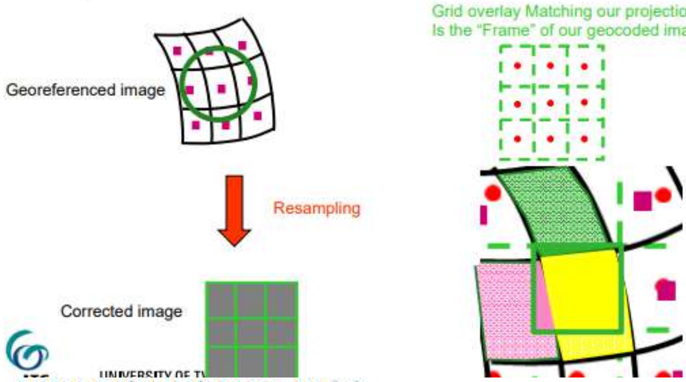
This is what you see on the screen after georeferencing

This is what you see on the screen after 1000 m geocoding

GEOCODING PROCESS

- Geocoding
- Georeference + Resampling
- The process allows
  - Select a new output coordinate/projection (any)
  - Define the new cell size
  - Define a new subset area of the map
  - Makes images North oriented & compatible with toposheets.
  - Disadvantage
    - The original geometry seen by the satellite changes
    - The original pixel needs to be shifted, so a new DN value needs to be calculated (resampling process)
    - DN value changes based on the resampling method

GEOCODING AN IMAGE: PROCESS AND RESAMPLING

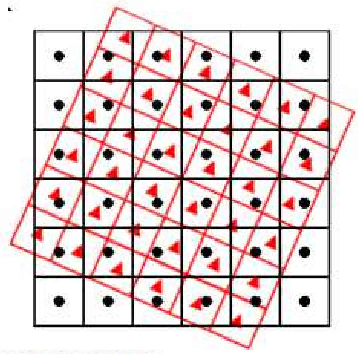


WHAT IS RESAMPLING?

▲ Center pixels for the georeferenced image

● Center pixels for the geocoded image.

Resampling: Method chosen by the user and assigns new DN values to the geocoded pixels.

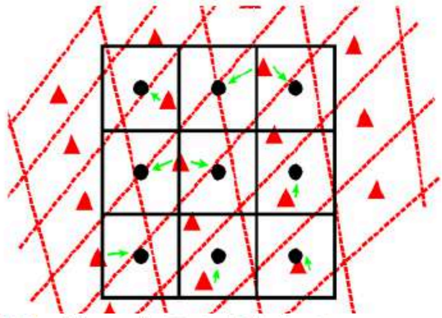


NEAREST NEIGHBOUR RESAMPLING

Geocoded pixel (black) adopts the value of the closest georeferenced pixel (red)

"Preserve" radiometry (DN) or the original value of the closest pixel

Image becomes somehow "blocky" (not much)



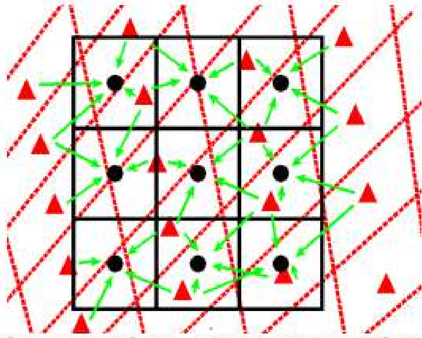
Measure the Centre of the closets pixel to the canter of the cell

BI-LINEAR RESAMPLING

Geocoded pixel value is a weighted linear calculation from the 4 closest pixel values of the georeferenced image.

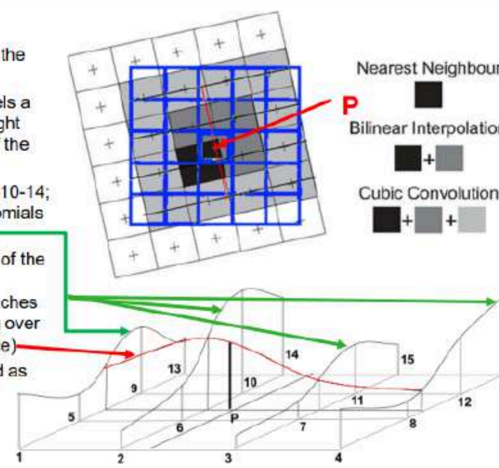
Weight ~ 1/d

Distorted radiometry but smooth changes.



BI-CUBIC RESAMPLING: (GRAPHICAL EXPLANATION)

- Takes 16 closest pixels from the one to be resampled.
- From the center of these pixels a vertical line is built with a height equivalent to the DN value of the original pixel
- In groups of 4 (1-5-9-13; 2-6-10-14; etc), it creates 4 cubic polynomials (black curves)
- From the central position "P" of the newly geocoded image, a perpendicular is built till it reaches the polynomial curve passing over the interpolation point (red line)
- Interpolated value is obtained as the length of this raised line.

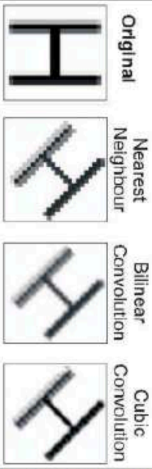


UNIVERSITY OF TWENTE.

RESAMPLING METHODS: HOW VISUALLY APPEARS?

- Nearest neighbor (NN)
  - Maintains original DN values
    - Good for Quantitative RS (e.g Image Classification) and thematic images
  - Results in jagged edges
- Bilinear interpolation (BIL)
  - Smoothed look
- Bicubic (BIC)
  - Edges look enhanced, might be shifted
- BIL and BIC good for scanned maps!

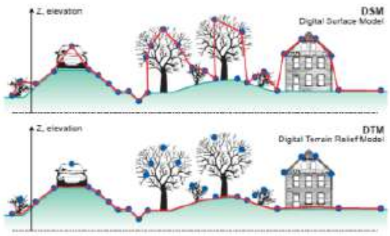
If we want to keep the values that the satellite saw then its preferred to use nearest neighborhood

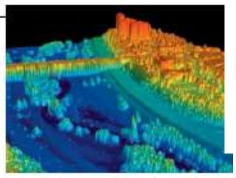


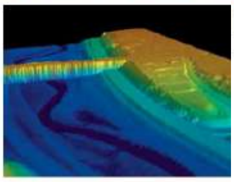
3D GEOREF: VERY HIGH RESOLUTION OR WHEN HEIGHTS ARE RELEVANT

- Necessary when:
  - 3D data (x, y, z) is needed.
  - 2D data is needed but the relief causes errors beyond requirements.
- Requirements:
  - It requires transformation of image coordinates to (X,Y,Z). The reverse process is also possible.
  - What data we need as input?
    - The original image as planar information (as in 2D) +
    - Height information (for the Z coordinates): DEM

DSM AND DTM (DEM)







DSM: Digital surface model

DTM Digital Terrain (relief) model

DEM: Digital Model for raster representations

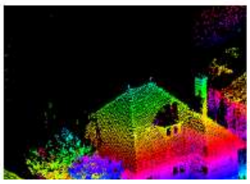
DSM ACQUISITION

Digitising/vectorising

Stereo models

Radar Interferometry (INSAR)

Laser Scanning/ LIDAR (airborne)



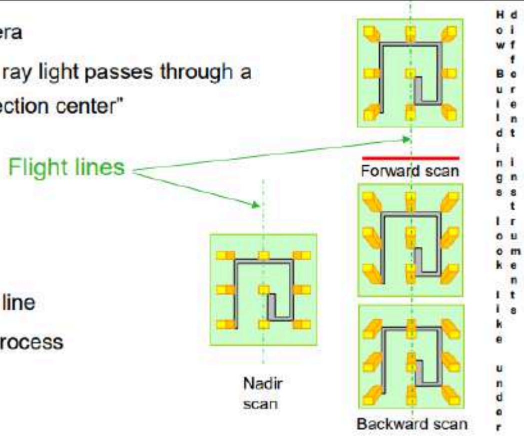
SCHEMA OF 3-D RELIEF DISTORTIONS FOR DIFFERENT SENSOR PERSPECTIVES

Vertical AP, Frame camera

- Central projection: ray light passes through a single point : "projection center"
- All in one shot

Line camera, scanner

- Pixels read line by line
- Continuous scan process



Georeference concepts

Georeferencing is the automatic or manual process necessary to give positional information to a **raster** image without altering its radiometric or geometric properties. The concept of georeference does not apply to vectors since vectors are built as a series of coordinates already.

A georeferenced image **does not store the coordinates of any individual or group of pixels**. It stores the **mathematical relation** between the row and columns of some selected and well distributed pixels in a raster, with the real coordinates of those pixels. As such those chosen pixels must be identifiable to be able to assign good coordinates to them. Since only a mathematical relation is built, the image before and after georeferencing is unaltered. We will repeat this below with more detail.

AERIAL PHOTOGRAPHY: GEOMETRY CAMERA ORIENTATION

To correct a 3D image for distortion, heights need to be measured

Purpose: to obtain the parameter values from transforming the terrain coordinates (x,y,z) to image coordinates and vice versa.

We base this process in the image co-linearity:

- A point in the terrain passes through a projection center onto the film/array.

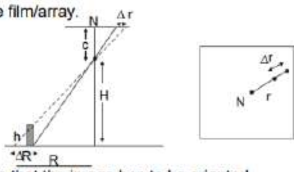
"C" is the focal distance (constant for a camera)

solving triangles

$\Delta r = r \cdot h/H$

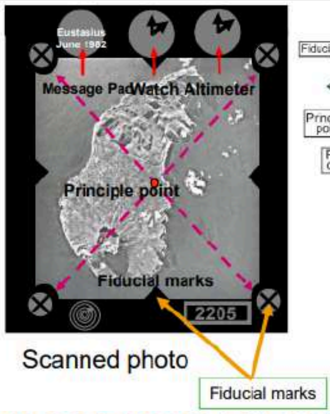
$h = \Delta r \,.\, H / r$

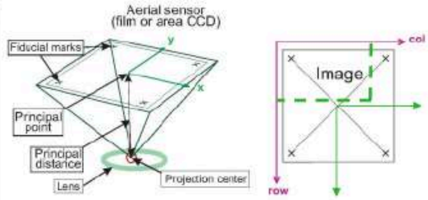
So the heights "h" are calculated



- However the location of every pixel needs still to be found. To do that the image has to be oriented. The orientation of a single image is done after interior and exterior orientation process.

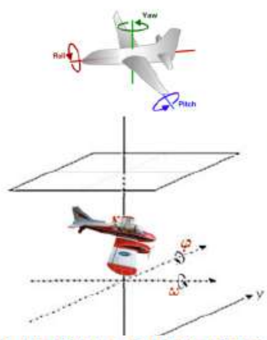
3D APPROACH: INTERIOR ORIENTATION. INDIVIDUAL PHOTOS





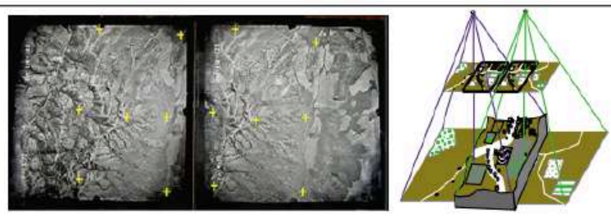
- Interior orientation: position of the projection center with respect to the image: Image Coordinate System.
- Each pixel in the image is measured in "cm" or pixels with respect to an orthogonal pair of axis. Origin could also be the principal point.
- Camera calibration report states the location of the principal point and the principal distance

3D APPROACH: EXTERIOR ORIENTATION INDIVIDUAL PHOTOS



- Determines the position and orientation of the projection center with respect to the terrain (location (x,y,z) and orientation (ω, ψ, κ)
- Exterior orientation solved by RPC (rational polynomial coefficients), modern sensors/cameras. If not available, then:
  - Indirect camera orientation (GCP's)
  - Direct camera orientation (GPS and IMU)
  - Integrated camera orientation (combination of the two)

RELATIVE ORIENTATION FOR STEREO PAIRS



- Stereo pairs are overlapping photographs.
- Conjugate points are found in overlap area of the photo
- Relationship between the two Image Coordinate Systems
  - Instead of doing two independent exterior orientations, first a relative orientation of the two images can be done, followed by an absolute orientation of the pair to the terrain coordinate system
  - This process is nowadays automatic creating CLOUD POINTS. (Advance topics)

ABSOLUTE ORIENTATION

- Follows relative orientation and defines the relationship between the stereo model and the terrain

CORRECTION OF RELIEF DISPLACEMENT

If relief prevents to derive accurate planimetric coordinates for a photo, then relief needs to be considered and corrected

a DTM is required to correct for relief displacement on a single photograph

After the geometry of this correction is built, then 2 mapping techniques become available

digital monoplotting

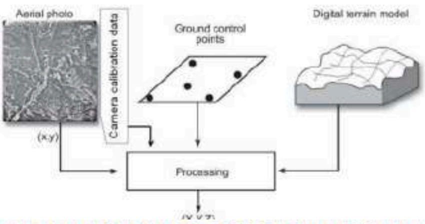
- Allows direct conversion between the measured image coordinates on the photo to terrain (real world) coordinates

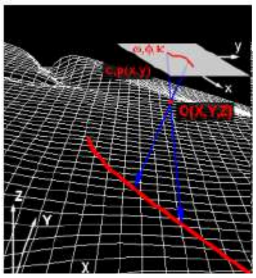
production of orthophotos or orthoimages

- the photograph is scanned to provide a digital image
- the pixels are transformed and resampled
- An orthographically corrected map is obtained after this process

DIGITAL MONOPLOTTING

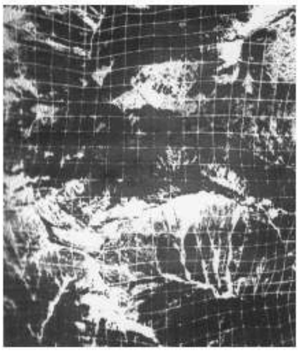
- Monoplotting is a procedure applied to a single raw aerial image that allows to get corrected real 3D coordinates from direct digitizing on the raw photo.
- The required datasets are: a DTM and an aerial image with its orientation parameters or a georeferenced orthoimage respectively.
- No resampling is done! So image remain distorted but the digitised features have the correct coordinate.



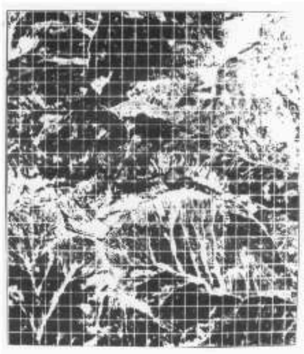


ORTHOPHOTO/ORTHOIMAGE

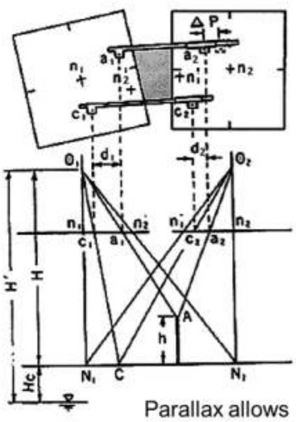
- Orthoimage delivers only (X,Y) coordinates!



On the original image the map grid is distorted



Geometrically correct orthoimage



Hc : flight height

c : control point

Hc : height of control point

H : H' - Hc

$\overline{n_1 \cdot n_2}$: principal points on a pair of stereo photos

$\overline{n_1' \cdot n_2'}$: transferred points from each principal point of the photos

$\overline{n_1 \cdot n_2}$: principal distance

$\overline{c_1 \cdot c_2}$: distance between at point c on a pair of stereo photos

$\overline{a_1 \cdot a_2}$: distance between at point a on a pair of stereo photos

parallax at point c : $P = \overline{n_1 \cdot n_2} - \overline{c_1 \cdot c_2}$

parallax distance between point a & c : $\Delta P = \overline{c_1 \cdot c_2} - \overline{a_1 \cdot a_2}$

obtained height : $h = H \cdot \Delta P / (P + \Delta P)$

Parallax allows measurements of heights when Δr is very small

Geocoding concepts

Geocoding, sometimes called resampling, reprojection, warping, among others[1], is a process done **over a georeferenced raster** where the resulting new image changes its original geometrical properties and values in order to fit certain desired north oriented (re-)projection.

The characteristics of the final geocoded image are:

- In the geocoding process the user designs a new north oriented empty grid where the pixel values will be calculated from the original georeferenced image based on a number of decisions taken by the operator.
- Consequently the final image will be north oriented.
- Geocoding is done over the whole original image or a subset (subimage) of it.
- Very important. The geocoded image can be built in any projection. It is just matter of choice.
- The pixel size is also a choice of the user, therefore the total number of row and columns may likely change with respect to the original image.
- All the pixels of the geocoded image represent the similar size on Earth. This is not true in the georeferenced image.
- The original pixel value (radiance or other) might be kept in the geocoded image **if nearest neighbourhood resampling method is chosen**, but IN NO CASE the area of the Earth covered by this new pixel is identical to the one in the original image. This will be clarified later in the example
- Using other resampling methods (bi-linear and cubic convolution) both the pixel values and the original area of Earth seen by the pixel will change.
- The geocoding process is compulsory when the GIS/RS project is composed of different rasters in different projections that need to be subjected to common operations (map calculation, data extraction, etc). In other words, when 2 rasters of different georeference need to be combined, a geocoding process is needed.

NOTE: As the process changes the original values of the pixel, users working on quantitative remote sensing (where pixel value information is highly controlled) should be aware of the conversion effects that geocoding produces, at all times.

Data Management

DATABASE

DATA, DATASET, DATABASE, DBMS AND DB SYSTEM

Data

- Is a **resource** held on paper or in digital format that serves to record or administer some **facts and descriptions of phenomena of interest**.

Data set (or dataset):

- A homogeneous **collection of data** normally describing a single kind of phenomenon

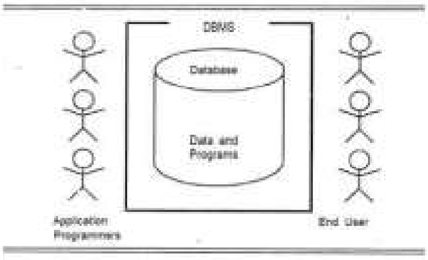
Database (DB)

- A **collection of interrelated data sets properly structured** by means of, and stored through a DBMS (Data Base Management System)

Data - It is useful information about facts of reality.
Data set - collection of data normally related to one kind of phenomena.
Database - A collection of interrelated data, with specific phenomena is extracted.
Database management system(DBMS) - A software used to manage database, to set up and maintain many database.

Database management system (DBMS)

- A **software package** that is designed for the purpose of managing databases. This means, DBMS allows to set-up, maintain and explore one or more databases.



Database system

DATABASE MANAGEMENT SYSTEM (DBMS)

WHAT DOES A DBMS OFFER TO USERS?

- Supports proper storage and manipulation of ***very large data sets***.
- Can be instructed to guard over ***data correctness***.
- Allows the ***control of data redundancy***.
- Supports the ***concurrent use*** of the same data set by many users.
- Includes data ***backup and recovery functions*** to ensure data protection and availability at all times.

Why do we use DBMS -

1. To manage the data.
2. To control database.
3. To back up and recover data.
4. It allows to control data redundancy, it means to avoid multiple entry of one thing.
5. Concurrent use of data by multiple use.

DATA MODEL

A DBMS works based on a ceratin **data model**.

A data model **organizes elements of data and standardizes how they relate to one another**

A data model is an integrated collection of:

- **Data structuring primitives,**
- **Rules of how to structure, and**
- **Mechanisms to handle the data**

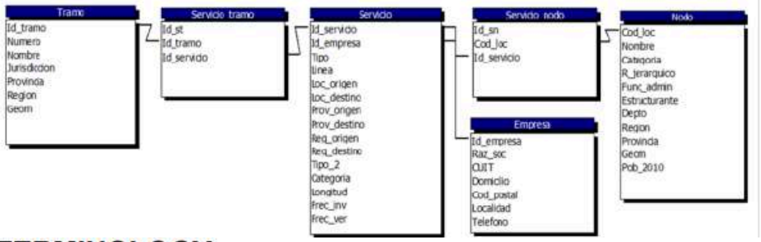
In other words, a data model is a **toolbox** that allows us to **create/define** *a* database structure and **manipulate** the data stored in it.

RELATIONAL DATA MODEL

WHY STUDY RELATIONAL DATA MODEL?

There are different data models: hierarchical, network, object-oriented, etc. → DBMS linked to most GIS packages make use of the **relational data model**

The **relational data model** structures a DB as a **collection of inter-related tables (or *relations*)** → we model a piece of reality as linked tables



TERMINOLOGY

THE LANGUAGE USED IN RELATIONAL DATA MODEL

Structured Query Language *(SQL)* – the relational database language:

- SQL is a DSL (*domain-specific language*) used in programming and **designed for managing data held in a relational DBMS**.
- **Natural language →** its syntax is easy to read
- Developped in the 70's → initially called SEQUEL (Structured English QUEry Language)
- **Particularly useful in handling structured data where there are relations between different entities/variables of the data.**

Find all cities (codes and names) that belong to the department called Tilcara

A

Code	Name	Dep
38077020	Cienega de Paicone	Santa Catalina
38077080	Paicone	Santa Catalina
38084020	Coranzuli	Susques
38084090	San Juan de Quillaques	Susques
38084080	Puesto Sey	Susques
38084040	Huancar	Susques
38084070	Pastos Chicos	Susques
38084060	Olaroz Chico	Susques
38084050	Mina Providencia	Susques
38084030	El Toro	Susques
38084055	Olacapato	Susques
38084045	Jama	Susques
38084100	Susques	Susques
38084010	Catua	Susques
38094050	Tilcara	Tilcara
38094040	Maimara	Tilcara
38094010	Colonia San Jose	Tilcara
38094030	Juella	Tilcara
38094020	Huacalera	Tilcara
38098010	Barcena	Tumbaya
38098050	Volcan	Tumbaya

```
SELECT   Code, Name
FROM     A
WHERE    Dep = Tilcara
```

RELATIONAL DATA MODEL

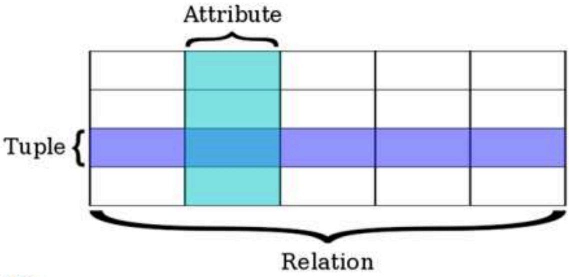
The relational data model is an integrated collection of:

1. Data structuring primitives = **attributes, tuples,** and **relations**
2. Rules of how to structure = **data definition language**
3. Mechanisms to handle the data in a database = **data manipulation language.**

Data model - a tool box used to manipulate your data.
Relational Data Model - it contains interrelated tables. It is a collection of attributed, tuples and relations.

DATA STRUCTURING PRIMITIVES

- **Attribute**: properties that describe an entity
- **Tuple**: the entire row of attribute's values corresponding to a particular entity
- **Relation** *(or table)*: a collection of tuples that are similarly shaped



ATTRIBUTE

Attribute

- In the relational data model, we represent **real-world objects** by **tuples** stored in **relations** (*tables*).
- These real-world objects have particular **properties** that describe them. These properties are called **attributes**.

The relation *Productions* has these attributes: *cid*, *crop*, *annum*, *score*, *quality*

Productions

cid	crop	annum	score	quality

Domain

- A set of acceptable values that an attribute is allowed to contain
- Original set of **atomic values** used to model data.
- By atomic value, we mean that each value in the domain is **indivisible** as far as the relational model is concerned; e.g. the domain of the attribute *Day* has the set of all possible days: {Mon, Tue, Wed…}.
- An attribute domain belongs to a certain **data type (also called domain)**
- There are many system-defined data types/domain: Numeric (Integer, Floating-point); Character (Char, Varchar); Temporal (Date, Time); Boolean (true/false, y/n)

Attribute - vertical columns in a table. Properties that describe an entity.
Tuple - Horizontal columns. They are described by list of attributes value.
Relation - table. It is a collection of tuples and attributes.
Domain - Data range or characteristics of an attribute. Set of values that are allowed.
Atomic value - unit assigned to domain.
Language used - SQL is used structured query language, It is domain specific language developed to manage data. It is a English based language.
Missing value - sometimes the value is unknown for a certain entity. In SQL null is given to some value that is unknown or not applicable.
Relation schema -name of the table, the list or attributes and the domain of attributes.
Relation instance - a set of tuples in a relation.6
Data definition language (DDL) - To create a table DDL is given. Create, alter, drop is used.
Data manipulation language - it is used to manipulate the data, select insert, update and delete are the common commands used.
Relation Database- it is based on relational data model. A collection of tables, that is structured to organise relations.
Database schema - it is described as a formal language, it represents logical view of the entire database.

ATTRIBUTE & DOMAIN

Each attribute has an associated domain:

- *In the relation Productions:*
  - ✓ *cid: integer*
  - ✓ *crop: varchar(255)*
  - ✓ *annum: integer*
  - ✓ *score: integer*
  - ✓ *quality: varchar(255)*

Productions

Field Name	Data Type	Description
cid	Number	country code
crop	Text	crop type
annum	Number	year in which produced
score	Number	production achieved, in tonnes
quality	Text	* = Unofficial figure, [ ] = Official data, F = FAO estimate, M = Data not available

NULL VALUES

Some attributes may have **missing values**. An attribute value may be:

- 'unknown'
- **'not applicable'**

In SQL we use **NULL** in both cases

NULL *(or Null)*

- A special marker used in SQL to indicate that a data value does not exist in the DB.
- It enables the representation of missing and inapplicable information
- **Not to be confused with a value of 0**

TUPLE

Tuple

- A tuple is a **record**
- In the relational data model, tuples represent **real-world objects/phenomena** that have certain attributes
- A tuple can be defined as **a list of attribute values**

One tuple from the relation *Productions* is: *(cid=3, crop="Rice,paddy", annum=2000, score=300, quality="F")*

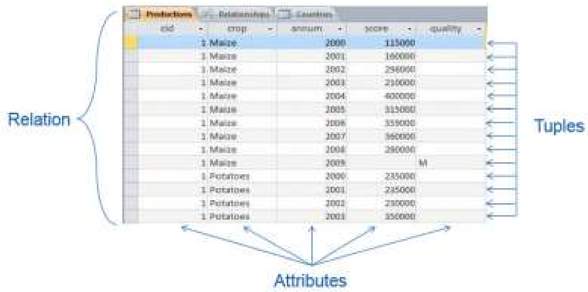
cid	crop	annum	score	quality
3	Rice, paddy	2000	300	F

RELATION (or TABLE)

Relation (*Table*)

- A **collection of tuples** that are **similarly shaped:** all tuples have the same attributes

The relation *Productions* in *FAOcrops.mdb*



RELATION SCHEMA AND RELATION INSTANCE

Relation schema

- Basic information describing a table or relation:
  - **Name** of the relation;
  - List of **attributes**;
  - **Domain** of each attribute

Relation instance

- Set of tuples that adheres to all the requirements that are formulated by the relation schema
- The **set of tuples in a relation at some point in time**

- **Relation schema**

Productions (cid: integer, crop: varchar(255), annum: integer, score: integer, quality: varchar(5))

- **Relation instance** –

Productions

cid	crop	annum	score	quality
1	Maize	2000	115000	
1	Maize	2001	160000	
1	Maize	2002	298000	
1	Maize	2003	210000	
1	Maize	2004	400000	

RELATIONAL DATA MODEL

Database instance

- A **collection of relation instances**, one for each relation in the database schema

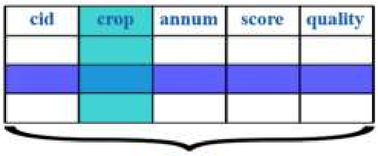
1. Data structuring primitives = **attributes, tuples,** and **relations**
2. Rules of how to structure = **data definition language**,
3. Mechanisms to handle the data in a database = **data manipulation language**.

DATA DEFINITION LANGUAGE

Data Definition Language (DDL)

- A (subset of a) computer language used to create and modify the **structure** of database objects in a database:
- Set of commands that can be used to define the database **schema:** CREATE, ALTER, DROP

```
CREATE TABLE Productions(
                              cid single,
                              crop varchar(255), annum integer,
                              score integer, quality varchar(255)
                              )
```



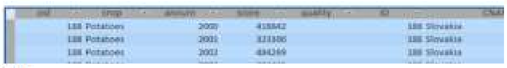
Productions

UNIVERSITY OF TWENTE.

Data Manipulation Language (DML)

- A (subset of a) computer language used for **adding** (*inserting*), **deleting**, and **modifying** (*updating*) **data** in a DB
- Set of commands that deals with the **manipulation of data:** SELECT, INSERT, UPDATE, DELETE

```
SELECT *
FROM Productions AS p, Countries AS c
WHERE c.ID=p.cid AND p.crop="Potatoes" AND
c.CNAME="Slovakia"
```



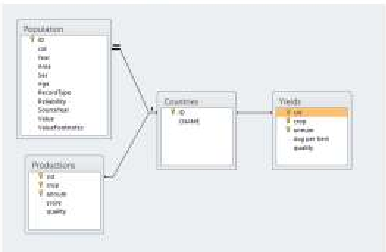
RELATIONAL DATABASE

Relational database

- A DB based on the **relational data model**
- A collection of relations (*tables*) **structured to recognize relations** *(links/associations)* between stored items of information.

Database schema

- The DB **structure** described in a formal language
- A **collection** of relation schemas and the **associations** amongst them
- The **skeleton structure** that represents the **logical view** of the entire database



The GDB *FAOcrop.mdb* contains 4 relations: *(Productions, Yields, Countries, Population)*

Integrity constrains - Set of rules to maintain the quality of information in Data base(DB). It protects the data from accidental error. While specifying the relation schema we define the constrains, there constrains are used to define relation in tables. Three types of constrains- Domain integrity constrains - set of values for an attribute. It anything is out of the constraint then the software will show error.
Primary key - it is a unique value, meaning in two tuples of it does not have the same attribute value. Look at the entire data set before finalising the key. It is always easy to add number (sr.no) to ass as a primary key in large data base, having a primary key is mandatory.
Foreign Key - It shows values in the tables that points to the other tables meaning, it will give some small information (code) in one table and the detail about this information is given in another table. Once you have the code, you can use the same code everywhere.

INTEGRITY CONSTRAINTS

Integrity constraints

- A **set of rules** that are used **to maintain the quality of information** in a DB
- **Ensure** that the data manipulation is performed in such a way that **data integrity is not affected**.
- Used to **guard against accidental damage/errors** to/in the DB
- Integrity constraints are **specified** when the schema is defined
- Also used to **establish links between relations/tables**

Integrity constraints

- There are various types of integrity constraints that a DB provides

  We will discuss 3 types:

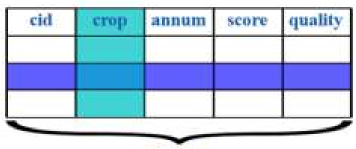
  1. **Domain** integrity constraint
  2. Entity integrity constraint: **Primary key**
  3. Referential integrity constraint: **Foreign key**

DOMAIN CONSTRAINTS

Domain integrity constraints

- Refers to the **definition of a valid set of values for an attribute**
- In order to add a new value, it needs to meet the criteria defined for that attribute. Otherwise, we get an error

```
CREATE TABLE Productions(
                              cid single,
                              crop varchar(255), annum integer,
                              score integer, quality varchar(255)
                              )
```



UNIVERSITY OF TWENTE.

Productions

Primary key

- The primary key *K* of a relation *R* is one of (or a combination of) R's attributes, such that:
  1. It is **unique** – there are no two distinct tuples of R that have the same attribute value.
  2. It is **minimal** – there is no proper subset of K that is unique.

Golden rule of keys:

- Check suitability carefully → Whether one (or a set) of the attributes can become a primary key or not can never be judged from looking at a relation instance.
- Each stored relation in a DB **must have** a defined primary key.
- Primary keys are conceived and defined (or even created!) when designing the database schema

Foreign key

- A set of attributes that is used to **refer to** a tuple in another relation.
- It must correspond with a primary key value in the second relation.
- A foreign key behaves like a 'logical pointer'.
- Links between **primary** and **foreign** keys determine the relationships amongst tables in a DB

How does DBMS and GIS go together-

GIS and DBMS is not the same, attribute table will store the information that is already georeferenced, however these two are stored in different places. Spatial(with coordinates) and non- spatial data are stored separately.
In vector, cell value is used. Whereas i**n vector object ID is used.**
Spatial database - it is a database that understands the spatial geometry, and manipulate the collection of spatial data. Spatial database management system - SDMS is a software package that offers all the same befits as a DBMS, it supports the storage of spatial data models, lines point and polygon. It focuses on querying and sharing the large spatial data sets.

A GIS **is not** a DBMS

Main GIS packages **can be linked** to a DBMS to store and manage data

LINKING SPATIAL AND NON-SPATIAL DATA

Spatial data

Non-spatial data

Code	Name	Dep	Type	Men	Women	Total
38007020	Abra Pampa	Cochinoca	Capital	4139	4566	8705
38007030	Abralaite	Cochinoca	Village	27	29	56
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Spatial and non-spatial data in a GIS environment are stored separately (.shp, .dbf) but dynamically linked

LINKING SPATIAL AND NON-SPATIAL DATA

1 Raster, ***Cell value***
2 Vector, ***Object ID***

A	A	A	B	B	B
A	A	F	B	B	B
F	F	A	B	C	E
F	F	C	C	E	E
F	F	C	E	E	A

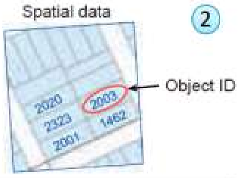
1

LandUseClass

Id	Description	Perc
A	Primary forest	11.3
B	Secondary vegetation	25.5
C	Pasture	31.2
E	Built-up area	25.5
F	Rivers, lakes	4.1

Spatial data

2



Parcel

PId	Location	OwnerID
3421	2001	435
8871	1462	550
2109	2323	1040
1515	2003	245
3434	2020	486
6371	1802	950
2209	3542	1840
1505	2609	145

Spatial attribute

SPATIAL DATABASES

Databases have evolved over the last 20 years towards supporting more complex data, such as **spatial data**

A **spatial database** allows users to store, query and manipulate collections of **spatial data by means of a new** data type: ***geometry***

Parcel

PId	Geometry	OwnerID
3421	"MULTIPOLYGON(((257462.704979333 464780.750851061,257463.89798...)))"	435
8871	"MULTIPOLYGON(((257409.813950544 464789.91585049,257407.896903...)))"	550
2109	"MULTIPOLYGON(((257785.714911912 464796.839972167,257782.59794...)))"	1040
1515	"MULTIPOLYGON(((257790.672100448 464807.13792585,257788.608078...)))"	245
3434	"MULTIPOLYGON(((257435.527950478 464803.92887633,257428.254887...)))"	486
6371	"MULTIPOLYGON(((257432.476077854 464813.848852072,257433.147910...)))"	950
2209	"MULTIPOLYGON(((257444.888027332 464826.555046319,257446.43201...)))"	1840
1505	"MULTIPOLYGON(((256293.760107491 464935.203846095,256292. 00881...)))"	145

SPATIAL DATABASE MANAGEMENT SYSTEM (SDBMS)

A S(patial)DBMS is a software package that:

- offers **all benefits of a DBMS** for data storage/management
- supports the **storage of spatial data** models – e.g. point, line, or polygon,
- can **manage coordinate systems** and transformations
- **extends querying** and manipulating capabilities of traditional SQL by adding special commands **aimed at spatial data** – e.g. `SELECT p.* FROM parcels AS p WHERE Area(p.geom)>1000`
- provide storage of the relationships between features, including the **creation and storage of topological relationships**.

GIS:

- Built-in 'understanding' of geographic space,
- Functions for spatial analysis of (almost) any kind, and
- Equipment for efficient map production
- **BUT** lack of fully developed query language to operate on tabular data.

DBMS

- Specifically designed for handling attribute data (i.e. administrative, non-spatial, tabular, thematic)
- Long tradition in multi-user concurrent management of large amounts of data
- **BUT** not support provided for spatial data storage and representation

SDBMS:

- Focuses on storage, querying and sharing large **spatial data sets**,
- Understand geometry and topology
- Provide fully-fledged querying language (i.e. SQL)
- Spatial querying functions
- Basic visualization

UNIVERSE OF DISCOURSE

We always aim at representing **only a part of the real world**.

Universe of discourse

- A **part** of the real world that is **of interest**.

We use data models for representing the universe of discourse and storing that representation in a database.

A.K.A. = the **database miniworld**

It must be well understood by the designers of the DB in order to be able to properly represent it!

PROCESS OF DATABASE DESIGN

Relational **database design** involves:

- Deciding which **relations** will be present in the database, and
- Defining **their schemas**, which in turn implies:
  - Deciding which **attributes** relations will have, and
  - What are the **domains** of these attributes.

Database design also includes:

- Defining **relationships** between relations by **integrity constraints**
- Defining desirable **queries**

Relational **database creation** involves 4 stages:

1. **Data investigation**
   - Type, quantity and quality of the data to be included in the DB
   - Entities and attributes are decided
2. **Data modelling**
   - Create a conceptual model of the data based on relationships between entities
3. **Database design**
   - Create a practical design for the DB
4. **Database implementation**
   - Create DB structure and populate it

Data modelling or entity relationship modelling involves 4 steps:

1. Identification of entities
2. Identification of relationships
   - One to one
   - One to many
   - Many to many
3. Identification of attributes of entities
4. Design of tables

The process should be led by:

- **Purpose** of the DB
- **Questions** that the DB should be able to answer

Data retrieval

Data extraction

- The process of acquiring precisely the data that is required/wanted/needed from a data set/database

Query

- Precisely formulated request to extract data from the database.
- Data extraction from the database without altering it

Transaction

- Precisely formulated request to make changes to data in the database.

To able to **properly query a database** we need to:

1. **Thoroughly know the data** set and what particularly will act as the input,
2. Precisely **understand what needs to be produced** as output
3. **'Speak' the DB query language**

Data extraction - The process of acquiring the data that you need from a proper data set or data base.
Query - it is a process of data a extraction without altering the data base. Three operators are used to tuple selection, attribute projection and joint.
Tuple selection - the tuples one is instressed in are selected. It is based on a certain condition.
Attribute projection - filtering out the columns that you need to select.

SQL - It is a language used to communicate with the data base.
Step in SQL - (brackets contain the sequence how you need to write it in SQL)

1. FROM - input the tuple that you want The table we need. (2)
2. WHERE - select the condition.(3)
3. SELECT - what are the attributes that we need. (1)

Renaming in SQL - used in renaming attributes, it is a way to shorter it so it's easy in writing.

The three query operators have some **common traits**:

1. All of them require **input** and produce **output**

2. Both input and output are **relations.** This guarantees that the output of one query can be the input of another query, which makes it possible to build more and more complex queries.

The three operators can be combined to define queries of higher complexity

Atomic Logical expression - it is used as symbols like =,<,> and etc, or like null or not
Non atomic formulas - They are combines with atomic logical expression by using connectives such as AND,OR, NOT, ( ).
Transaction - It is used to make change in a data base.
Multi relation query - it is also called **joint query or JSP query (joint select project)**. When we query more than one table at a time. It should have a very clear induction on which table is being referred to. It is gluing two or more tables together. If it is a joint query the it will be select...,from x,y,z. x,y,z being different attributes.
For mixing two table you can use Y AND Z or Y x Z.
Cartesian product - type of operation done between two sets. Relations involved in a query. All the tuples of table one multiplied by all the relations in table two.

Spatial queries - It refers to a spatial attribute like features such a geometry aka geom. The feature like geom is always written in WHERE.

SELECT firstname, lastname
FROM student as s, courserecord as cr, courses as c
WHERE s.studentnumber=cr.studentnumber AND cr.coursenumber=c.coursenumber AND c.name= "Introduction RS" AND cr.grade

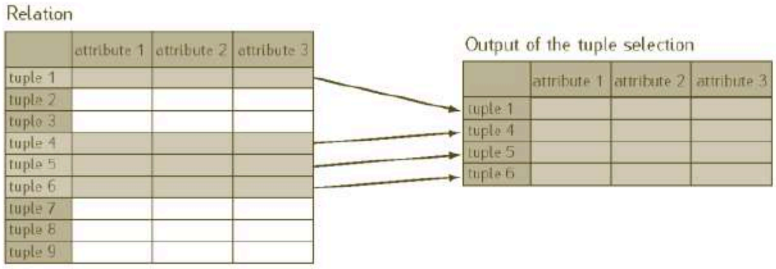
`SELECT output tuple definition` (step 3)

`FROM input tuple declaration` (step 1)

`WHERE selection condition` (step 2)

Tuple selection

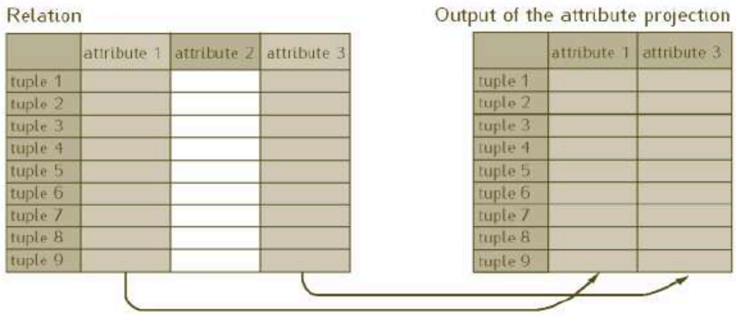
- **Retrieval of tuples** specified by a given condition.
- Works like a filter: it allows tuples that meet the selection condition to pass and disallows tuples that do not meet the condition.



- ✓ Some tuples in the input are selected, others are left out.
- ✓ We get a subset of the input data

Attribute projection

- **Retrieval of indicated attributes** from all tuples in the relation.
- We are selecting a subset of the attributes present in the input



- ✓ We are 'projecting on' attributes 1 and 3.

CLOSER LOOK AT **OUTPUT TUPLE DEFINITION**

Re-naming (*aliasing*) is sometimes handy/useful.

```
SELECT p.crop, p.cid
FROM Productions AS p
WHERE p.annum = 2009
```

For attributes only:

```
SELECT attribute1 AS Year,
       attribute2 AS Population
FROM... WHERE...
```

Selection condition determines which of the input tuples are used to generate the output of the query.

In SQL, the selection condition is represented by the `WHERE` clause:

```
SELECT p.crop, p.cid
FROM Productions AS p
WHERE p.annum = 2009
```

Selection conditions are formulated as a **logical expression**.
Logical expressions can be formed by **atomic** or **non-atomic** logical formulas. ***Atomic formulas*** (conditions) use a predicate symbol like: `=, >, <, ≥, ≤, <>, is null, is not null`

Examples (of `WHERE` clause) in SQL:

- `WHERE` $\boxed{\texttt{p.year = 2009}}$
- `WHERE` $\boxed{\texttt{p.crop is not null}}$
- `WHERE` $\boxed{\texttt{p.score > 1500000}}$

Logical expressions

Non-atomic formulas (conditions) are built up from atomic ones, using connectives like: `AND`, `OR`, `NOT`, `()`

Examples (of `WHERE` clause) in SQL:

- `WHERE p.crop is not null AND p.score is not null`
- `WHERE p.annum = 2009 OR p.annum = 2008`
- `WHERE NOT p.annum = 2009`
- If we have the following:
  Logical expression 1 = p
  Logical expression 2 = q

(**p** AND **q**) is correct if and only if p is correct and q is correct
(**p** OR **q**) is correct when p is correct, or q is correct or both of them are correct.

Example Question: Select all the records for years 2009 and 2005

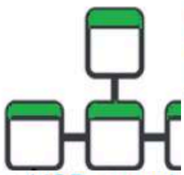
Answer : Selection condition would be :

~~Year = 2009 AND Year = 2005~~

Year = 2009 OR Year = 2005 ✓

MULTI-RELATION QUERY
RATIONALE

- Single-input relation queries were discussed so far.
- Very often, to extract required data from a database, we need to combine data from **two** or **more** relations in one query. (*E.g. remember country names and country id were in two different tables in our FAOcrops database?*)
- In such queries, selection conditions depend on the **relationships between relations** that are involved in the query.



Multi-relation query is also called a **'join query'** or a **'JSP query'**

The abbreviation **JSP** stands for:

JOIN
➢ Join condition

SELECT
➢ Tuple selection

PROJECT
➢ Attribute projection

JOIN
➢ Two or more input relations are joined (using join condition)

SELECT
➢ Joined tuples are filtered through tuple selection (using selection condition)

PROJECT
➢ Attributes are projected to the output relation (using attribute projection)

JSP QUERY
THE CONCEPT

- One of the query schemes that is used in JSP queries in SQL looks like:

```
SELECT…
FROM table1 AS t1, table2 AS t2
WHERE t1.attrib1=t2.attrib5 AND…
```

- Semantics of a JSP query is based on the **Cartesian product** of the relations involved in the query – e.g. `table1` × `table2`
- After calculating the Cartesian product of relations involved in a query, **tuple selection** and **attribute projection** are performed.

JSP QUERY
THE CONCEPT – CARTESIAN PRODUCT

- **Cartesian product** of `table1` and `table2` (`table1` ×`table2`) is a relation obtained from `table1` and `table2` by concatenating ('gluing') any tuple of `table1` with any tuple of `table2`.
- Every resulting tuple has all the attributes of `table1` and all those of `table2`.

A	B
700	a
300	b
150	c

Table1

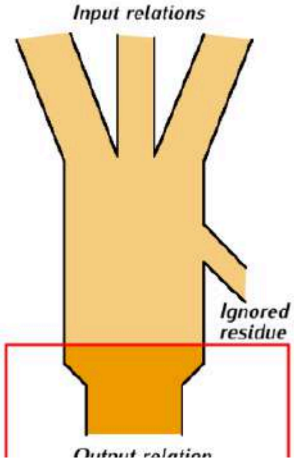
X	Y	Z
m	100	s
n	200	a

Table2

A	B	X	Y	Z
700	a	m	100	s
700	a	n	200	a
300	b	m	100	s
300	b	n	200	a
150	c	m	100	s
150	c	n	200	a

Phases of JSP query:

1. Cartesian product of input relations
   - *Which relations are involved in a query?*
2. Tuple selection
   - *What is the join condition?*
   - *What is the selection condition?*
3. Output tuple creation
   - ***What attributes will the output relation have?***



In SQL:

```
SELECT c.CNAME, p.crop, p.score, p.annum AS Year
FROM Productions AS p, Countries AS c
WHERE p.cid=c.ID AND c.CNAME='France' AND p.annnum=2009
```

join condition

- Join condition is a special type of selection condition (content of the `WHERE` clause) – because `p.cid` and `c.ID` are from different relations
- Join condition is a subset of the Cartesian product of `Productions` × `Countries`

INNER JOIN

- The example shown earlier, is called an EQUI-JOIN – because of the "=" in the join condition (`p.cid=c.ID`). This type of JSP query is the most common.
- An alternative way of writing this query, makes use of explicit join statements in the `FROM` clause of the SQL query. This alternative is referred to as INNER JOIN and the query format is:

```
SELECT…
FROM table1 AS t1 INNER JOIN table2 AS t2
ON t1.attribute_A=t2.attribute_B
WHERE…
```

- Typically, the `attribute_A` is the foreign key of `table1` and `attribute_B` is a primary key of `table2` (or *vice versa*)

3 tables join in SQL:

```
SELECT…
FROM table1 AS t1, table2 AS t2, table3 AS t3
WHERE t1.attr1=t2.attr2 AND t2.attr2=t3.attr3
```

Or:

```
SELECT…
FROM (table1 AS t1 INNER JOIN table2 AS t2 ON
t1.attr1=t2.attr2) INNER JOIN table3 AS t3 ON
t2.attr2=t3.attr3
WHERE…
```

The comparison operator:

- "=" is by far the most common comparison operator in join conditions, but
- Any of the following operators can also be used: `<>`, `>`, `>=`, `<`, `<=`, `BETWEEN`, or `LIKE`

Number of join conditions in more than two relations involved in a query – general rule:

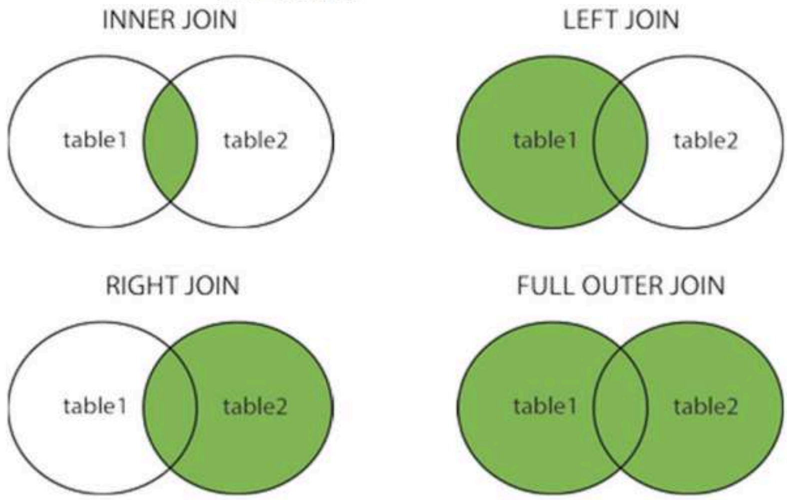
N input relations → (*N-1*) join conditions:

The INNER JOIN is a common operator in database querying, but it is not the only possible type of JOIN.

There are other options known as **OUTER JOINs**:

- **LEFT JOIN**
- **RIGHT JOIN**
- **FULL JOIN**

OUTER JOINs



SPATIAL QUERIES

BRIEF INTRODUCTION

- In a GIS or SDBMS:

Built in functions based on topological rules and standards.

Spatial Selections

```
SELECT f.*
FROM Factories AS f, Railway AS r
WHERE  ST_DISTANCE (f.geom,r.geom) < 200
```

Spatial Join

```
SELECT r.name, r.length, s.name, s.area
FROM Rivers AS r, Regions AS S
WHERE  ST_INTERSECTS(r.geom,s.geom)
```

SPATIAL QUERIES

BASIC SPATIAL FUNCTIONS

Location

ST_X — Returns the X coordinate of a Point.
ST_Y — Returns the Y coordinate of a Point.

```
SELECT ST_Y(geom)
FROM vector.bi_settlements
WHERE gid='300'
```

Distance

ST_Distance — Returns the distance between two geometry or geography values.

```
SELECT ST_Distance((a.geom),(b.geom))
FROM vector.bi_settlements a, vector.bi_settlements b
WHERE a.gid='300' AND b.gid='302'
```

Length/Area

ST_Area — Returns the area of a polygonal geometry.

ST_Length — Returns the 2D length of a linear geometry.

SPATIAL QUERIES

TOPOLOGY FUNCTIONS

ST_Contains(geometry A, geometry B): no points of B lie in the exterior of A, and at least one point of the interior of B lies in the interior of A.
ST_Crosses(geometry A, geometry B): Geometries have some, but not all, interior points in common.
ST_Disjoint(geometry A , geometry B): Geometries do not share any space together.
ST_DWithin(geometry A, geometry B, radius): geometries are within the specified distance (radius) of one another.
ST_Intersects(geometry A, geometry B): Geometries share a portion of space
ST_Overlaps(geometry A, geometry B): Geometries share space but are not completely contained by each other.
ST_Touches(geometry A, geometry B): Geometries have at least one point in common, but their interiors do not intersect.
ST_Within(geometry A , geometry B): Geometry A is completely inside geometry B

```
select a.name_com, a.geom, b.geom
from vector.bi_communes a join vector.bi_provinces b
on ST_Contains (b.geom, a.geom)
```

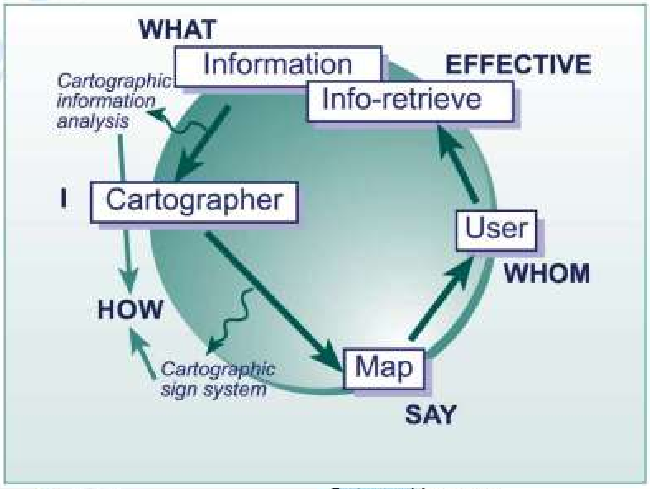
ST_Difference — Returns a geometry representing the part of geometry A that does not intersect geometry B.
ST_Intersection — Returns a geometry representing the shared portion of geometries A and B.
ST_SymDifference — Returns a geometry representing the portions of geometries A and B that do not intersect.
ST_Union — Returns a geometry representing the point-set union of the input geometries.

```
select ST_Union (a.geom, b.geom)
from data.table_1 a, data.table_2 b
```

Visualisation principles

"How do I say what to whom, and is it effective?"

(prof. Koeman, 1971)



"Rules of thumb" for effective graphic communication

Jacques Bertin: Sémiologie Graphique.
(Mouton, Paris/Den Haag, 1967)
First to systematically describe "how graphics work":

How the human mind finds meaning in graphical symbols

Perception

The visual variables enable observers to perceive:

- association (what belongs together)
- order
- quantities

Qualitative

Differences between data are purely descriptive and textual (phenomena attributes as: gender, land use, language, religion, geology, ...).

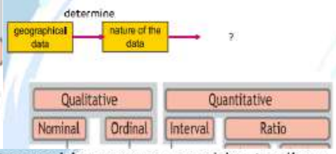
Nominal
The data are different in nature or identity, but all information is of the same order.

eg. landuse: forest, built-up, lake, grassland, industry
eg. network: river, road, railroad, powerline
eg. points of interest: busstop, campsite, watchtower

Ordinal
The data can be ordered or ranked, but this order or ranking cannot be defined in quantitative terms.

eg. vegetation height: high, medium, low
eg. roads: primary, secondary, tertiary, track
eg. settlements: city, village, hamlet

Cartographic process



Cartographic process: graphics toolbox

What kind of graphics can I use?

- point symbols
- line symbols
- area symbols
- text

- elements can all vary in *appearance*
- elements *position* can vary
  they are *visual variables*

Associative	Ordered	Quantitative
The symbols look different Ranking in order is not possible	The symbols express a specific order	The symbols express quantities

→ this process is pre-attentive (unconscious, 'automatic')

Quantitative

The data can be counted or expressed numerically. It is objective data produced through a systematic process that is verifiable, replicable and in and of itself is not subject to interpretation.

Interval
The data can be ranked, and the differences between the data does make sense.
There is an arbitrary zero.

eg. temperature: 18°, 22°, 26°, 30°
eg. measurement of time: Year of construction, 1200 After Christ, 100 Before Christ
eg. altitude: 100 meter, 200 meter, 300 meter, 400 meter

Ratio
Quantitative data based upon an absolute zero.

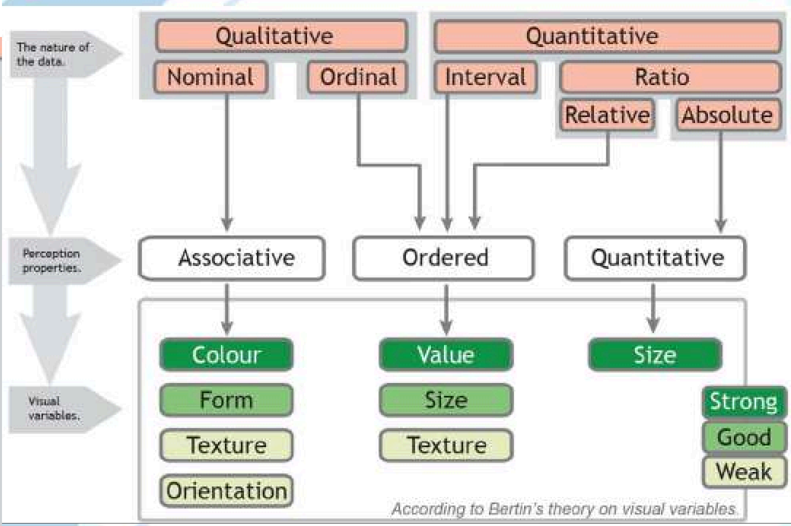
Relative
Values derived from "raw" data. Generally an absolute quantity divided by some other quantity (percentage, rate, ratio or density).

eg. population density: persons per km2, 1-10, 11-30, 31-70, 71-124 (can be calculated)
+ averages

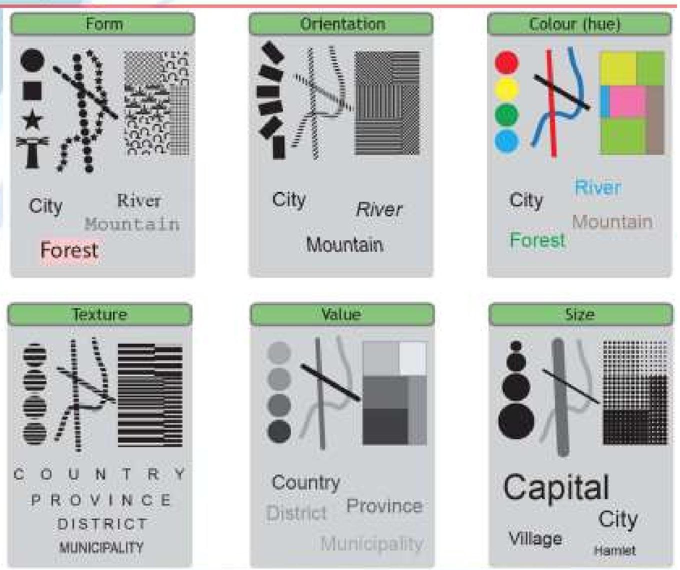
Absolute
The "raw" data, quantities concerning a single class of features, expressed in absolute terms according to some measurements scale.

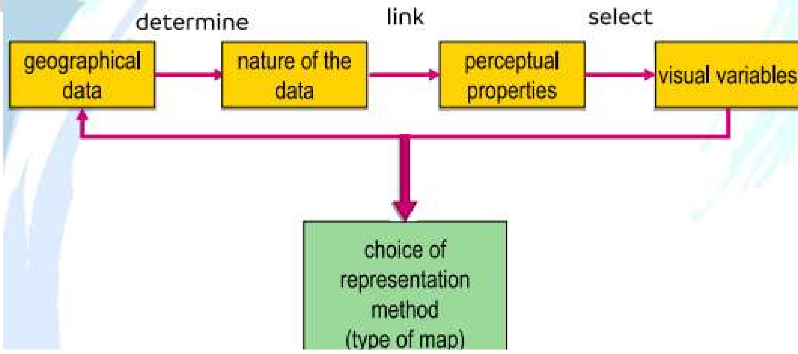
eg. number of inhabitants: x 1 milj, 1, 3, 5 (can be measured or counted)

Our link between Data & Visual variables



cartographic toolbox: "Visual Variables





Graphics are used in order to understand things easily. They pack a lot of meaning in a small thing. Graphic show patterns. Point symbols. Line, area symbols and text are used as graphics.

Cartographic - are visualisation of geo data. Keep in mind the user that the map is being made for.

Cartographic process - there are two types of data in this process.

- **Qualitative data** - it is the difference in the quality of the data but you can't put quantities. Like forest.
- **Quantitative data** - the data that has a certain amount of quantity like population.

1. Relative quantitate data - it is a relative type of data which has no exact values such as population, topography and etc.
2. Absolute quantitative data - it has an absolute value or a fixed value.

Different wave length means different colours.
Visual variables - It depends on form, orientation, colour(hue), texture, value and size.
Perception- it is how people look at certain things. Association, ordered and quantitative are three variables of perception.
Associative - position, colour, form, orientation, texture are associative meaning you can see the differences very clearly.
Dissociative - Value and size.
Ordered - we can see an order or direction, value and size. Form, orientation and colour are not ordered.
Quantitative - it tells the size, position and size work. Form, orientation, colour, texture, position are not quantitative.
Topographic Map - Accurate representation of earths topography.
Thematic maps - the maps that have a theme.

1. For a landuse map, the following categories are classified: - Urban area, Industrial area, Agricultural area, Nature reserves, orientation colour size form texture are the visual variables could be used to depict these data in such a way that a correct perception is achieved

Choice of map type is almost endless

- Topographic maps:
  - accurate representation of Earth's topography
- Thematic maps:
  - one or more particular themes are emphasized
    - Nominal maps
    - Proportional Point maps
    - Flow Maps
    - Choropleth maps
    - 3D-maps
    - isoline maps
    - cartograms
    - schematic maps

- Topographic maps:
  - accurate representation of Earth's topography

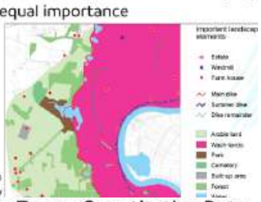
Choice of map type

- Thematic maps:
  - one or more particular themes are emphasized
  - always some topographic content (framework data)
  - several sub-kind of maps (thematic map types)

Thematic Map Types: Qualitative Data

need to stress the nominal differences while keeping the perception of equal importance

- for points, lines, areas, or a combination
- use colour, form, orientation, or a combination
- use colours and symbols of similar importance (value) to avoid perception of order or quantity



Thematic Map Types: Quantitative Data

Absolute data: need to stress the proportional differences of the quantities

- for points (which can represent areas)
- using size

Proportional point map

Thematic Map Types: Quantitative Data

Absolute data: need to stress the proportional differences of the quantities

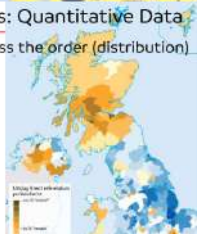
- for lines (which can represent flows)
- using size

Proportional line map (flow map)

Thematic Map Types: Quantitative Data

Relative data: need to stress the order (distribution)

- the order can be bi-polar
- using diverging value



Divergent or (bi-polar) choropleth map

Thematic Map Types: Quantitative Data

Relative data: need to stress the order (distribution)

- the data is normalized: (made) relative to eg. area (per km^2), totals (%), per inhabitant, etc.
- using value

Choropleth map

Thematic Map Types: Quantitative Data

Relative & absolute: both show a different side of the data

- these can be combined

to tell "the story of the data" !

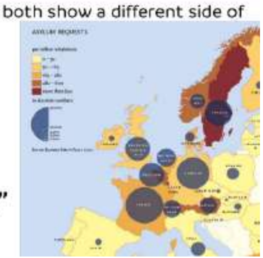


Image visualisation and colour composites

Sensitivity of cones and rods - there are three cones that measure the incoming light in the eye, red, green and blue cones. They need light to work hence they work the most at day time.

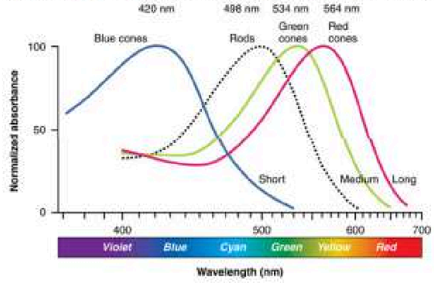
Colour system - the bigger the distance the brighter the colour, **RGB cube** has 8 different colour on each corner of the cube, **Intense hue system(IHS)** is the colour gradient between the 8 points on a RGB cube.

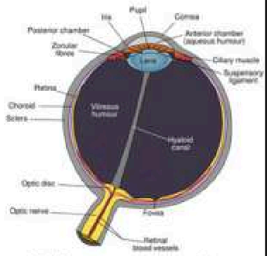
Additive system - in this you start from black.
Subtractive system - here you start from white.
Raster data set contains one or more layer containing measurement in different portions of EMR.

Colour composites - There are multiple spectral bands, we need to select 3 bands on the colour, there are classification to the colour composite.

SENSITIVITY OF CONES AND RODS

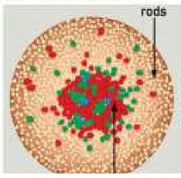
https://upload.wikimedia.org/wikipedia/commons/9/94/1416_Color_Sensitivity.jpg





http://www.macroevolution.net/images/human-eye-373-358-20.jpg

- Cones need high intensities to work well
  - Three types of cones exist with different sensitivity (approx. Red-L, Green-M and Blue-S)
- Rods work well in low illumination, but show only brightness



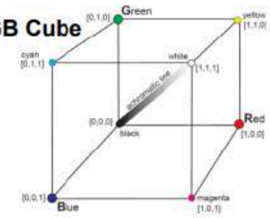
TRI-STIMULI MODEL FOR COLOUR VISION

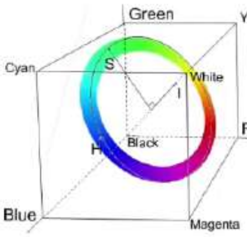
- Three kinds of cones
- Three layers on colour film
- Red-Green-Blue dots on monitors
- Magenta-Yellow-Cyan (inks) for printing
- Colour cube/spaces

- Sensation to the human has stronger relation to Intensity-Hue-Saturation (Brightness-Colour-Vividness) than to Red-Green-Blue

COLOUR SYSTEMS (DIS)AGREEMENT

RGB Cube





IHS (HLS)

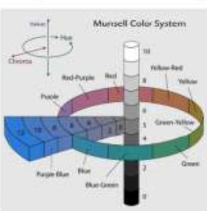
DIN 6164

Ostwald

https://upload.wikimedia.org/wikipedia/commons/thumb/b/b4/Ostwald_Color.jpg/800px-Oswald_Color.jpg

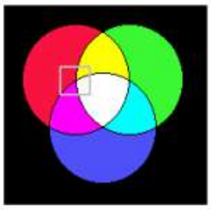
Munsell

https://upload.wikimec mons/thumb/d/d5/Mur



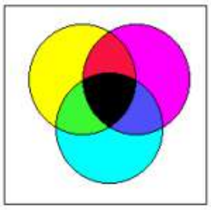
n.com/wp-
din6164.jpg

UNIVERSITY OF TWENTE.



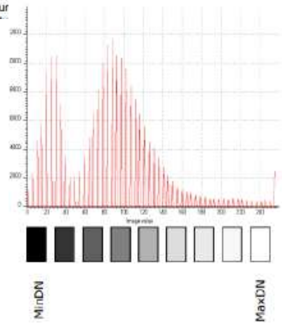
Additive Color Scheme

Screen display



Subtractive Color Scheme

Printing

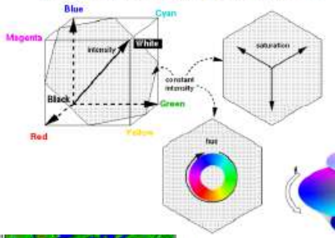


TERMINOLOGY

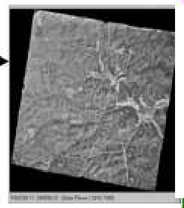
- Raster data sets contain one or more layers
- EO images (⊂ raster data sets) contain one or more layers containing measurements in different portions from the EM spectrum. Also referred to as (spectral) bands.
- Monitors build up display using one up to three channels from Red, Green and Blue
- A measurement is stored as a Digital Number (DN) also referred to as value (and tautology DN value

IMAGE DISPLAY

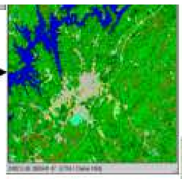
IMAGE DISPLAY; PSEUDO COLOUR (HUE)



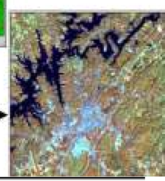
One band → Grey scale



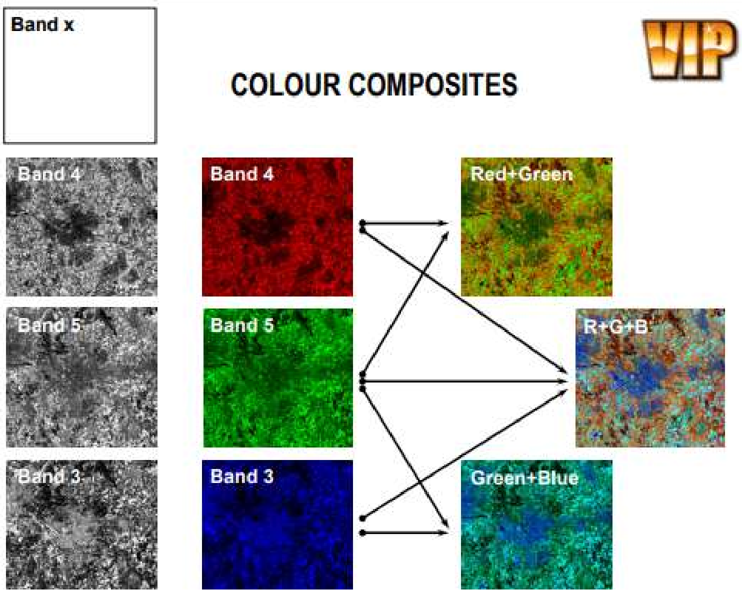
One band → Pseudo colour table (e.g. classification results)



Multiple bands → Select 3 → Colour composites (e.g. false colour comp.)

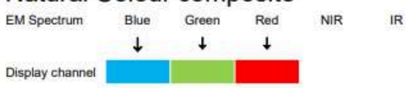


COLOUR COMPOSITES

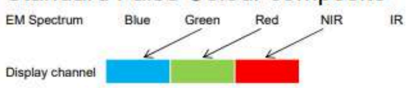


COLOUR COMPOSITES CLASSIFICATION

- Natural Colour composite



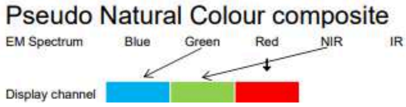
- Standard False Colour composite



healthy vegetation = reddish

- All other combinations
  Selecting prominent bands to interpret specific object classes

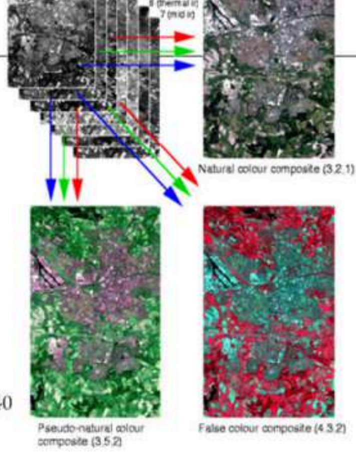
- Pseudo Natural Colour composite



COLOUR COMPOSITES

- TM (6 bands)= $\left(\frac{6!}{(6-3)!}\right) = \left(\frac{720}{6}\right) = 120$
- Sentinel 2 (13 bands)= $\left(\frac{13!}{(13-3)!}\right) = 1716$
- AVHRR (220 bands)= $\left(\frac{220!}{(220-3)!}\right) = 10503240$

$Permutations = \left(\frac{n!}{(n-k)!}\right) = \#\ of\ Colour\ Composites$

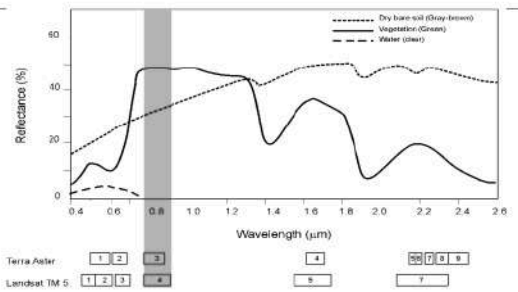


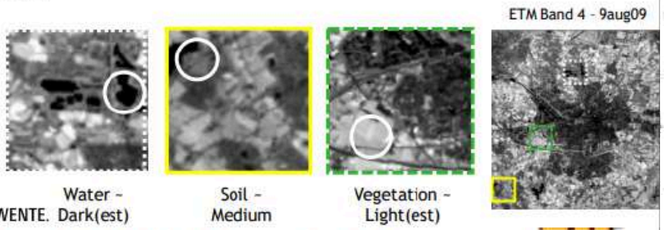
BAND SELECTION IS ESSENTIAL

- Analyse spectral char. of objects of interest with respect to (char. of) bands available
- Don't select correlated bands
- Consider User/Usage (experience of user, natural versus false colour, application)

VISUALISATION OF MAIN COVER TYPES

- Identify types of objects
- Find signatures
- Find sensor characteristics
- Translate signature into relative brightness for display



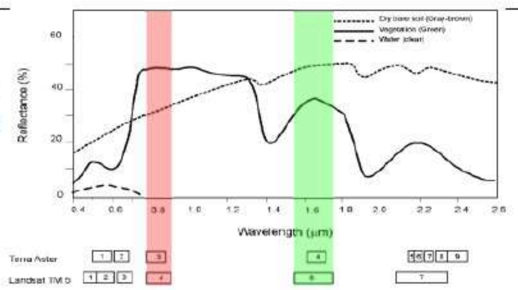


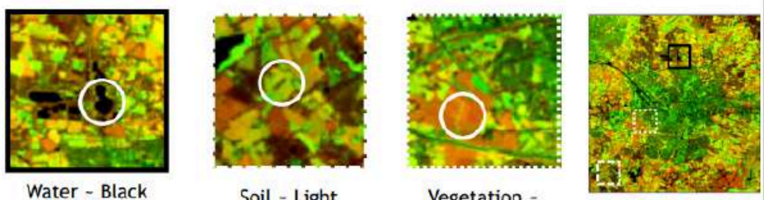
Water - Dark(est) | Soil - Medium | Vegetation - Light(est)

UNIVERSITY OF TWENTE.

VISUALISATION OF MAIN COVER TYPES

- Same applies to colour composites
- E.g. Band 4 in Red and band 5 in Green





Water - Black | Soil - Light Greenish | Vegetation - Orange | ETM - 9 Aug 09

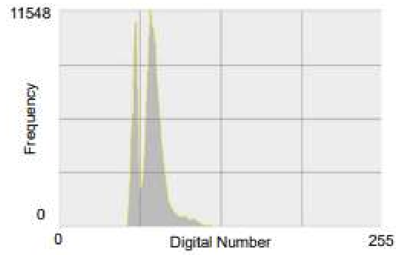
UNIVERSITY OF TWENTE.

ELEMENTARY IMAGE ENHANCEMENT

- Enhance the image for a specific purpose through histogram operations
- Global contrast enhancement

SETTING THE SCENE

- Image is subset of the World
  - Energy range measured in an image is a subset of the dynamic range of sensor -> subset of 8 bit, 10 bit . . .

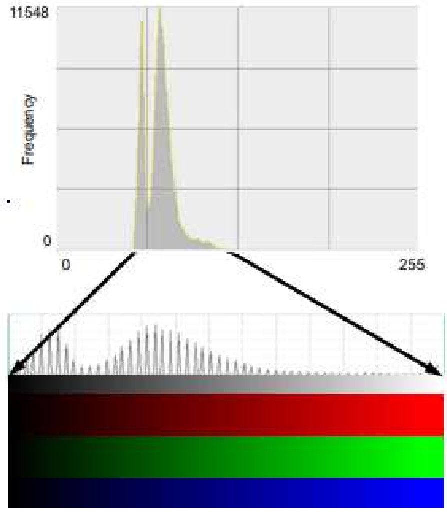


- Monitor can display $2^8 = 256$ shades of grey

ENHANCEMENT BY HISTOGRAM OPER.

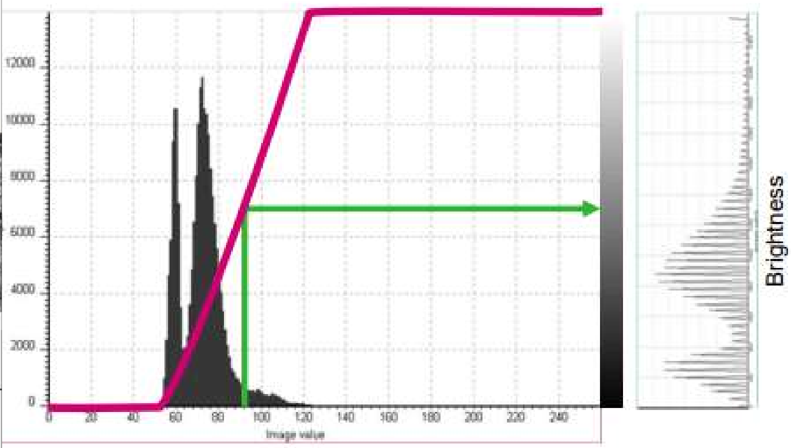
- One way or another 'map' DN value to
  - Grey value

  or
  - Red
  - Green
  - Blue



UNIVERSITY OF TWENTE.

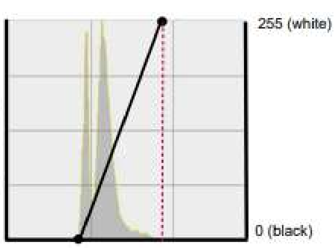
TRANSFER FROM DN TO BRIGHTNESS



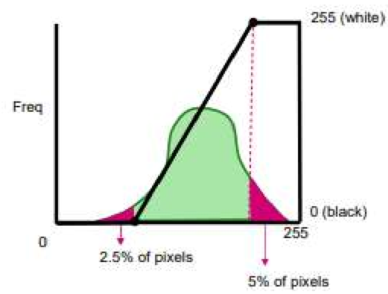
HOW TO SET EFFICIENT CONTRAST (WOOCLAP)

Use existing DN values

- Minimum input value and maximum input value for setting the Transfer Function



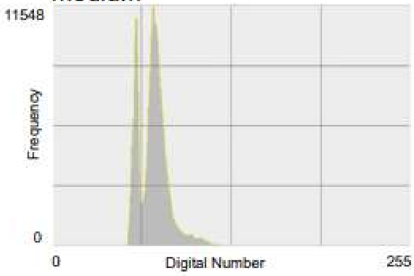
- Use % to cutoff (eg 2.5/5)
  - Assumes lowest 2.5 % and highest 5 % are obsolete



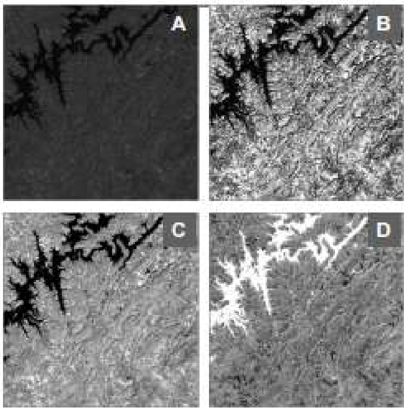
HOW TO SET EFFECTIVE CONTRAST?

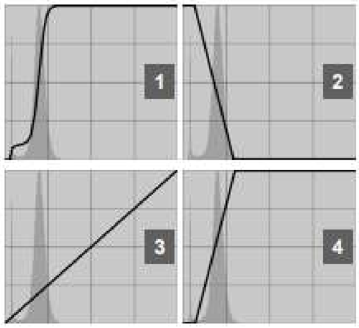
- Assess expected objects, their signatures and thus representation in the (histogram of the) spectral band at hand
- Set contrast to most interesting ranges (objects)

- SPOT B1 – Green
- Expected objects: water, various crops and forest on land
- Expected DN: water very low, forest low, crops low to medium

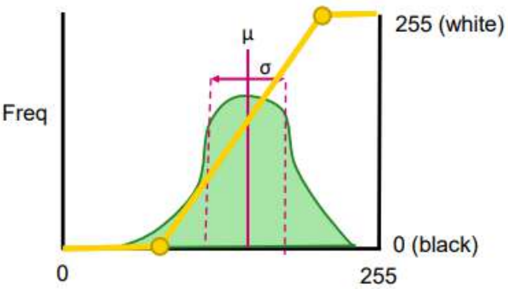


UNIVERSITY OF TWENTE.





ALTERNATIVE – STANDARD DEVIATIONS (WOOCLAP)



Based upon
- Mean (μ)
- +/- 2 Standard Deviation (σ)

Nice start **BUT**

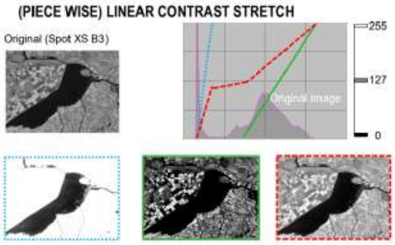
Assumes Normal or Gaussian distribution

STRETCHING METHODS (FUNCTIONS)

- Linear
- Piece wise linear
- Histogram equalisation
- Gamma
- . . .
- . . .

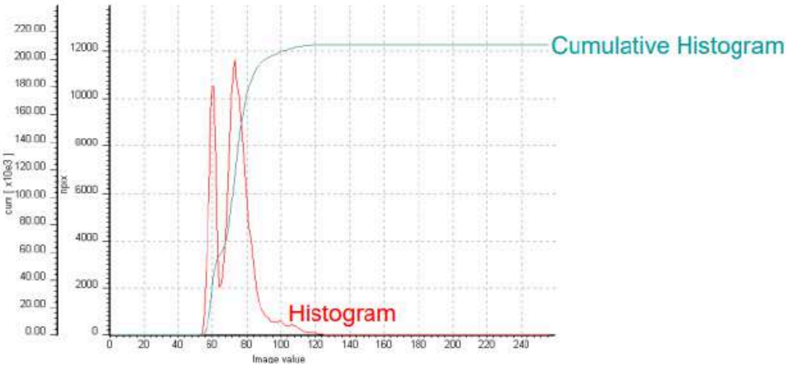
All transferring DNs to brightness

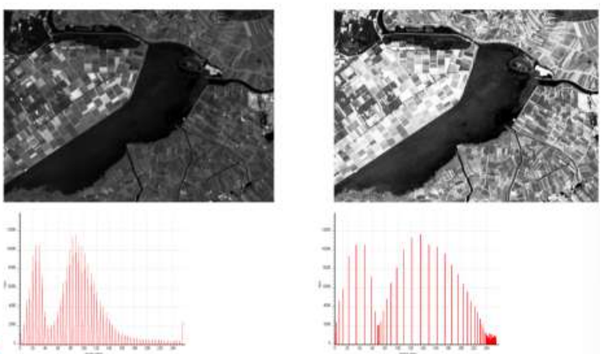
(PIECE WISE) LINEAR CONTRAST STRETCH



HISTOGRAM EQUALISATION

- Transfer function is the cumulative histogram (nonlinear)
  - Steeper transfer function at higher frequencies!





Conceptual models are...
approximate processes in the real world. perfect models of real-world processes. closely related to a specific point of view.

Which property is identical for all types of electromagnetic radiation? - Frequency

For electromagnetic waves, c=λν and Q=hν=hc/λ
This means that radiation with a small wavelength has a higher frequency, while a large wavelength is associated with a lower frequency. Moreover, because blue light has a larger wavelength than red light, the energy of "blue" photons is smaller than "red" photons.
higher
Lower
Smaller wavelength
Smaller than

correct sequence of ranges of EMR, ordered from smallest wavelength to largest.
ultraviolet-visible-near infrared-thermal infrared-microwave

EMR wavelength ranges on the left to the examples of applications on the left.
Visible - Identification of objects and structures.
Microwave - Quantification of surface roughness and soil moisture contain.
Near infrared - Detection of vegetation
UV - Observation of atmosphere and certain minerals
Thermal infrared - Determination of surface temperature

Q - Energy of photon(joules)
H = planks constant(joules per sec)
Lamda = wavelength(m)
c = speed of light (m*sec^-1)

Scanning, map registration, skeletonizing, extracting vectors, data clean up, feature forming - steps to fully automatic digitizing a paper map

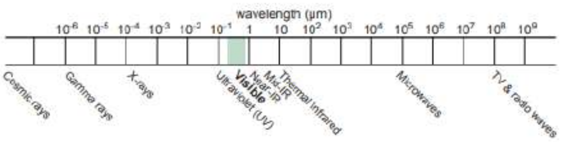
Spatial coverage = area covered by image
Temporal resolution = revisit time
Dynamic change = max and min
Spectral coverage = total wavelength range

If the band cannot discriminate the difference between two covers having high spectral resolution it is due to a limited radiometric resolution.
To study the growth cycle of the crop we need data with higher temporal resolution.
To discriminate between the two land cover we need measurements in different parts of the spectrum therefore a sensor with radiometric resolution is required.
Detection of trees to be able to count how many there are requires a higher spatial resolution.
1- simplices only intersect

Two lines, representing the center lines of two roads, cross each other without being split at the intersection.

2.3 Electromagnetic spectrum

We call the total range of wavelengths of EM radiation the *EM spectrum*. Figure 2.2 illustrates the spectrum of visible light; Figure 2.5 illustrates the wider range of EM spectrum. We refer to the different portions of the spectrum by name: gamma rays, X-rays, UV radiation, visible radiation (light), infrared radiation, microwaves, and radio waves. Each of these named portions represents a range of wavelengths, not one specific wavelength. The EM spectrum is continuous and does not have any clear-cut class boundaries.



1. Conceptual models are perfect models of real-world processes.
2. Frequency y is identical for all types of electromagnetic radiation.
3. For electromagnetic waves, c=λv and Q=hv=hc/λ This means that radiation with a small wavelength has a higher frequency, while a large wavelength is associated with a lower frequency. Moreover, because blue light has a larger wavelength than red light, the energy of "blue" photons is smaller than "red" photons.
4. Ordered from smallest wavelength to largest ultraviolet-visible-near infrared-thermal infrared-microwave.
5. EMR wavelength ranges and application -
Visible - Identification of objects, microwave - quantification of surface, Near infrared - detection of vegetation, ultraviolet - observation of atmosphere, thermal infrared - determination of surface.
6. Planks law of radiation describes, radiative flux of a (black) body of a certain temperature, C = the speed of light, h = Planck's constant, k = the Boltzman constant, λ = wavelength, T = absolute temperature.
7. The Sun can be considered a blackbody. Its temperature is high , therefore the radiative flux is larger and the maximum wavelength of emission small compared to the Earth. The Sun's maximum emission is in the visible range, whereas Earth's emission maximum is in the infrared range. The emission spectra of Sun and Earth barely overlap.
8. We perceive objects as red when they primarily reflect visible radiation in the red wavelength range. Near-infrared radiation cannot be seen by humans. In remote sensing, infrared radiation is used to monitor vegetation . The earth is a large source of thermal infrared radiation.
9. Several wavelength ranges of the radiative spectrum are not suitable for satellite observations of the surface, due to the absorption of radiation by water and carbon dioxide. The ranges that can be used for Earth Observation are called window regions because the atmospheric transmitted is high and a lot of radiation reaches the surface.
10. The amount of scattering depends on the ratio between the wavelength of radiation and the size of the object. Which type of scattering will green light undergo if it encounters objects that are Rayleigh much smaller than the wavelength; mie the same order of magnitude in size as the wavelength; non-selective much larger than the wavelength.
11. Kirchhoff's law, , means that in thermal equilibrium (object's temperature does not change) absorption and emission are equal.

The boundary of a polygon that represents an administrative unit (municipality) has a small opening (gap).

Every 2-simplex has a closed boundary consisting of an alternating (cyclic) sequence of 0- and 1-simplices.

There is a small gap between two polygons that represent two administrative boundaries (for example - municipalities in the Netherlands) that are neighbours.

Every 1-simplex borders two 2-simplices (polygons, namely the 'left' and 'right' polygon)

radiometric units

We can use different measures to quantify radiation. The amount of radiative energy is commonly expressed in joules (J). We may, however, be interested in the radiative energy per unit of time, called the *radiant power*. We measure the power in watts ($W = J\ s^{-1}$). *Radiant emittance* is the power emitted from a surface; it is measured in watts per square metre ($W\ m^{-2}$). *Spectral radiant emittance* characterizes the radiant emittance per wavelength; it is measured in $W\ m^{-2}\ \mu m^{-1}$ (this is the unit used in Figure 2.4). *Radiance* is another quantity frequently used in RS. It is the radiometric quantity that describes the amount of radiative energy being emitted or reflected in a specific direction per unit of projected area per unit of solid angle and per unit of time. Radiance is usually expressed in $W\ sr^{-1}\ m^{-2}$ (sr is steradian, unit of solid angle). *Spectral radiance* used in Equation 2.4 is radiance per wavelength and is measured in $W\ sr^{-1}\ m^{-2}\ \mu m^{-1}$. *Irradiance* is the amount of incident radiation on a surface per unit of area and per unit of time. Irradiance is usually expressed in $W\ m^{-2}$.

Planck's law of radiation is only applicable to black bodies. A black body is an idealized object with assumed extreme properties that helps us when explaining EM radiation. A black body absorbs 100% of incident EM radiation; it does not reflect anything and thus appears perfectly black. Because of its perfect absorptivity, a black body emits EM radiation at every wavelength (Figure 2.4). The radiation emitted by a black body is called black-body radiation. Real objects can re-emit some 80 to 98% of the radiation received. The emitting ability of real objects is expressed as a dimensionless ratio called emissivity $\epsilon(\lambda)$ (with values between 0 and 1). The *emissivity* of a material depends on the wavelength; it specifies how well a real body made of that material emits radiation as compared to a black body.

Wien's displacement law

The Sun behaves similarly to a black body. It is a prime source of the EM radiation that plays a role in Earth Observation, but it is not the only source. The global mean temperature of the Earth's surface is 288 K and over a finite period the temperatures of objects on the Earth rarely deviate much from this mean. The surface features of the Earth therefore emit EM radiation. Solar radiation constantly replenishes the energy that the Earth radiates into space. The Sun's temperature is about 6000 K. Planck's law of radiation is illustrated in Figure 2.4 for the approximate temperature of the Sun (about 6000 K) and the ambient temperature of the Earth's surface (288 K). The figure shows that for very hot surfaces (e.g. the Sun), spectral emittance of a black body peaks at short wavelengths. For colder surfaces, such as the Earth, spectral emittance peaks at longer wavelengths. This behaviour is described by *Wien's displacement law*:

$$\lambda_{max} = \frac{b}{T}, \qquad (2.5)$$

where λ_{max} is the wavelength of the radiation maximum (µm), T is the temperature (K) and $b \approx 2898$ µm K is a physical constant.

black body

We can use Wien's law to predict the position of the peak of the black-body curve if we know the temperature of the emitting object. The temperature of the black body determines the most prominent wavelength of black-body radiation. At room temperature, black bodies emit predominantly infrared radiation. When a black body is heated beyond 4450 K (approximately 4700 °C) emission of light becomes dominant, from red, through orange, yellow, and cyan, (at 6000 K) to blue, beyond which the emitted energy includes increasing amounts of ultraviolet radiation. At 6000 K a black body emits radiation of all visible wavelengths in approximately equal amounts, creating the sensation of white to us. Higher temperatures correspond to a greater contribution of radiation of shorter wavelengths.

The following description illustrates the physics of what we see when a blacksmith heats a piece of iron or what we observe when looking at a candle. The flame appears