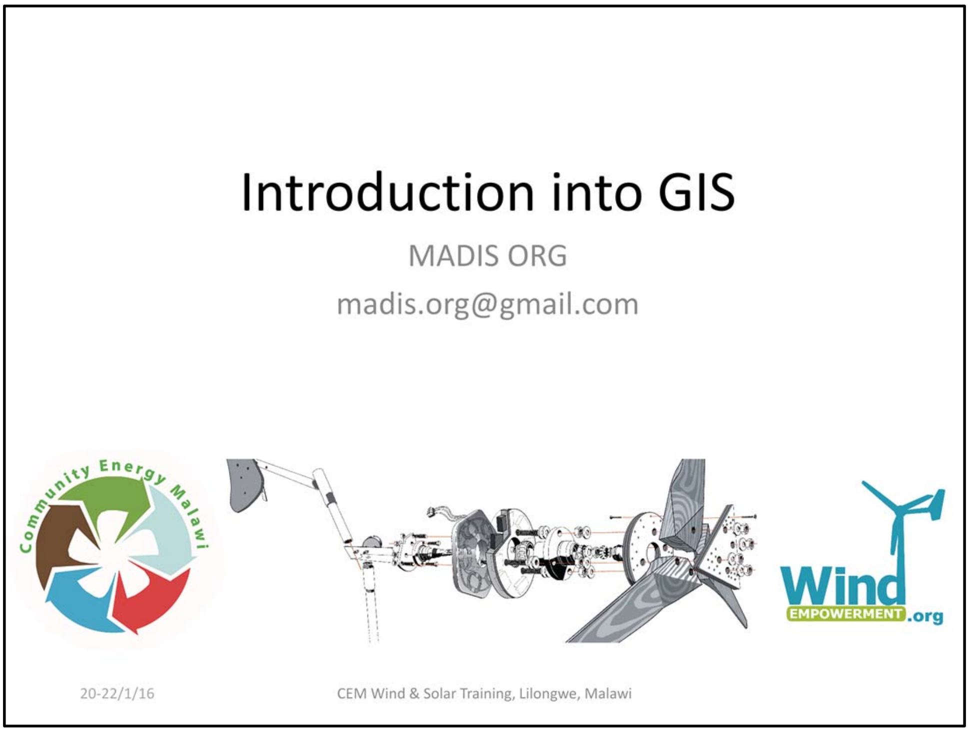


An information system has a full range of functions to:

- process observations
- process measurements
- provide descriptions
- explain data
- make forecasts
- make decisions

A GIS is an organized collection of computer hardware, software, geographic data, and personnel to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information.

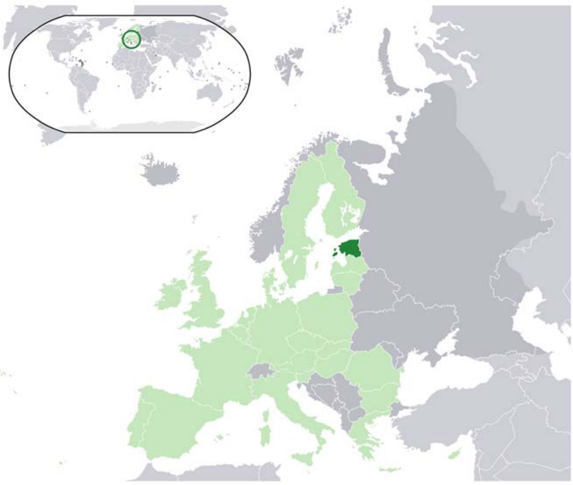
A GIS *integrates* spatial and other kinds of information within a single system to provide a consistent framework for analyzing geographic (spatial) data

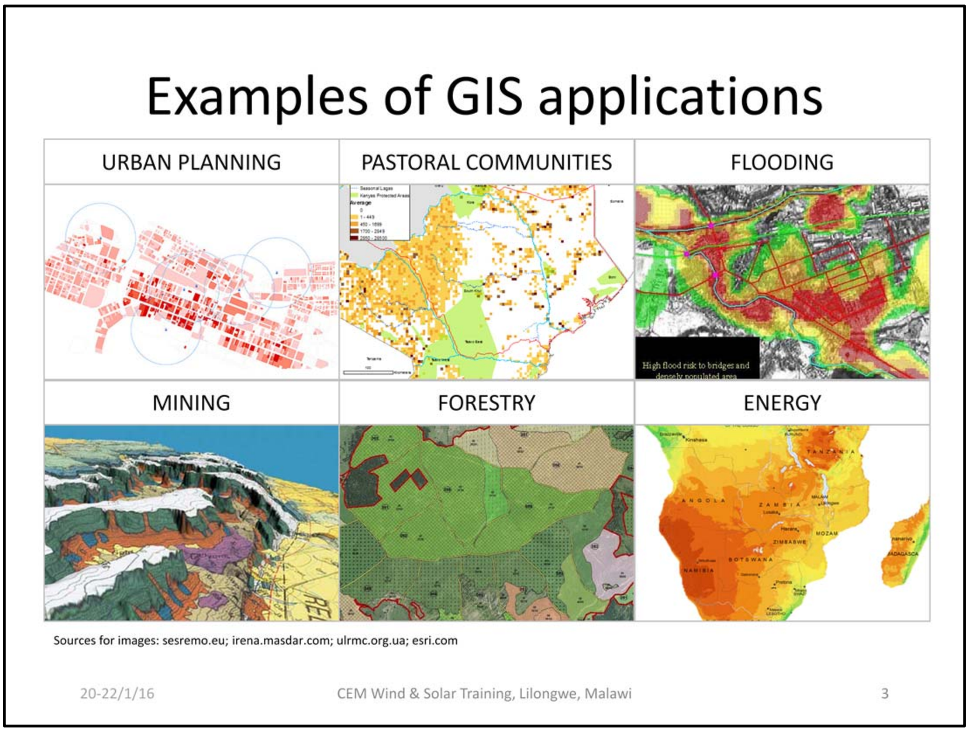
A GIS makes connections between activities based on *geographic proximity*.

The digital data structure can be conceptualized as a set of “floating electronic maps” with a common registration allowing the used to “look” down (drill down) and across the stack of maps.

The spatial relationships can be summarized (data base inquiries)

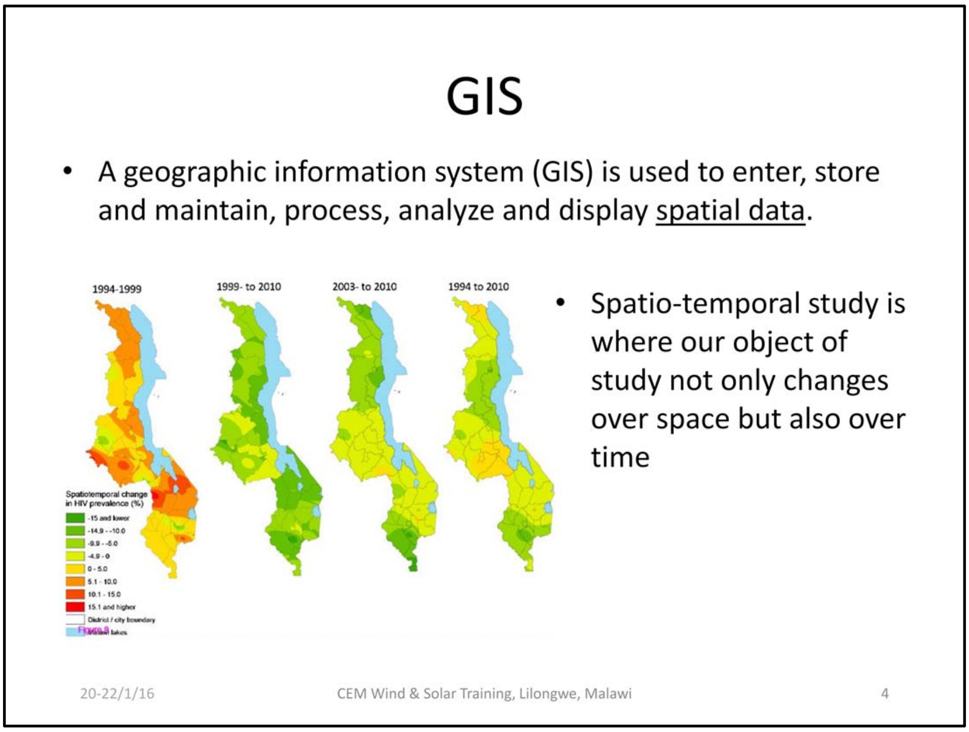
ESTONIA





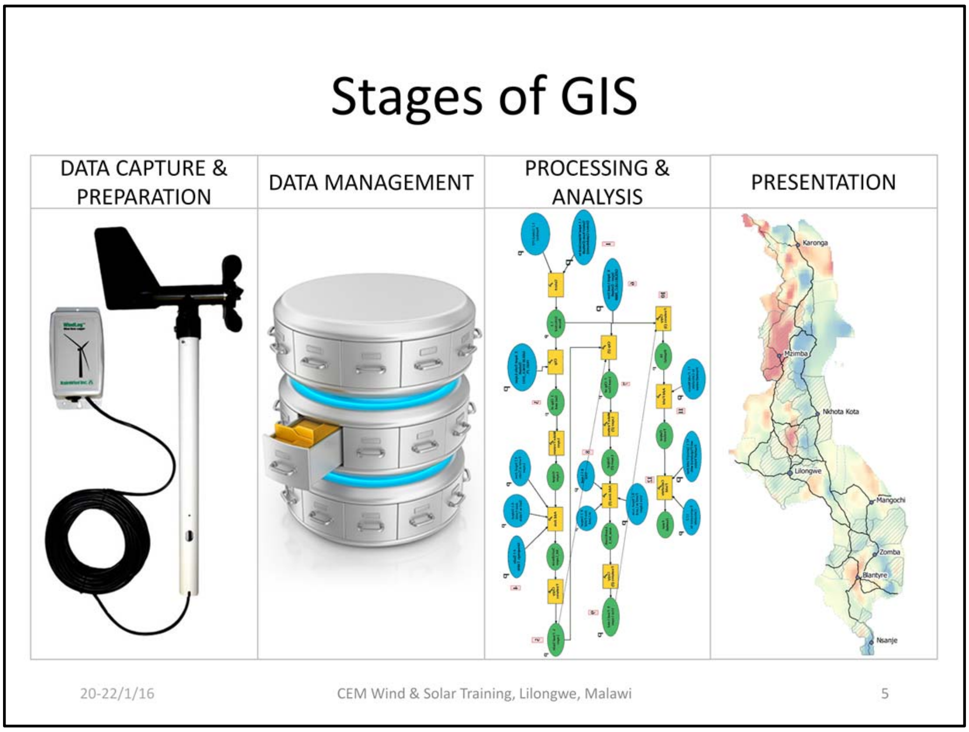
1) Urban planners need to make decisions in regard to the services that the public system has to provide – be it fire department, police, post offices or transportation services. All of these services have a spatial dimension – given a demand in a certain area, the supply has to meet the demand also in the same area.
2) Pastoralists have adapted to harsh conditions and use communal land systems to share risk during dry years. While dry areas are often underutilized, pastoralism is an effective use of ASOLs and provides valuable protein sources. Low populations of cattle have even been shown to even support higher biodiversity in dryland ecosystems. Kenya's Department of Remote Sensing and Resource Survey (DRSRS) monitors the numbers of large animals in 30-50% of the rangelands of the county. They use aerial surveys to photograph and count animal numbers.
3) Here is a map of a city center in Western Ukraine. GIS plays an important role in planning in coping with natural hazards like flooding or earthquakes. Based on mainly the elevation analysis, one can identify high risk areas. This information then will flow into property prices, insurance companies, real estate owners...who can take precautionary measures.
4) Colorado. With the aid of different kind of sensors, like hyperspectral sensors utilizing the near-infrared range of the spectrum, one can get an idea of whta is going on within the the ground, about the densities, materials, etc. This is very good in planning for prospective mining sites, but also for road planning, settlement planning, etc.

5) Norway: A forest manager might want to optimize timber production using data on soil and current tree stand distributions, in the presence of a number of operational constraints, such as the need to preserve species diversity in the area;
6) Energy. Atype of energy source, be it conventional or renewable, has a spatial dimension to it. For example in the oil business the same questions arise as in the case of mining, in case of biomass as an energy source, it is similar to the questions from forestry. As seen on this map of Southern Africa, solar irradiation across the continent may vary to a very big degree.

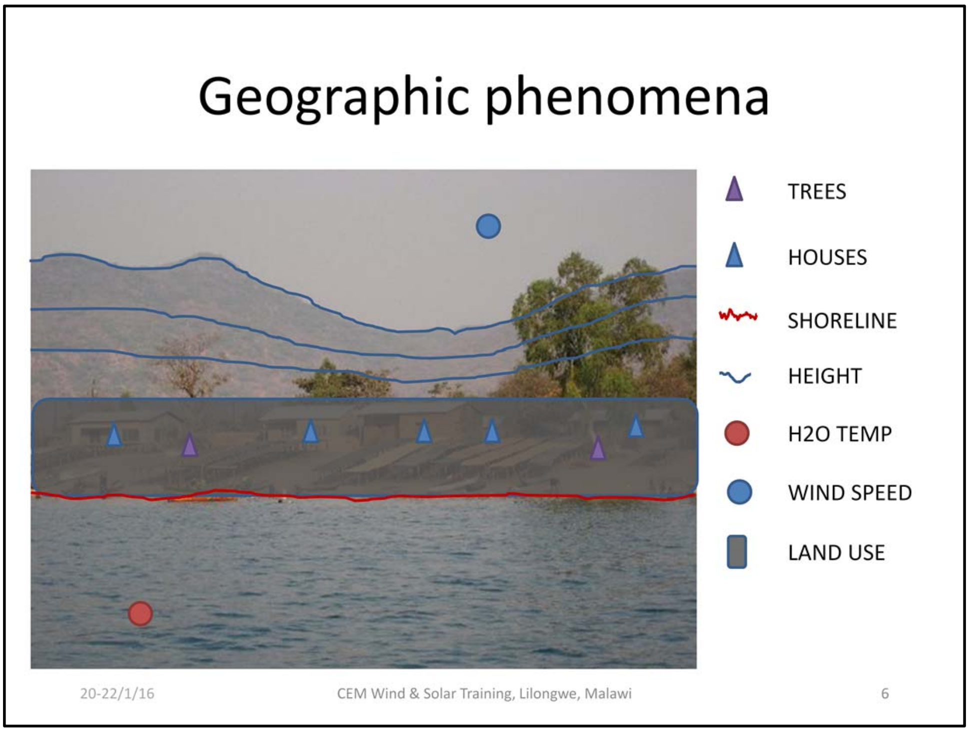


For examples of each of these capabilities, let us take a closer look at the El Nino˜ example. Many professionals closely study this phenomenon, most notably meteorologists and oceanographers. They prepare all sorts of products, such as the maps of Figure 1.1, in order to improve their understanding. To do so, they need to obtain data about the phenomenon, which, as shown above, includes measurements about sea water temperature and wind speed from many locations. This data must be stored and processed to enable it to be analyzed, and allow the results from the analysis to be interpreted. The way this data is presented could play an important role in its interpretation.

This means that our object of study has different characteristics for different locations (the geographic dimension) and also that these characteristics change over time (Spatial and temporal dimension). The El Ni dimensions no event is a good example of such a phenomenon because sea surface temperatures differ between locations, and sea surface temperatures change from one week to the next.



1)
2) data management refers to the storage and maintenance of the data transmitted by the buoys via satellite communication. This phase requires a decision to be made on how best to represent our data, both in terms of their spatial properties and the various attribute values which we need
3) Once the data has been collected and organized in a computer system, we can start analysing it. Spatial interpolation, combination of inputs, averages, spatio-temporal analysis
4) After the data manipulations discussed above, our data is prepared for producing output. In this case, the maps of Figure 1.1. The data presentation phase deals with putting it all together into a format that communicates the result of data analysis in the best possible way: The message, The audience, The medium, The rules of aesthetics, The techniques

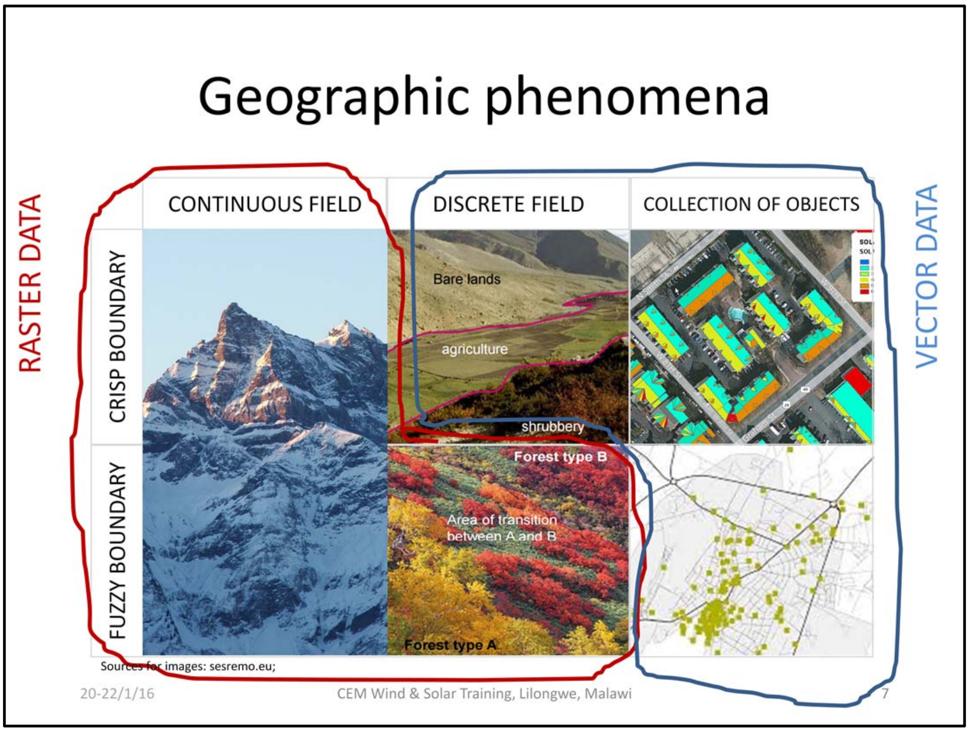


Geographic phenomena are the study objects of a GIS. Geographic phenomena exist in the real world, everything you see outside is a Geographic phenomenon. Some of the things you do not see are also Geographic phenomena like temperature We need to come up with a digital representations of the geographic phenomena in order to store them in a GIS. This is not easy because different phenomena require different digital representations and multiple representations are possible for the same phenomenon.

A digital representation is a model, it is not the real thing Our representation will never be perfect, some facts will not be found

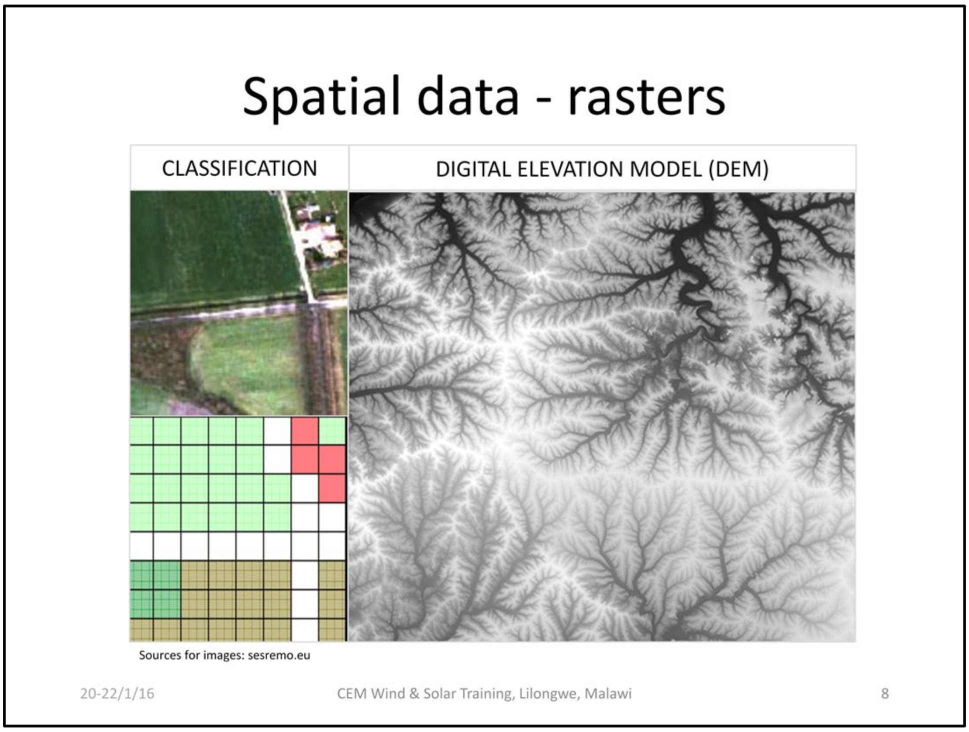
There are two groups of geographic phenomena, fields and objects: A (geographic) field is a geographic phenomenon for which, for every point in the study area, a value can be determined. (temperature, barometric pressure and elevation) (Geographic) objects populate the study area, and are usually well distinguishable, discrete, bounded entities. The space between them is potentially empty. (buildings, rivers)

There are two types of geographic fields, discrete fields and continuous fields In a continuous field, the underlying function is assumed to be continuous. Continuity means that all changes in field values are gradual. (for example elevation) Discrete fields cut up the study space in mutually exclusive bounded parts, with all locations in one part having the same field value. (for example landuse)



There are two types of geographic fields, discrete fields and continuous fields In a continuous field, the underlying function is assumed to be continuous. Continuity means that all changes in field values are gradual. (for example elevation) Discrete fields cut up the study space in mutually exclusive bounded parts, with all locations in one part having the same field value. (for example landuse)

Continuous means that all changes in field values are gradual In a differentiable field we can measure the change.
Forest plots form forests Parcels form blocks and blocks form suburbs Streams brooks and rivers form a river drainage system
We usually do not study objects in isolation (a single object) but whole collections of objects
As a general rule of thumb, crisp boundaries are more common in man-made phenomena
Fuzzy boundaries contrast with crisp boundaries in that the boundary is not a precise line, but rather an area of transition.
Computer representations can be divided in two groups: tessellations and vector-based representations. The next step is to understand how the computer representations can be applied to represent geographic fields and objects.

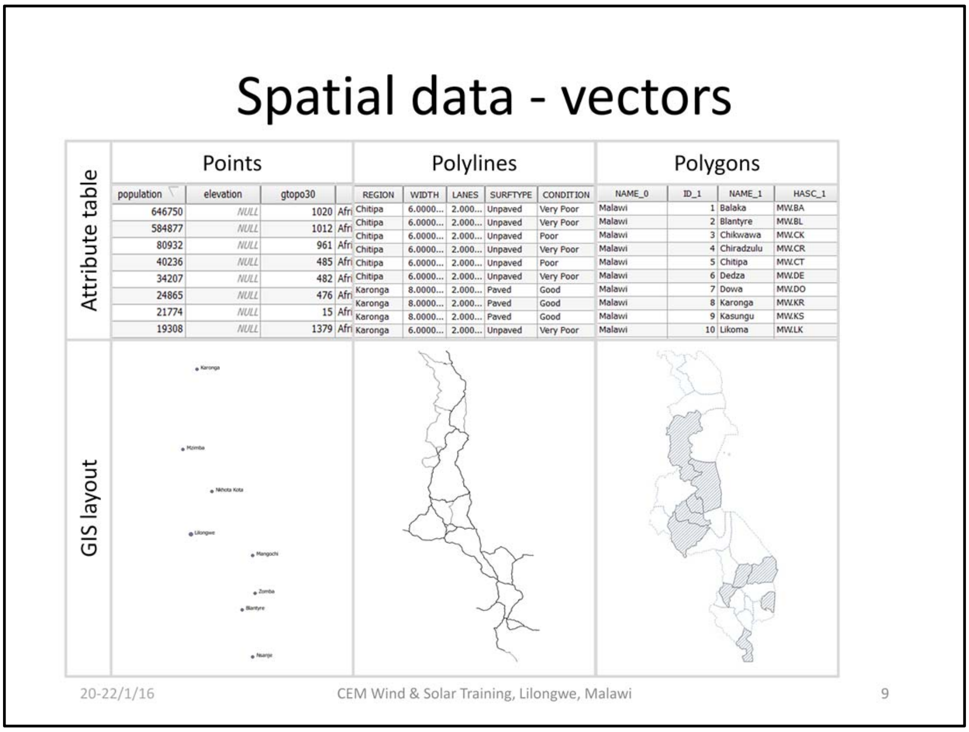


A tessellation is a partition of space into mutually exclusive cells that together make up the complete study area.

Some convention is needed to state which value prevails on cell boundaries

The size of the area that a single raster cell represents is called the raster's resolution

All regular tessellations have in common that the 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.

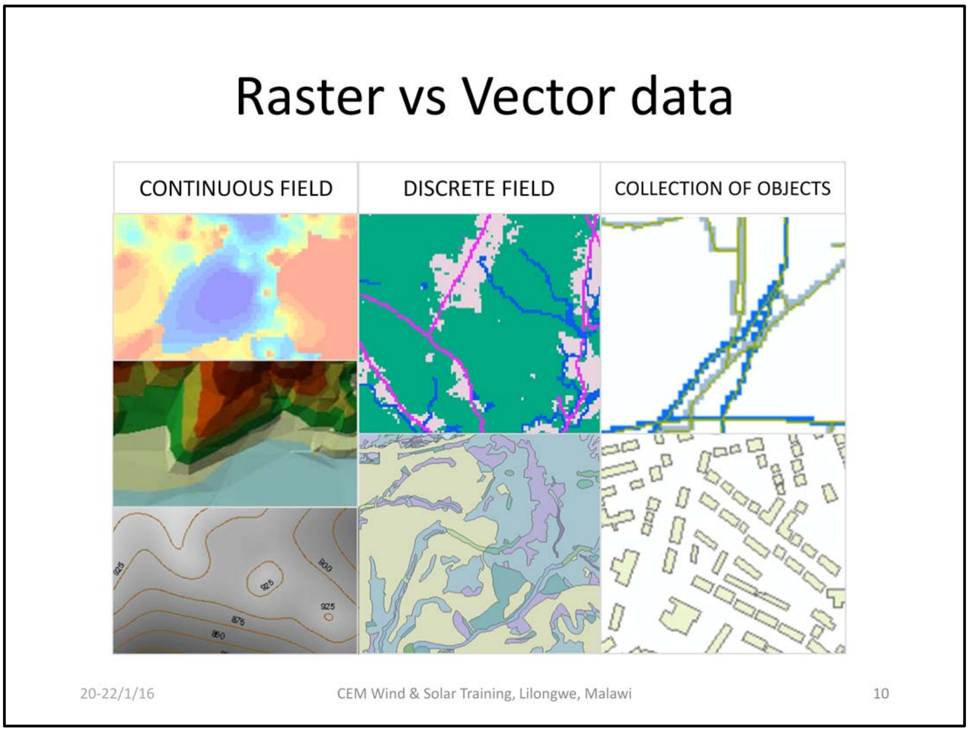


In vector representations a georeference is explicitly associate with the geographic phenomena. A georeference is a coordinate pair from some geographic space, also known as a vector.

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 best used to represent objects that are best described as shape- and size less single locality features.

Line representations: Used to represent onedimensional objects (roads, railroads, canals, rivers…) Line is defined by 2 end nodes and 0-n internal nodes to define the shape of the line. An internal node or vertex is like a point that only serves to define the line

 Area representations: When area objects are stored using a vector approach, the usual technique is to apply a boundary model. The area is defined by de boundary of the area



Continuous fields like elevation can be represented as a raster. The raster can be thought of as a long list of field values. You will see that field values are normally "floating point" values, and values change gradually.

In a TIN the amount of data stored is less compared to a regular tessellation. The quality of a TIN depends on the choice of anchor points, as well as the triangulation. Anchor points on elevation ridges are a guarantee for correct peaks and mountain slope faces.

An isoline is a linear feature that connects the points with equal value When the field is elevation we also speak of contour lines.

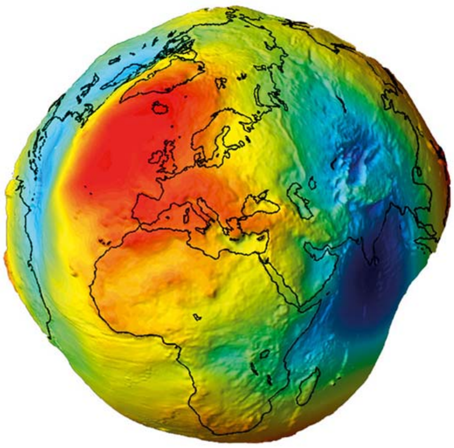
Discrete field represented as a raster. Raster cells have an integer value. Value is a classification, a numeric code representing for example a landuse class. Area boundaries may appear ragged.

Discrete fields represented as vector polygons. A field because they cover the full map extent (study area).
Area objects similar as in discrete fields, except "NoData" values may appear in between objects. Lines and point objects are awkward to represent in raster.

Raster formats are efficient when comparing information among arrays with the same cell size.
Raster representations are relatively coarse and imprecise
Raster files are generally very large because each cell occupies a separate line of data, only one attribute can be assigned to each cell, and cell sizes are relatively small.

Vector representations of shapes can be very precise.
Vector formats are efficient when comparing information whose geographical shapes and sizes are different.
Vector files are much smaller because a relatively small number of vectors can precisely describe large areas and a many attributes can be ascribed to these areas.

Georeferencing / projecting



Sources for images: principles.ou.edu

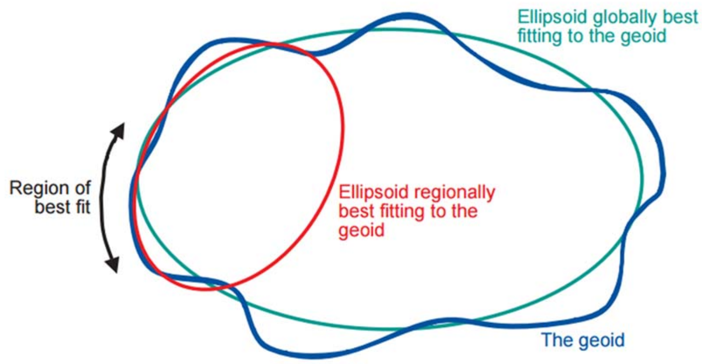
20-22/1/16 CEM Wind & Solar Training, Lilongwe, Malawi 11

Ellipsoids have varying position and orientations. An ellipsoid is positioned and oriented with respect to the local mean sea level by adopting a latitude (ϕ) and longitude (λ) and ellipsoidal height (h) of a so-called fundamental point and an azimuth to an additional point. We say that this defines a local horizontal datum. Notice that the term horizontal datum and geodetic datum are being treated as equivalent and interchangeable words. Several hundred local horizontal datums exist in the world.

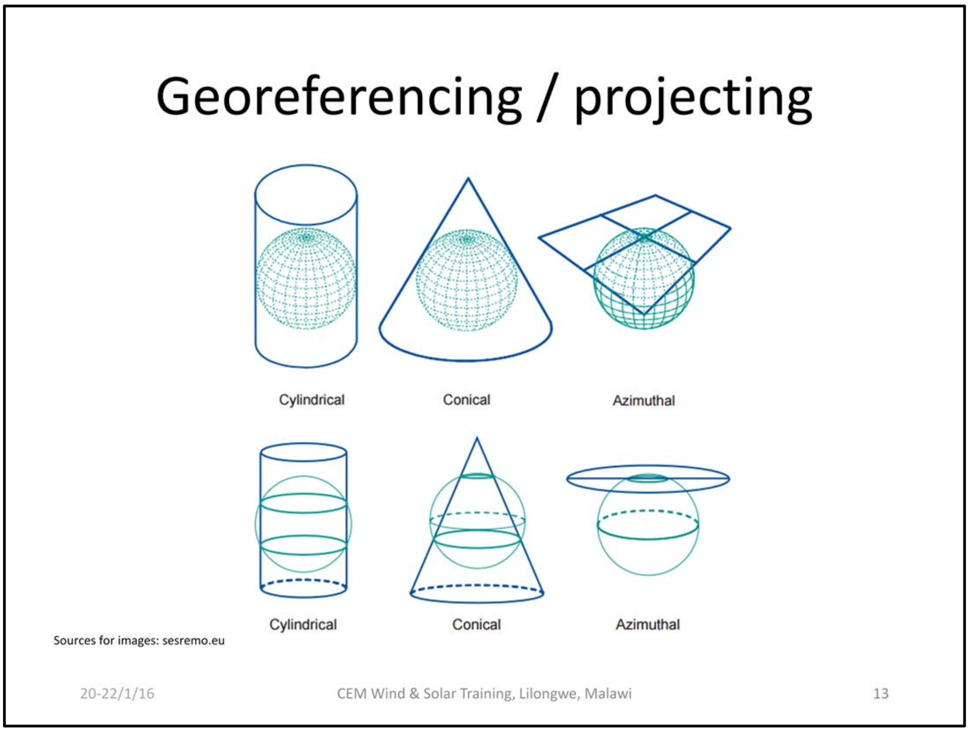
A map projection is a mathematically described technique of how to represent the Earth's curved surface on a flat map.

There is simply no way to flatten out a piece of ellipsoidal or spherical surface Scale distortions without stretching some parts of the surface more than others. The amount and which kind of distortions a map will have depends on the type of the map projection that has been selected.

Georeferencing / projecting

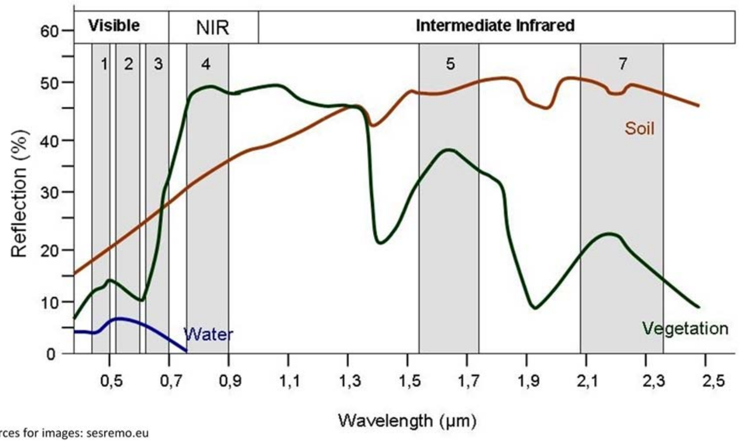


Sources for images: sesremo.eu

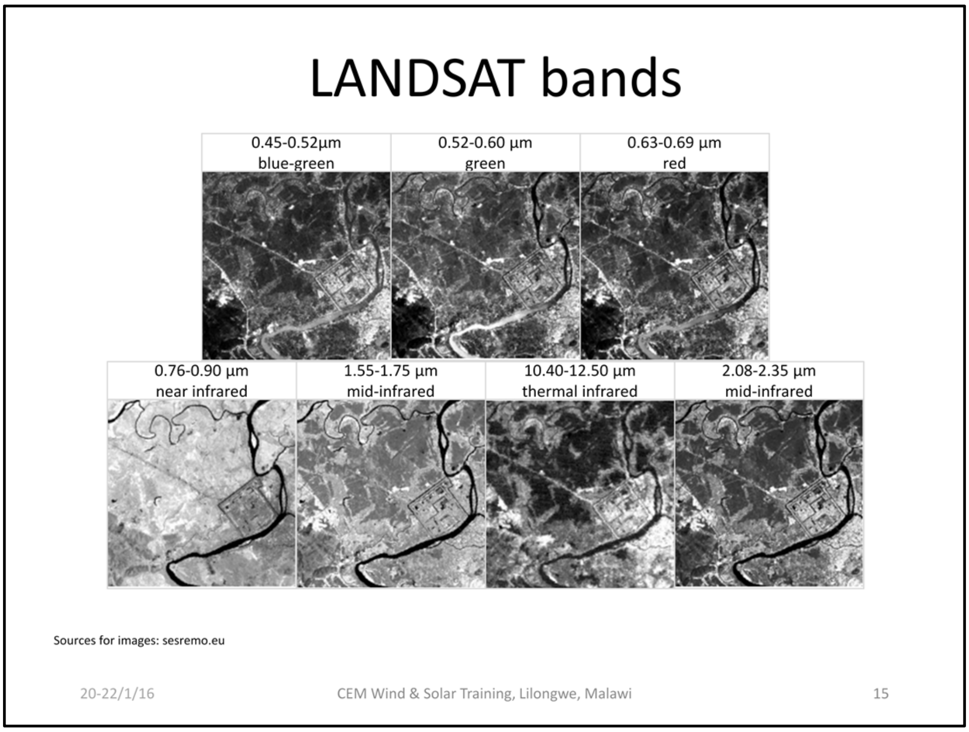


• In a conformal map projection the angles between lines in the map are identical to the angles between the original lines on the curved reference surface. This means that angles (with short sides) and shapes (of small areas) are shown correctly on the map.
• In an equal–area (equivalent) map projection the areas in the map are - identical to the areas on the curved reference surface (taking into account Distortion properties the map scale), which means that areas are represented correctly on the map.
• In an equidistant map projection the length of particular lines in the map are the same as the length of the original lines on the curved reference surface (taking into account the map scale).

Remote sensing

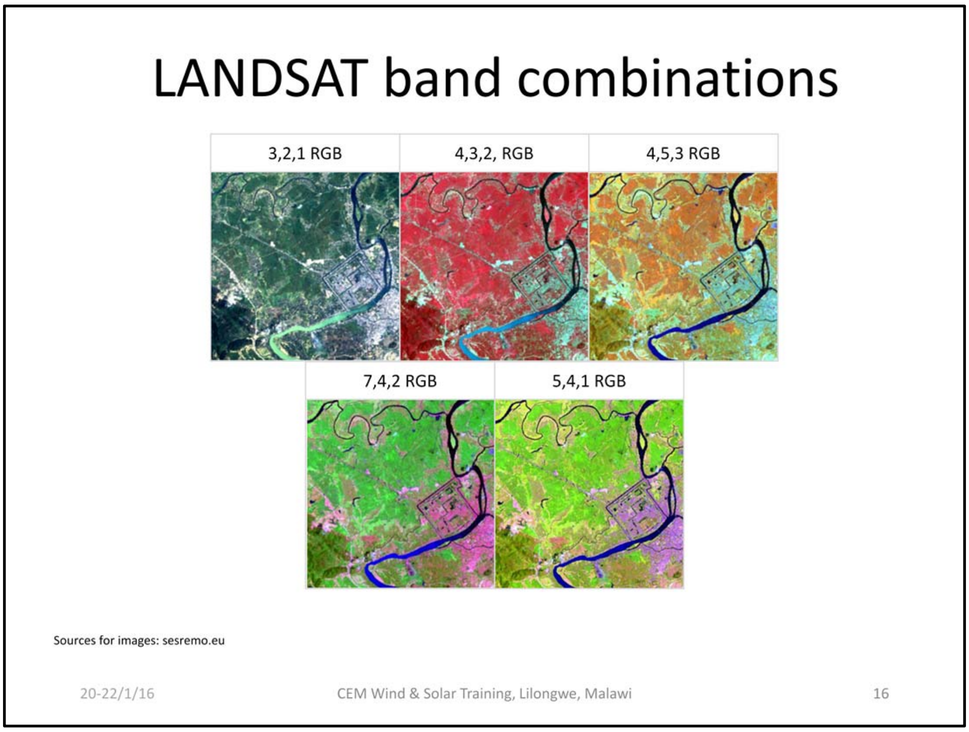


Sources for images: sesremo.eu

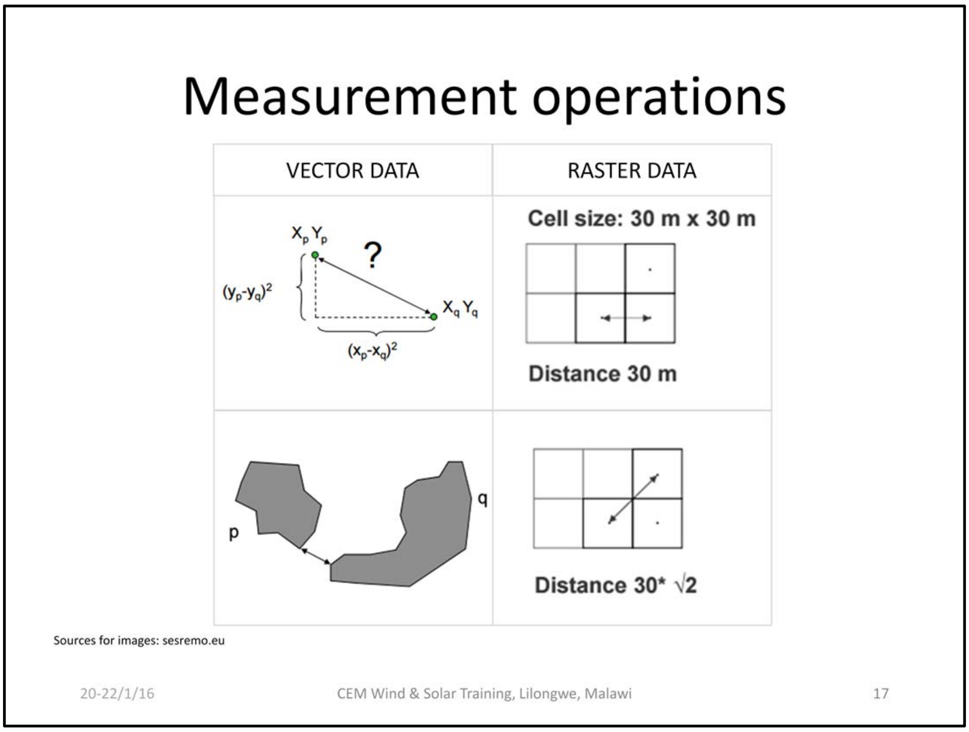


1) This short wavelength of light penetrates better than the other bands, and it is often the band of choice for monitoring aquatic ecosystems (mapping sediment in water, coral reef habitats, etc.). Unfortunately this is the “noisiest” of the Landsat bands since it is most susceptible to atmospheric scatter.
2) This has similar qualities to band 1 but not as extreme. The band was selected because it matches the wavelength for the green we see when looking at vegetation.
3) Since vegetation absorbs nearly all red light (it is sometimes called the chlorophyll absorption band) this band can be useful for distinguishing between vegetation and soil and in monitoring vegetation health.
4) Since water absorbs nearly all light at this wavelength water bodies appear very dark. This contrasts with bright reflectance for soil and vegetation so it is a good band for defining the water/land interface.
5) This band is very sensitive to moisture and is therefore used to monitor vegetation and soil moisture. It is also good at differentiating between clouds and snow.
6) This is a thermal band, which means it can be used to measure surface temperature. Band 6 is primarily used for geological applications but it is sometime used to measure plant heat stress. This is also used to differentiate clouds from bright soils as clouds tend to be very cold. The resolution of band 6 (60m) is half of the other bands.
7) This band is also used for vegetation moisture although generally band 5 is

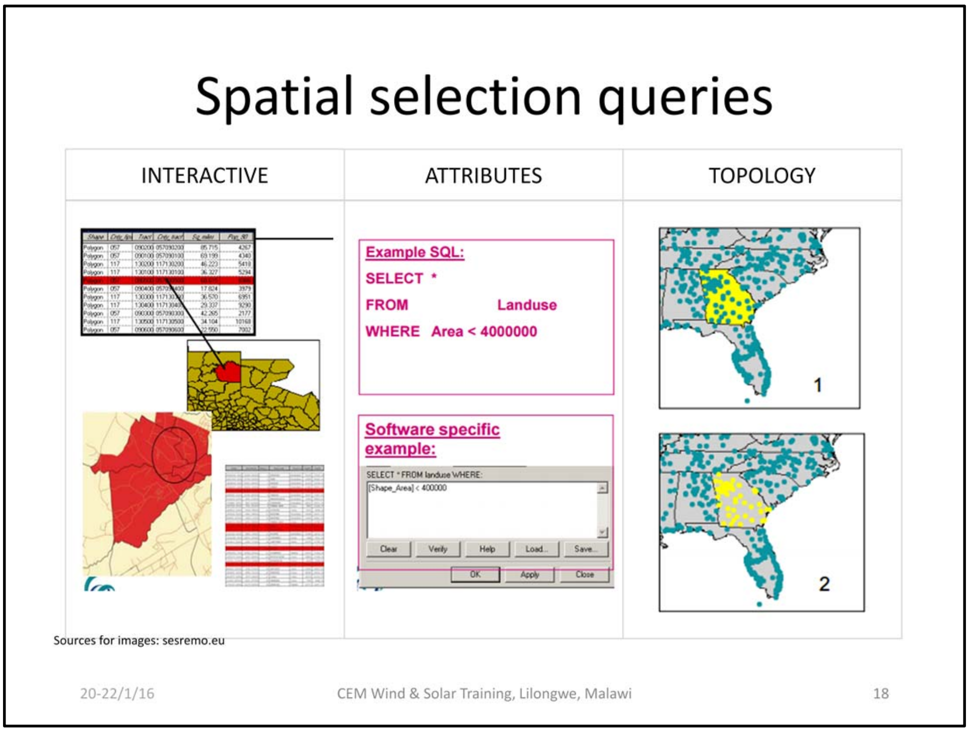
preferred for that application, as well as for soil and geology mapping.



1) This color composite is as close to true color that we can get with a Landsat ETM image. It is also useful for studying aquatic habitats. The downside of this set of bands is that they tend to produce a hazy image.
2) This has similar qualities to the image with bands 3,2,1 however, since this includes the near infrared channel (band 4) land water boundaries are clearer and different types of vegetation are more apparent. This was a popular band combination for Landsat MSS data since that did not have a mid-infrared band.
3) This is crisper than the previous two images because the two shortest wavelength bands (bands 1 and 2) are not included. Different vegetation types can be more clearly defined and the land/water interface is very clear. Variations in moisture content are evident with this set of bands. This is probably the most common band combination for Landsat imagery.
4) This has similar properties to the 4,5,3 band combination with the biggest difference being that vegetation is green. This is the band combination that was selected for the global Landsat mosaic created for NASA.
5) This band combination has similar properties to the 7,4,2 combination, however it is better suited in visualizing agricultural vegetation.



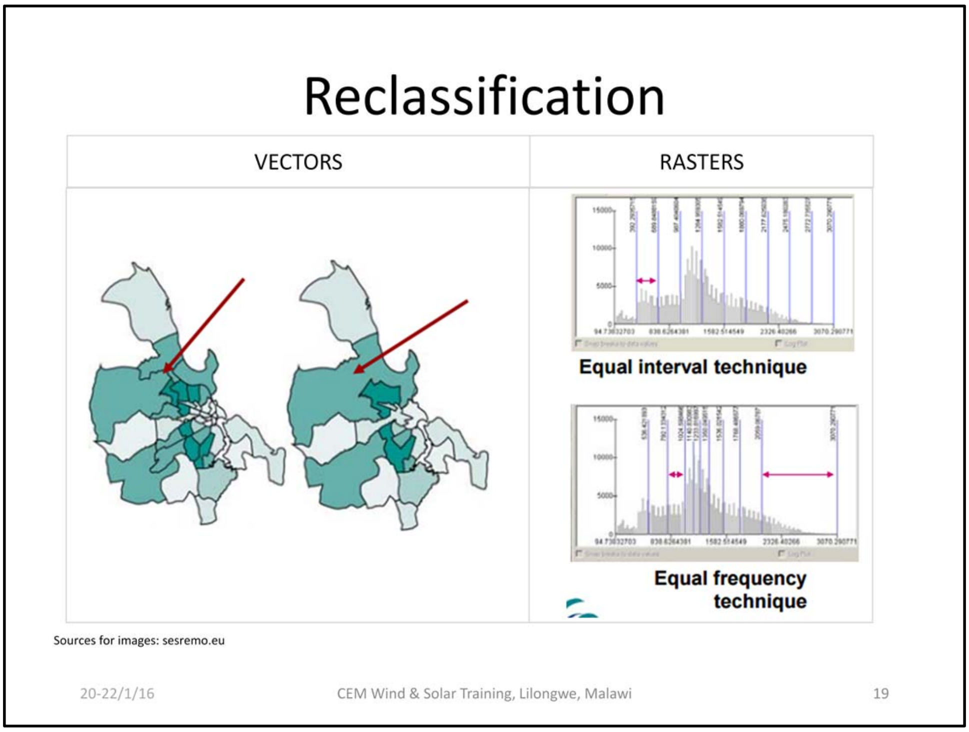
 Vector measurements include: location (coordinates and centroids), length, distance and area size Another geometric measurement is the minimal bounding box of lines and polygons. Pythagorean distance function is used Raster: Raster measurements include: location, distance and area size standard Pythagorean distance function applied to the locations of their mid-points.



Interactive
Spatial Selection by Attribute conditions
Relational operators
Logical operators
Combining attribute conditions

Spatial selection using topological relationships
Selecting features that are inside selection objects
Selecting features that intersect
Selecting features adjacent to selection objects
Selecting features based on their distance

Interactive spatial selection is a selection in which you select features by clicking on the screen (on the feature to select) or drawing a graphic, to select all objects within this graphic



 Remove detail from an input dataset to reveal important spatial patterns. Reduce the number of classes and eliminate details. If the input dataset itself is the result of a classification we call it a reclassification Spatial selection is a two step process: Select one or more selection objects. Apply a chosen spatial relationship to determine the features that have that relationship with the selection objects.

Neighborhood functions

	Buffer	Thiessen polygons	Spread computation	Seek computation
Method	Geometric distance	Geometric distance	Cost distance	Least cost path
Data structure	Vector and raster	Vector	Raster	Raster
Input	Target points	Target points	Target points and resistance raster	Elevation raster Direction raster
Output	Buffer polygon or raster	Polygon layer	Minimal total resistance	Accumulated flow count

Sources for images: sesremo.eu

Principle is simple, we select one or more target locations and determine the area around them. Buffer generation can be performed on vector as well as raster data. Target locations can be point, lines or polygons in a vector environment

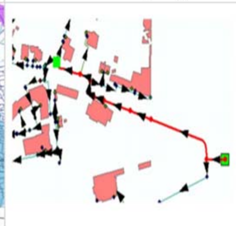
Thiessen polygons Divide an area into polygons, so that each polygon contains locations that are closer to the midpoint than to any other midpoint It will generate a polygon around each target location that identifies all those locations that ‘ belong to’ that target

In spread computation the neighbourhood of a target location not only depends on distance but also on direction and differences in the terrain Examples are pollutions and radio waves

 Seek computation applies when a phenomenon does not spread in all directions, but chooses a least-cost path. A typical example is a drainage pattern in a catchment.

Network analysis

	Optimal path finding	Network allocation	Trace analysis
Directed/undirected networks	both	both	directed
Input	At least two points (origin – destination)	Min. one point, source of the service area	Trace origin
Output	Path	A set of street segments, or polygon covering these segments	Path
Other requirements	One or two cost fields	One or two cost fields, maximum distance	Condition (maximum distance, direction or capacity)

Sources for images: sesremo.eu

20-22/1/16 CEM Wind & Solar Training, Lilongwe, Malawi 21

Network is a set of connecting lines Network can represent rivers, roads, pipelines, telecommunication lines etc. Network analysis analyze the way ‘goods’ can be transported along these lines Network analysis can be done in raster or in vector.

Networks can be directed , transportation is only in one direction, for example rivers, or it can be undirected, the goods can be transported in both directions (roads)

Non-planar networks have multi-level crossings, underpasses and overpasses. When they are modeled in 2-D these overpasses and underpasses should be modeled in a special way Example of non-planar networks are roads.

Optimal path finding is used when a least cost path between two nodes in a network must be found. You need a cost function! Also called Impedance. One of the attributes in the feature attribute table. Length, travel time, etc. The least-cost path is the one that has the min. value of the total cost between two nodes

In network partitioning, the purpose is to assign lines and nodes (parts of the network) to a number of target locations (for example which part of the network belongs to a hospital, school or fire station):
We have a number of resource centers, and the problem is which part of the network

can be assigned to which service center. ⍰ In principle this is a simple problem, each line segment is assigned to the service center that is the nearest.

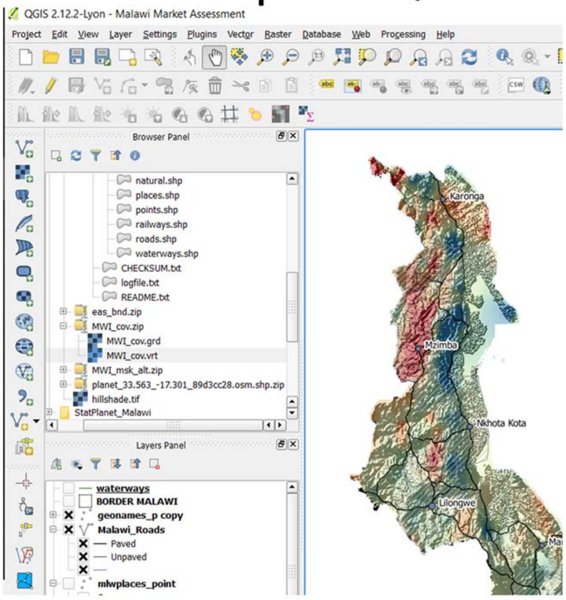
Trace analysis is performed when we want to understand which part of a network is connected to the trace origin. ⍰ A condition can be applied for example, trace only in the direction of the origin (upstream)

Software

	ArcGIS	QGIS
Functionality	Advanced	Intermediate to advanced
Ease of use	Very good	good
Adaptability	Good	Very good
Support	Very good	Intermediate to good
Price	Costly	Free

- Web-based map services
  http://irena.masdar.ac.ae/

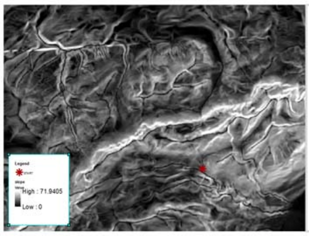
First steps in QGIS



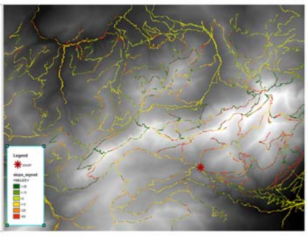
TIME-COST SURFACE | SITE-SELECTION | MALAWI EXAMPLE

TRAVEL DIRECTION

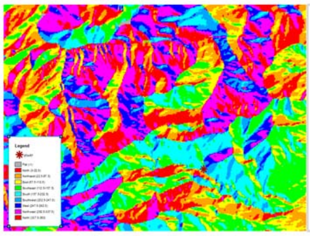
ELEVATION MODEL



+/- SLOPE

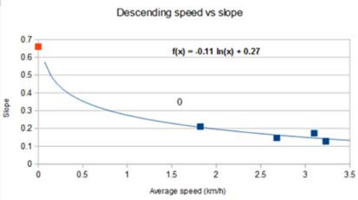


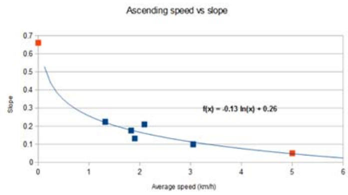
ASPECT



TIME-COST SURFACE | SITE-SELECTION | MALAWI EXAMPLE

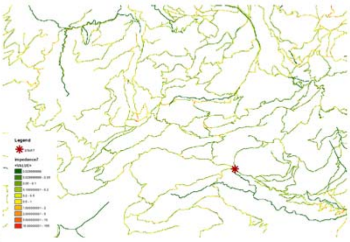
DESCENDING SPEED



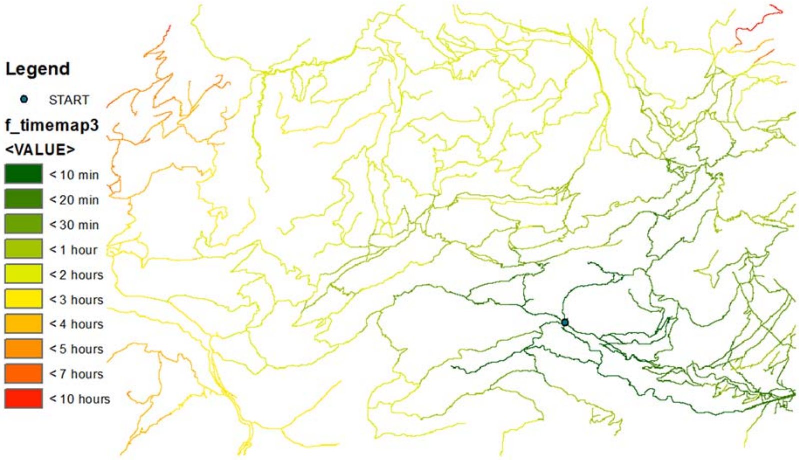


ASCENDING SPEED

IMPEDANCE

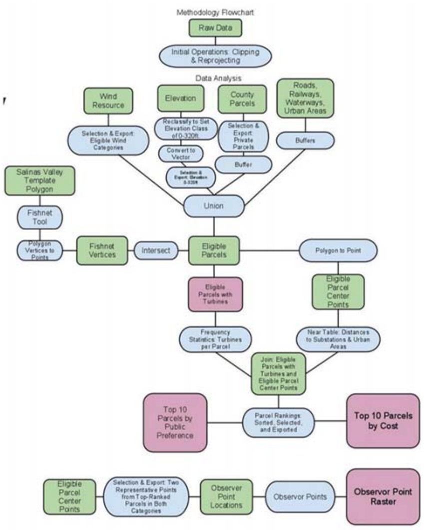


TIME-COST SURFACE | SITE-SELECTION | MALAWI EXAMPLE



TIME-COST SURFACE	SITE-SELECTION	MALAWI EXAMPLE

Feature Layer	Importance of Data Source
Parcel Boundaries & Ownership	This vector layer was used to extract privately owned parcels to analyze for eligibility.
Wind Resource	Wind Resource Data from the National Renewable Energy Laboratory delineates areas with a sufficient wind resource for analysis.
Urban Areas & Populated Places	Because wind development is undesirable near populated areas, these layers were buffered. Their locations were also used to determine distances from eligible parcels for parcel rankings.
Rivers & Bodies of Water	Rivers and bodies of water were buffered to ensure exclusion of riparian habitat.
Roads & Railroads	Per the regulations set forth by Monterey County, these layers were buffered.
Electrical Substations	Distance to electrical substations was an important component of our ranking methodology.
Digital Elevation Model	A 10 meter DEM was used to exclude areas of condor habitat, above 320ft. Because the study area is small, a high resolution DEM was important.
Anemometer Location Points	This layer identifies the placement of anemometers currently collecting windspeed data in the Salinas Valley.

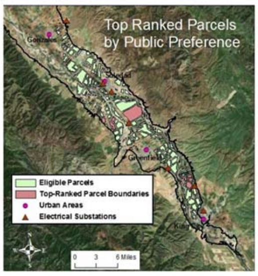


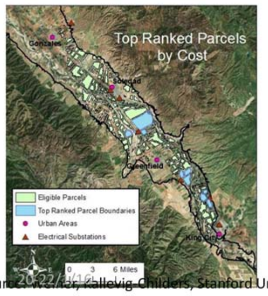
Source: Weiner, Kallevig-Childers, Stanford University

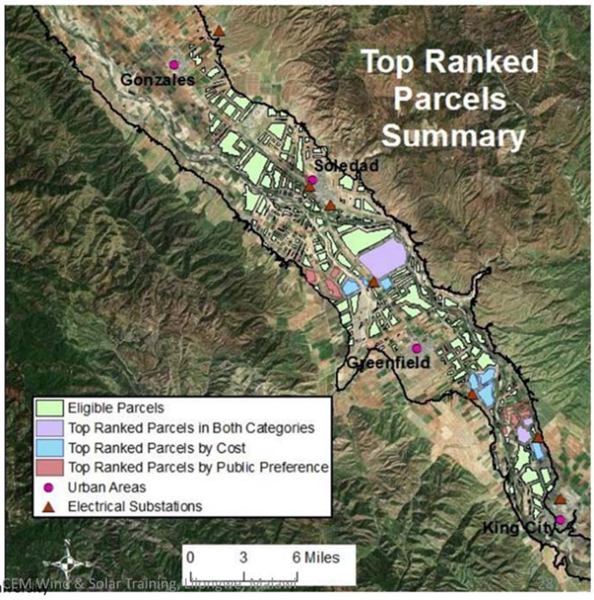
2022/1/16 CEM Wind & Solar Training, Lilongwe, Malawi 27

Amy and Kendra conducted a site-suitability analysis for wind energy development in Salinas Valley in Monterey County, CA. The purpose of their project was to identify land parcels eligible for wind turbines, specify the location and number of turbines per parcel, rank the parcels for optimal suitability, and perform a visual analysis of the viewsheds of the proposed turbines. They used wind resource data from the National Renewable Energy Lab, U.S. Census data for urban areas and populated places, the 10m resolution NED DEM from the USGS, and data from anemometers in the Salinas Valley georeferenced using Google Earth Pro. By employing the fishnet, buffer, union, and observer point analysis tools, Amy and Kendra were able to create a summarizing site suitability surface. They conclude that there is significant wind energy potential in Salinas Valley, but that prior to turbine construction, there is need for additional electrical substations, and further research to determine which locations would be optimal for new substations.

TIME-COST SURFACE | **SITE-SELECTION** | MALAWI EXAMPLE

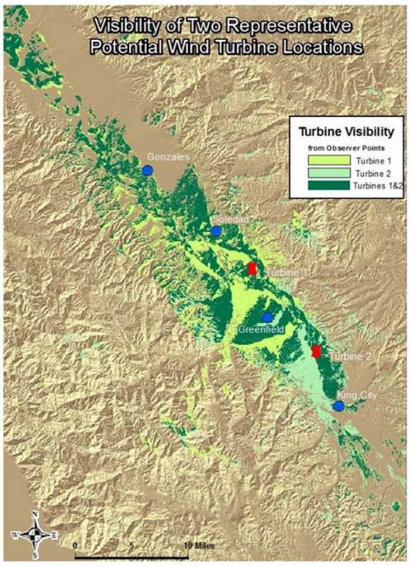






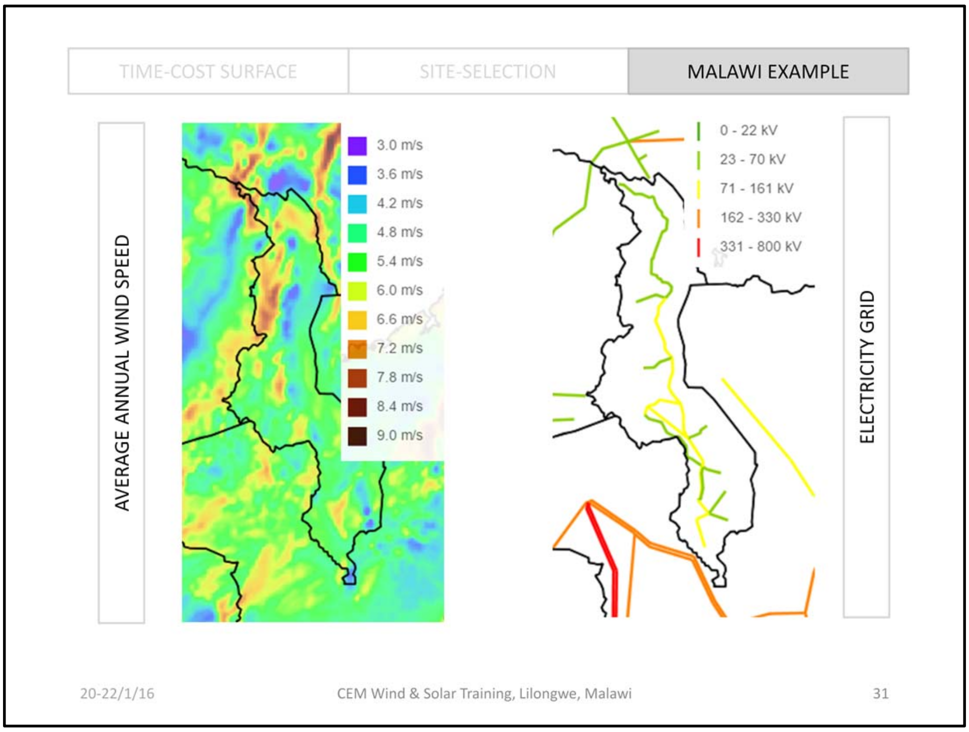
Source: W[illegible], Kallevig-Childers, Stanford University

TIME-COST SURFACE | SITE-SELECTION | MALAWI EXAMPLE



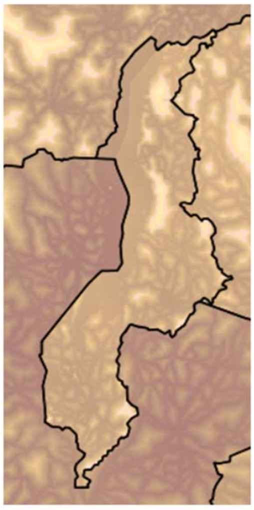
Source: Weiner, Kallevig-Childers, Stanford University

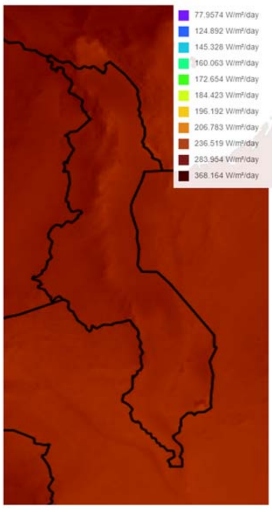
- Where in Malawi for off-grid Wind, PV, Diesel or combinations?
- What is the market size?
  - Population now, in 2030?
  - Household income?


TIME-COST SURFACE
SITE-SELECTION
MALAWI EXAMPLE
AVERAGE ANNUAL WIND SPEED
3.0 m/s
3.6 m/s
4.2 m/s
4.8 m/s
5.4 m/s
6.0 m/s
6.6 m/s
7.2 m/s
7.8 m/s
8.4 m/s
9.0 m/s
0 - 22 kV
23 - 70 kV
71 - 161 kV
162 - 330 kV
331 - 800 kV
ELECTRICITY GRID
20-22/1/16
CEM Wind & Solar Training, Lilongwe, Malawi
31

TIME-COST SURFACE | SITE-SELECTION | **MALAWI EXAMPLE**

DIESEL SYSTEM COST





SOLAR IRRADIATION


THANK YOU!