Terrain Analysis

SOMAIYA
VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya
TRUST

What is a terrain?

- Terrain is the term for an area of land.
- Terrain may include, flat plains, mountains, forests.
- Terrain is a term used to describe the layout of the landscape, and it can also be referred to as topography.
- The terrain plays an important role in many spatial models. In a GIS, the information about altitude is stored in digital terrain models.

Terrain Analysis

- Slope (Landslide susceptibility)
- Aspect (Solar insolation, vegetation)
- Catchment or dispersal area (Runoff volume, soil drainage)
- Flow path (Distance of water flow to point)
- Profiles, fence diagrams
- Viewshed (visibility)

What is a digital elevation model (DEM)?

- A Digital Elevation Model (DEM) is a representation of the bare ground (bare earth) topographic surface of the Earth excluding trees, buildings, and any other surface objects.
- DEM – Digital Elevation Models
- DSM – Digital Surface Models
- DTM – Digital Terrain Models and even
- TIN – Triangular Irregular Networks

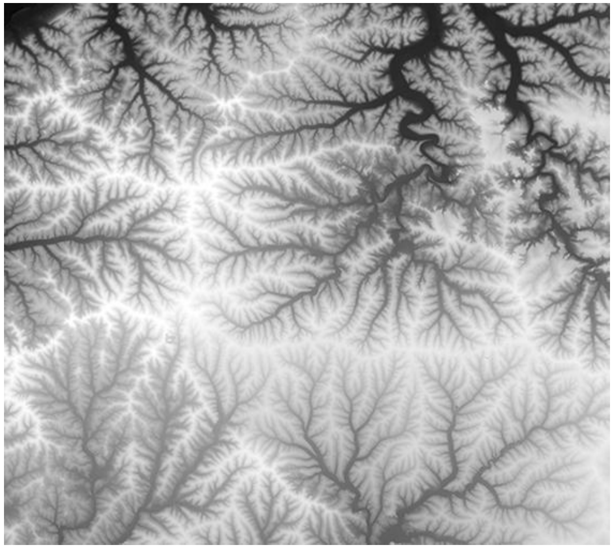
Definition

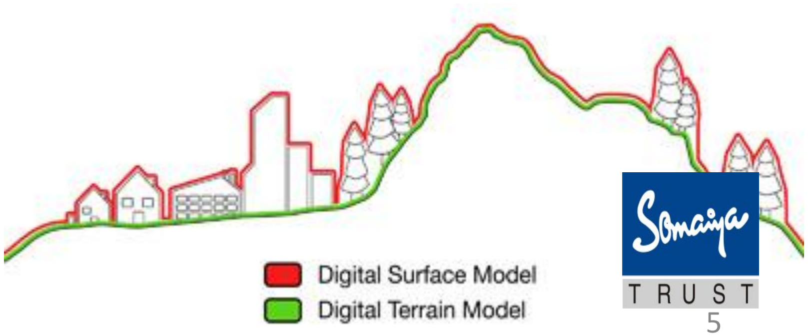
In scientific literature there is no universal agreement about the usage of the terms:

- *digital elevation model* (DEM)
- *digital terrain model* (DTM)
- *digital surface model* (DSM)

DEM is often used as synonimous of DTM, but in many cases it used as a generic term for both DTMs and DSMs.

The most common representation of a DEM is a raster, where the DNs correspond to the average elevation value of the area framed by the cell.





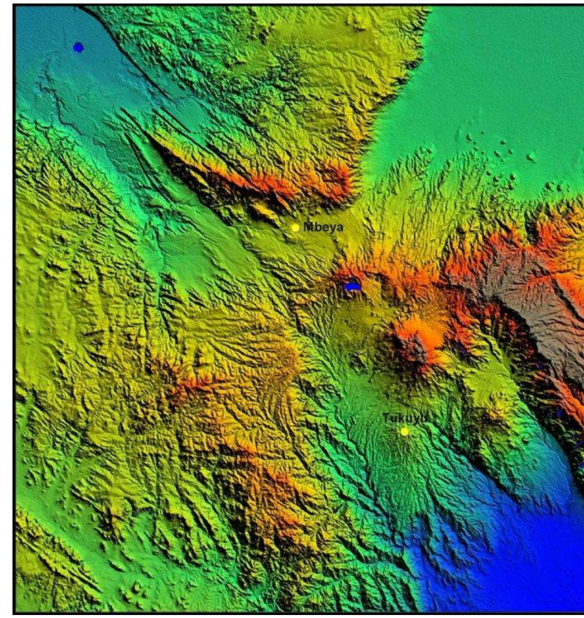
SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

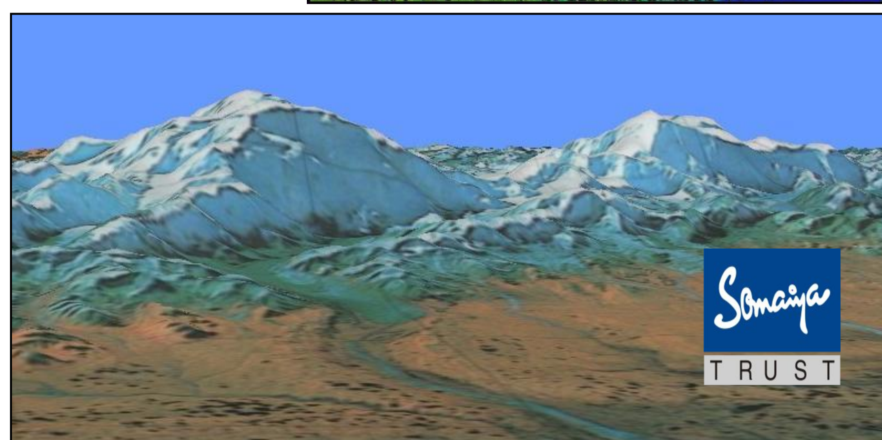
Somaiya TRUST

Uses

- Terrain analysis in geomorphology and physical geography
- Modelling water flow for hydrology
- Modelling soil erosion or mass movement
- Creation of relief maps
- Rendering of 3D visualizations
- Rectification of aerial photography or satellite imagery
- Climatology
- Urban planning
- Logistics and communications
- Mining

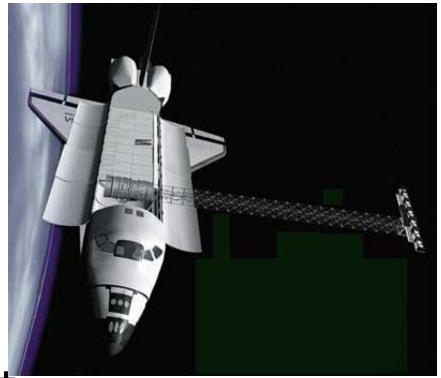
- ...and many more other GIS applications





SOMAIYA
VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Sources of elevation data



- Stereo photogrammetry from aerial surveys
  - traditional: manually by a trained photogrammetrist
    - contours of topographic maps, to be interpolated
  - modern: by automatic stereo-correlation
- Stereo-correlation from optical satellite imagery
  - SPOT (off-track stereoscopy)
  - Aster DEM (along-track stereoscopy)
- Interferometry from radar data
  - European Remote Sensing Satellite ERS (multi-pass)
  - Shuttle Radar Topography Mission SRTM (single-pass)
- Lidar
  - interpolation of a matrix of xyz points obtained using a laser beam

Spatial / Terrain Analysis

- The operations for spatial analysis of continuous fields that are represented by the regular square grid, where each attribute is represented by a separate overlay, and each grid cell is allowed to take a different, scalar value.
- The z attribute can represent any continuously varying attribute or regionalized variable, such as levels of pollutants in soil, atmospheric pressure, annual precipitation, an index of marketing potential, population density, or the costs of access to a given location.

Basic operations for spatial analysis with discretized continuous fields

- Map algebra and cartographic modelling

- Major advantage with raster representation in which each attribute is recorded in a separate overlay, is that any mathematical operation performed on one or more attributes for the same cell can easily be applied to all cells in the overlay.
- This means that one can use exactly the same algebraic notation to operate on gridded data as on single numbers. The method is called Map Algebra and the procedure of using algebraic techniques to build models for spatial analysis is called Cartographic Modeling

Spatial operations

- Using gridded data has advantages and disadvantages compared with a topologically linked vector database of defined entities.
- The disadvantages include the problems that the exact shapes of entities are only approximated by the grid cells and that directed operations over a network cannot be carried out without first deriving the topology from the properties of the surface.
- The advantages are that the continuous field model provides a much richer suite of truly spatial analysis operations that have many practical uses.

These spatial operations include:

- Interpolation
- Spatial filtering
- First and higher-order derivatives
- The derivation of surface topology: drainage networks and catchment delineation
- Contiguity assessment (clumping)
- Non linear dilation (spreading with friction)
- Viewsheds, Shaded relief, and irradiance.

Interpolation

- Interpolation is the prediction of a value of an attribute z at an unsampled site (x0) from measurements made at other sites xt falling within a given neighbourhood. Interpolation is used to create discretized continuous surfaces from observations at sparsely located points or for resampling a grid to a different density or orientation as in remote sensing images.
- Interpolation can be seen as a particular class of spatial filtering where the input data are not necessarily already located on a continuous grid.

Spatial analysis using square windows

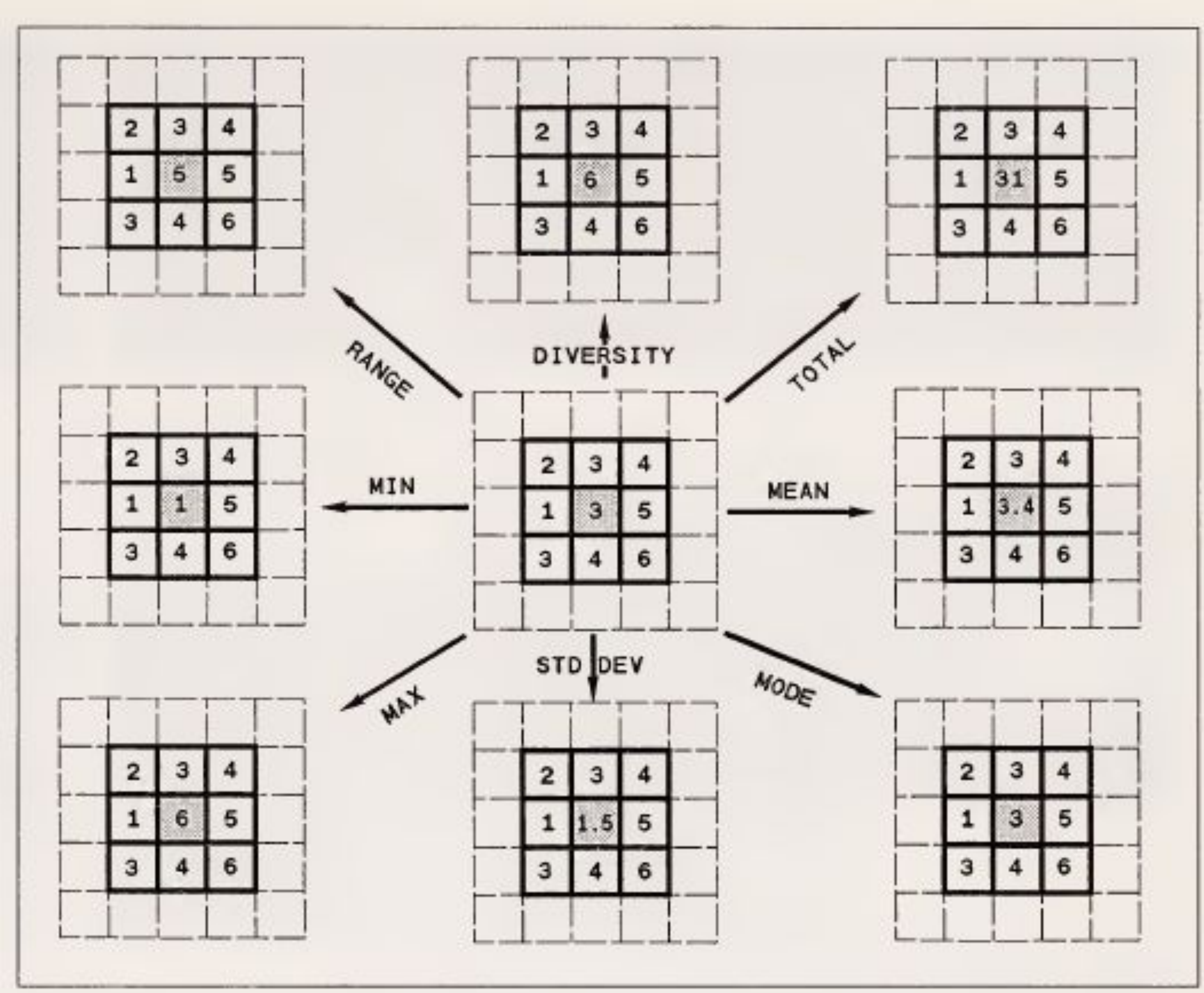
- Spatial filtering
- The simplest and perhaps most widely used method of spatial filtering a discretized, continuous surface involves passing a square window (otherwise known as a kernel or filter) over the surface and computing a new value of the central cell of the window
- $C_{i,j}$ as a function of the cell values covered by the window. This kind of operation is also commonly known as convolution.
- The window is frequently of size 3x3 cells, but any other kind of square window (5x5, 7x7 cells, or distance measurements) is possible.

$$C_{i,j} = f\left(\sum_{i-m}^{i+m}\sum_{j-n}^{j+n} c_{i,j} \cdot \lambda_{i,j}\right)$$

Gaussian Blurring

- **What is Gaussian blurring?**
- Named after mathematician Carl Friedrich Gauss (rhymes with “grouse”), Gaussian (“gow-see-an”) blur is the application of a mathematical function to an image in order to blur it.
- “It’s like laying a translucent material like vellum on top of the image,” says photographer Kenton Waltz. “It softens everything out.”
- A type of **low-pass filter**, Gaussian blur smoothes uneven pixel values in an image by cutting out the extreme outliers.
- **When to use Gaussian blur.**
- Photographers and designers choose Gaussian functions for several purposes.
- If you take a photo in low light and the resulting image has a lot of noise, Gaussian blur can mute that noise.
- If you want to lay text over an image, a Gaussian blur can soften the image so the text stands out more clearly.

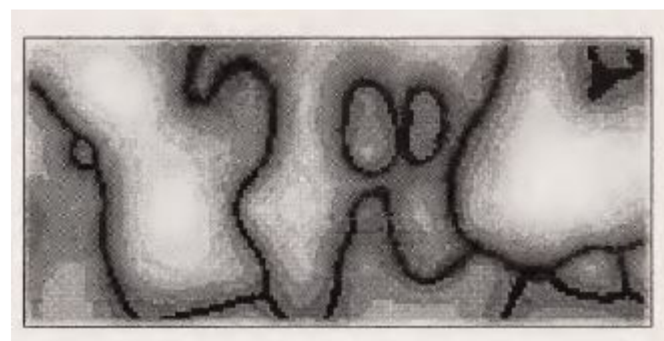
1	1/15	2/15	1/15
2	2/15	3/15	2/15
3	1/15	2/15	1/15
	1	2	3


RANGE
DIVERSITY
TOTAL
MIN
MEAN
MAX
STD DEV
MODE
3.4
1.5
31

High-pass and edge filters

- The high pass filter accentuates the comparative difference between a cell's values and its neighbors.
- It has the effect of highlighting boundaries between features (for example, where a water body meets the forest), thus sharpening edges between objects.
- It is generally referred to as an edge-enhancement filter.
- A 3 x 3 filter for this option is:

```
-0.7  -1.0  -0.7
 -1.0   6.8  -1.0
 -0.7  -1.0  -0.7
```



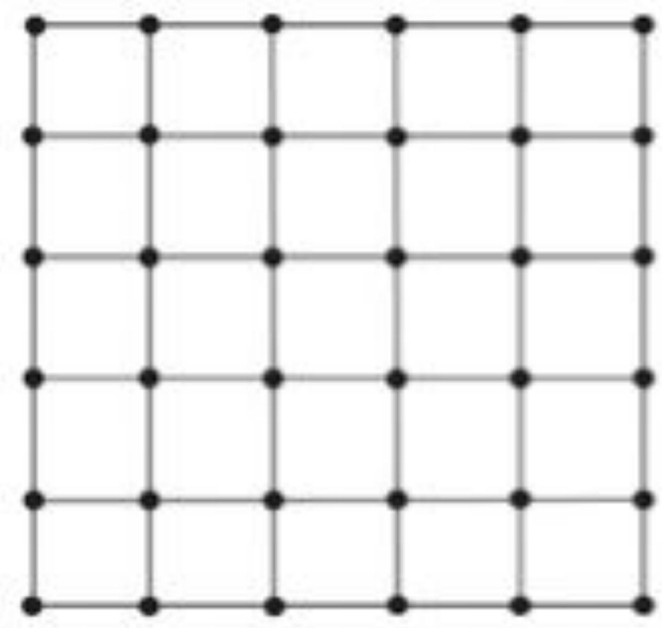
Laplacian filter

0	1	0
1	−4	1
0	1	0

Grid vs. TIN

- In a digital terrain model, altitude values are usually organized in one of these data structures:
  - Grid
    - The structures of grids are similar to raster data sets.
    - Each rectangular grid cell has the same size.
    - The altitude is stored in each of these grid cells.
    - This data structure is the most widely used structure, due to its simplicity and easily implemented algorithm.
    - However, the disadvantage of a grid is that density cannot be adjusted to the complexity of the terrain. For that reason there are often too many data used for the representation of simple terrains.
  - TIN
    - TINs consist of irregularly distributed points.
    - The triangles are built up on these points.
    - This data structure allows for the efficient storage of terrain information; the triangulation allows a variable density and distribution of points.
    - It can be adapted to more complex terrain (more data points at complex terrains and fewer points in flat areas of the terrain).
    - It is more complicated to implement these algorithms than to implement algorithms based on raster data.

GRID Vs TIN



Information which can be derived from digital terrain models

Derivatives	Output type	Description
Slope	Number	Slope at a point
Gradient	Number	Gradient between two points
Aspect	Number	Orientation of the slope
Curvature	Number	Curvature in a certain direction (for example plan and profile curvature)
Intervisibility	Yes/no	specifies whether the viewer can see a certain point
Viewshed	Polygon	Area which is visible from one or more points
Hillshade	Image	Shaded relief under a given illumination angle
Stream networks	Line	Lines of water runoff in the terrain
Catchment	Polygon	Area where other areas drain in
Profile	Line	Change in elevation of a surface along a line
Volume	Number	Calculation of volume change between two surfaces
Perspective image	Image	Perspective relief representation
Line of greatest slope	Line	Path along the steepest slope

21

- These parameters are used in various applications
- Aspect is used in vegetation geography
- Altitude, aspect and gradient are parameters which are used for the modeling of potential alpine permafrost (Frozen ground).
- Slope and aspect are required again for the modeling of potential avalanches

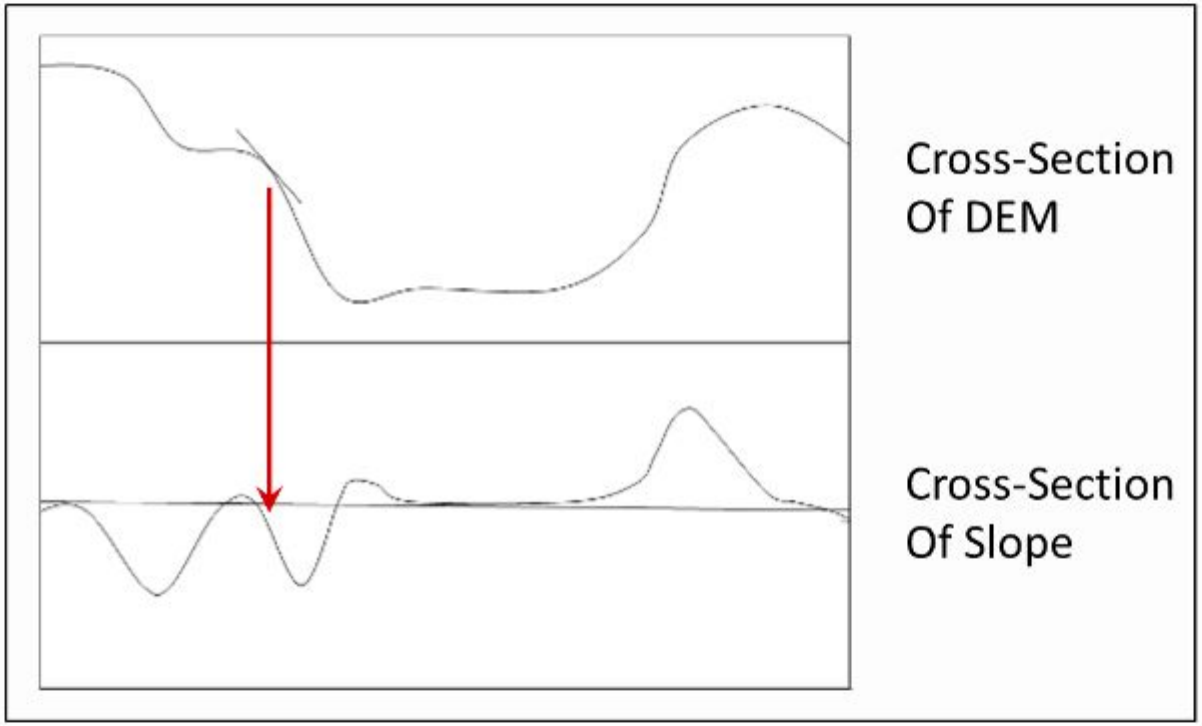
22

Slope

- The slope at a point is given by the tangent plane. In order to calculate the slope, the derivatives of the surface in the direction of the x- and y-coordinate have to be known.
- Therefore, special methods are needed to estimate the derivatives from elevation models. The most common method in grid based terrain models is called "finite difference"

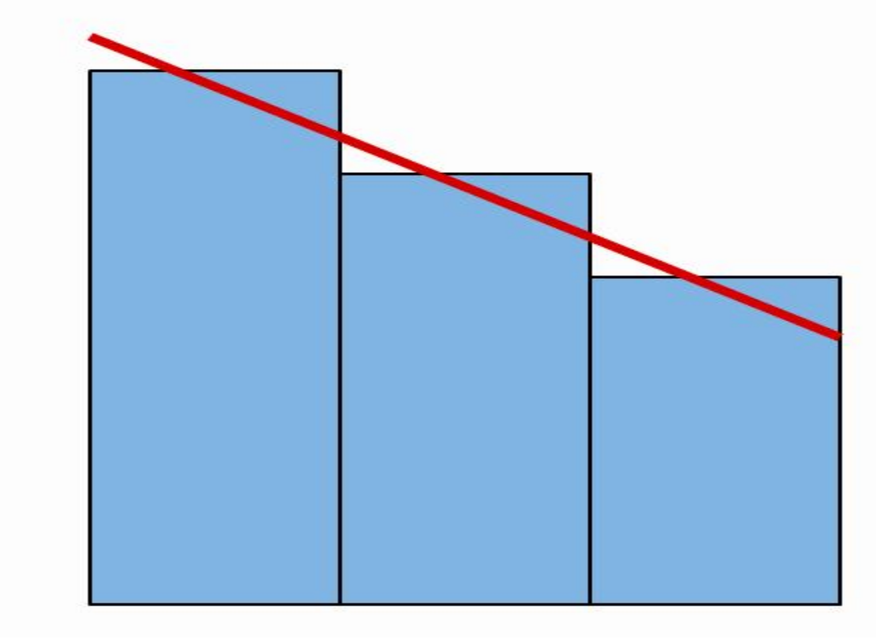
23

Slope From DEMs



If we move from left to right, the slope starts out close to zero because the elevation is very flat, as shown in the DEM. Then, the land drops in elevation and the slope goes negative. As the land flattens out, the slope comes back to zero and then starts to go negative again as the land slopes off

- If we look at the grid of pixels above in a cross-section, we can see that the slope is computed by putting a line through the pixels. The angle of this line is then the "slope" value that is put into the pixel in the center of the 3x3 grid.

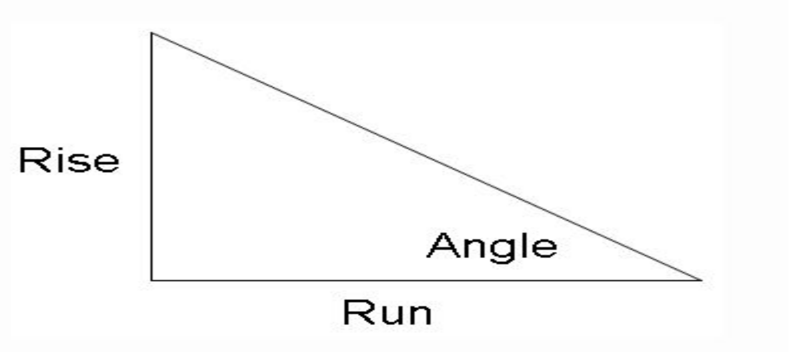


SOMAIYA
VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya
TRUST

25

- There are two common units for slope, percent and degrees. Slope in percent has a range from 0 to infinity (not 100%).



- Slope = (Rise/Run) * 100%
- Slope in degrees has a range from 0 to 90 and is computed using:
- Slope = arctan(Rise/Run).

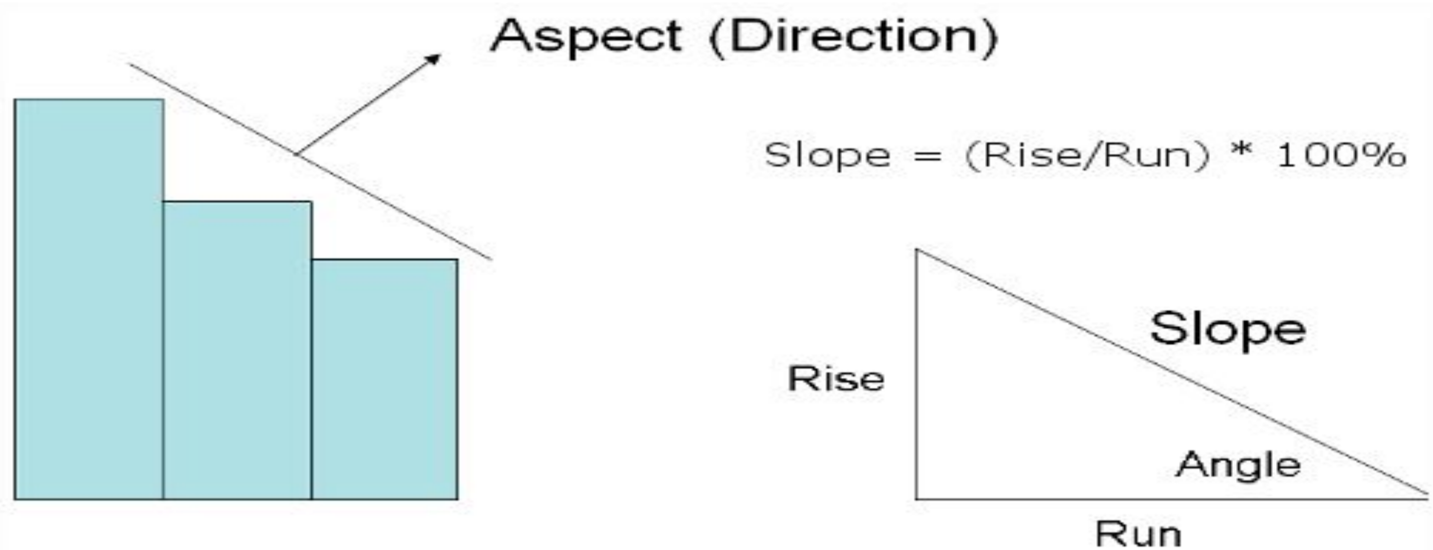
Somaiya Vidyavihar University
K J Somaiya College of Engineering

Somaiya TRUST

26

Aspect: Direction of the Slope

- Aspect is computed in a similar way to slope in that a plane is fit through a 3x3 grid of pixels. However it's the direction of the plane that is used for the pixel value instead of the slope. The direction at which slope faces.



27

- Aspect is the direction (clockwise from North = azimuth) to which the steepest slope of the plane tangent faces.
- To calculate the aspect, the derivative in direction x and in direction y have to be calculated

$$Aspect = arctan\left(\frac{b}{a}\right)$$

with

$$a = \left(\frac{dz}{dx}\right)$$

$$b = \left(\frac{dz}{dy}\right)$$

SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

Viewsheds, shaded relief, and irradiance

- Concerning the computation of the paths of light between a light source on or above the DEM and its effect at other locations.
- Line of sight maps
- Shaded relief maps
- Irradiance mapping

Line of sight maps

- The aim is to determine those parts of the landscape that can be seen from a given point.
- Intervisibility is often coded as a binary variable—0 invisible, 1 visible.
- The collective distribution of all the ‘true’ points is called the **viewshed**.
- Intervisibility maps can be prepared from altitude matrices and TIN’s using tracking procedures

Line of sight maps

- The site from which the viewshed needs to be calculated is identified on the DEM and rays are sent out from this point to all points in the model.
- Points (cells) that are found not to be hidden by other cells are coded accordingly to give a simple map
- Estimating the intervisibility of sites is an important GIS application in simulators for pilot training, the location of microwave transmission stations or the location of forest fire warning stations.

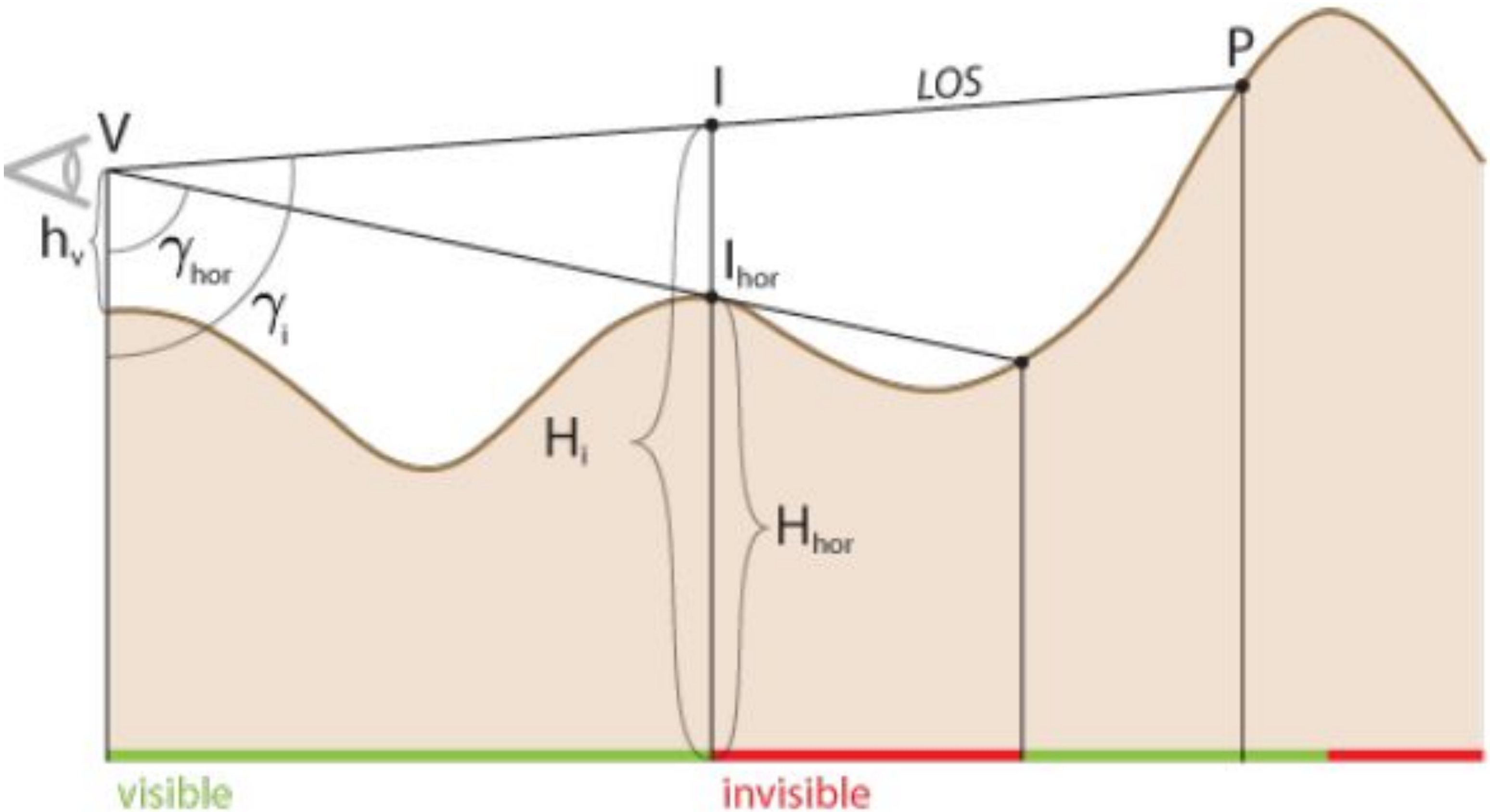
Visibility analysis

- Digital terrain models can be used for many engineering and planning applications. In this unit, visibility analysis is discussed in detail. A visibility analysis detects all points that are visible from a given point based on a digital terrain model.

32

Calculation of visibility

- In 1 Dimension, in order to determine if there is visibility between point *P* and point *V*, a line of sight is calculated between the two.
- If there is a higher point lying on the line of sight, the target point is not visible. Otherwise, the target point is visible.
- The easiest way to perform such a test is to calculate the slope of the line of sight (*LOS*) between point *V* and point *P* and to compare it to the slope of the line of sight between point *V* and point *I hor* , lying in between point *V* and point *P*.
- The vertical angles of the two lines can be compared.
- In the case that the angle of the *LOS* between *V* and *P* is bigger than the angle of the *LOS* between *V* and I *hor* , *P* is visible, Otherwise, *P* is not visible.

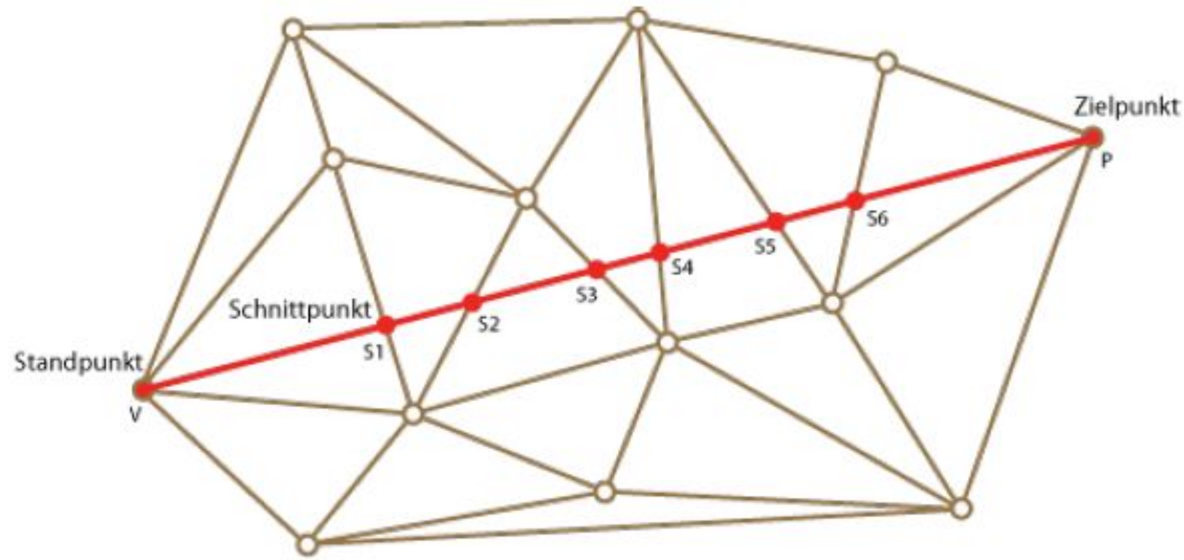


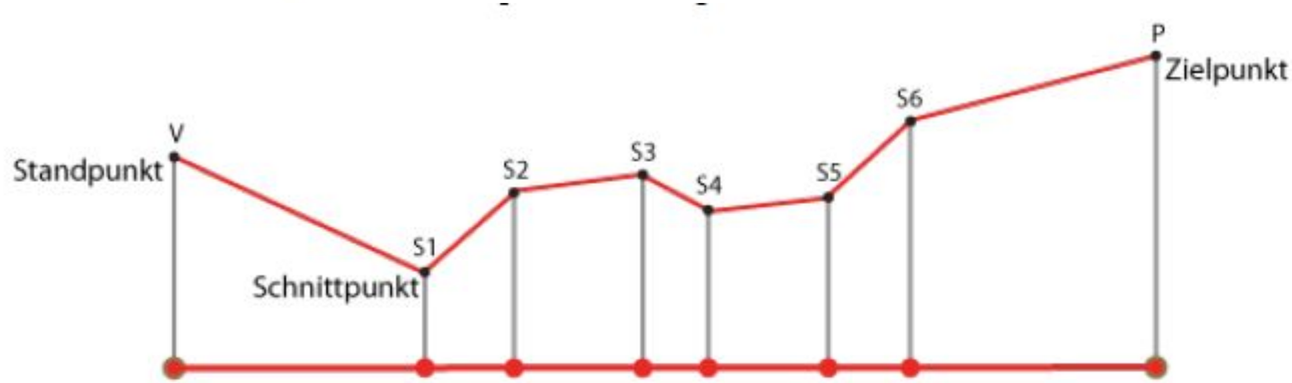
SOMAIYA
VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

34

In a terrain model based on TINs, the altitude of the intersection points between the connection line V-P and the triangles have to be calculated first. The altitude at the intersection points S1 to S *n* can be calculated from the two corresponding vertices of the triangle by performing a linear interpolation. At this point, the algorithm discussed above can be applied to calculate the LOS on the established profile.



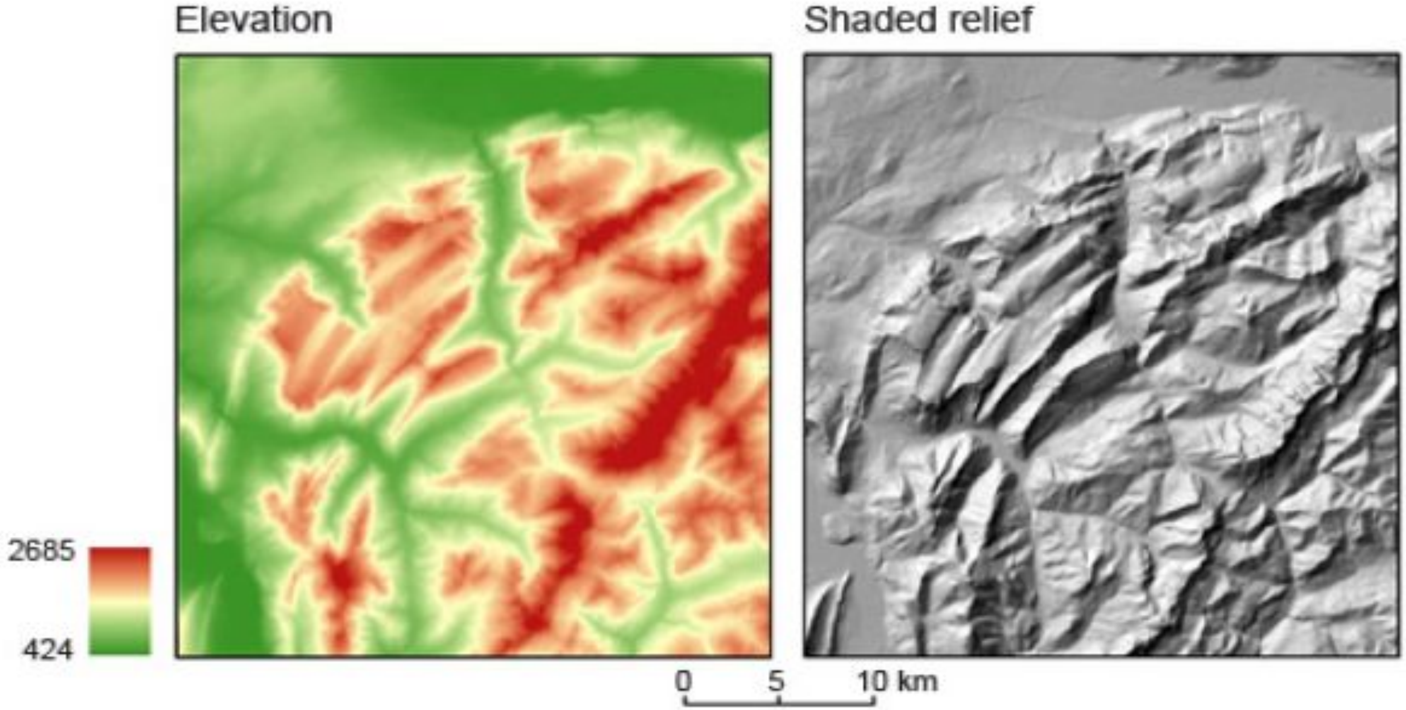


- Visibility analyses are applied in different fields.
- For example, it can be used to determine the location of an observation tower. By using a digital terrain model, different possible locations can be examined in a short period of time.
- The use of visibility analysis for the calculation of the expansion of electromagnetic radiation (e.g. for the planning of mobile phone antennas).
- A GIS can be used to find the optimal area for the relocation of animals. This reduces the possibility of being discovered by a predator.

36

Shaded relief maps

- The shaded relief map is computed from an altitude matrix.
- Shaded relief maps show features on the surface, such as mountains, valleys, plateaus, and canyons.
- Areas that are flat or have few features are smooth on the map, whereas areas with steep slopes and mountains appear more rough.
- The light source is usually chosen as being at an angle of 45° above the horizon in the north-west, a position that has much more to do with human faculties for perception than with astronomical reality.
- The computer simulates what the landscape would look like if the sun were low in the western sky (like several hours before sunset).
- West-facing slopes, which would be in the sun, are shown brighter than east-facing slopes, which would be shaded. The resulting light-dark pattern causes the 3D illusion.
- The terrain model is usually smoothed and generalized because of the data-gathering process and will not show the fine details present in the aerial photograph.


Elevation
Shaded relief
2685
424
0
5
10 km

Irradiance mapping

- This is the extension of the shaded relief principle to compute the amount of solar energy falling directly on a surface.
- The sun is now not fixed in any one position in sky, but is allowed to take a position according to the latitude, the time of day, and the day of the year.
- There is a need to incorporate the effect of atmospheric absorption on the amount of energy actually received, and also to model the shadowing effect of terrain.
- Integrate the daily or monthly estimates of irradiance for a whole season or year, and thereby to create a map that distinguishes sites in terms of the energy inputs for plant growth, home heating, or rock weathering.

Irradiance mapping

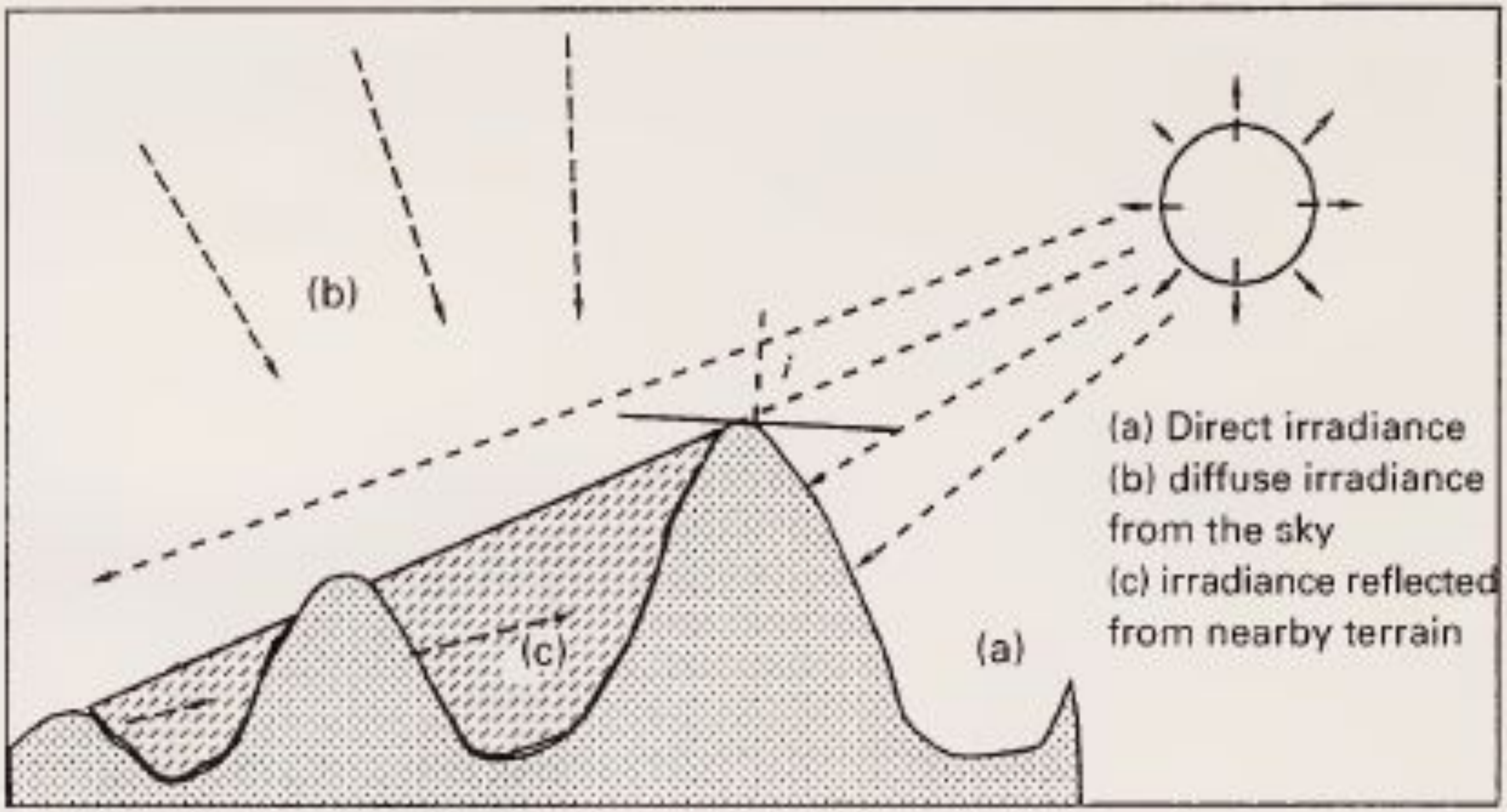


Figure 8.17. Computing solar irradiance for a slope

Hillshade

- Sometimes, the relief in topographic maps is shaded to get a better impression of the third dimension.
- For such kinds of relief maps (hillshade), the illumination source is defined generally at an angle of 45° from the north-west.
- Even though this position is very unrealistic for the northern hemisphere, it is known that this sun position gives the best impression of relief in the third dimension.
- The brightness of the shadow of a given surface element (either a grid cell or a TIN section) depends on the following properties:
  - Aspect and slope of the surface element
  - Reflecting properties of the surface element

Flow Direction

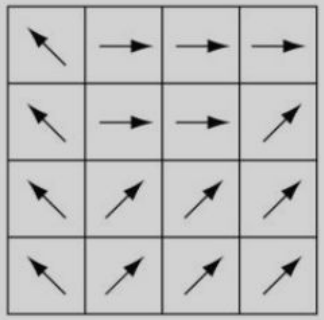
- Useful for finding drainage networks and drainage divides
- Direction is determined by the elevation of surrounding cells
  - Water can flow only into one cell
- Water is assumed to flow into one other cell, unless there is a sink
  - GIS model assumes no sinks

Flow direction in a DEM

(a)

1014	1011	1004	996
1019	1015	1007	999
1025	1021	1012	1003
1033	1029	1020	1003

(b)



(c)

0	0	1	2
0	0	2	6
0	0	2	3
0	1	2	3

Figure 15.9
This illustration shows a filled elevation raster (a), a flow direction raster (b), and a flow accumulation raster (c). Both shaded cells in (c) have the same flow accumulation value of 2. The top cell receives its flow from its left and lower-left cells. The bottom cell receives its flow from its lower-left cell, which already has a flow accumulation value of 1.

Flow accumulation

- The number of cells, or area, which contribute to runoff of a given cell
- The accumulation function determines the area of a watershed that contributes runoff to any given cell

Flow accumulation in a DEM

0	0	0	0	0
0	1	3	1	0
0	1	8	1	0
0	1	13	1	0
0	2	24	2	0

Flow accumulation for individual cells

Watershed

- An area that contributes flow to a point on the landscape

  Water falling anywhere in the upstream area of a watershed will pass through that point.

- Identified from a flow direction surface

Drainage network

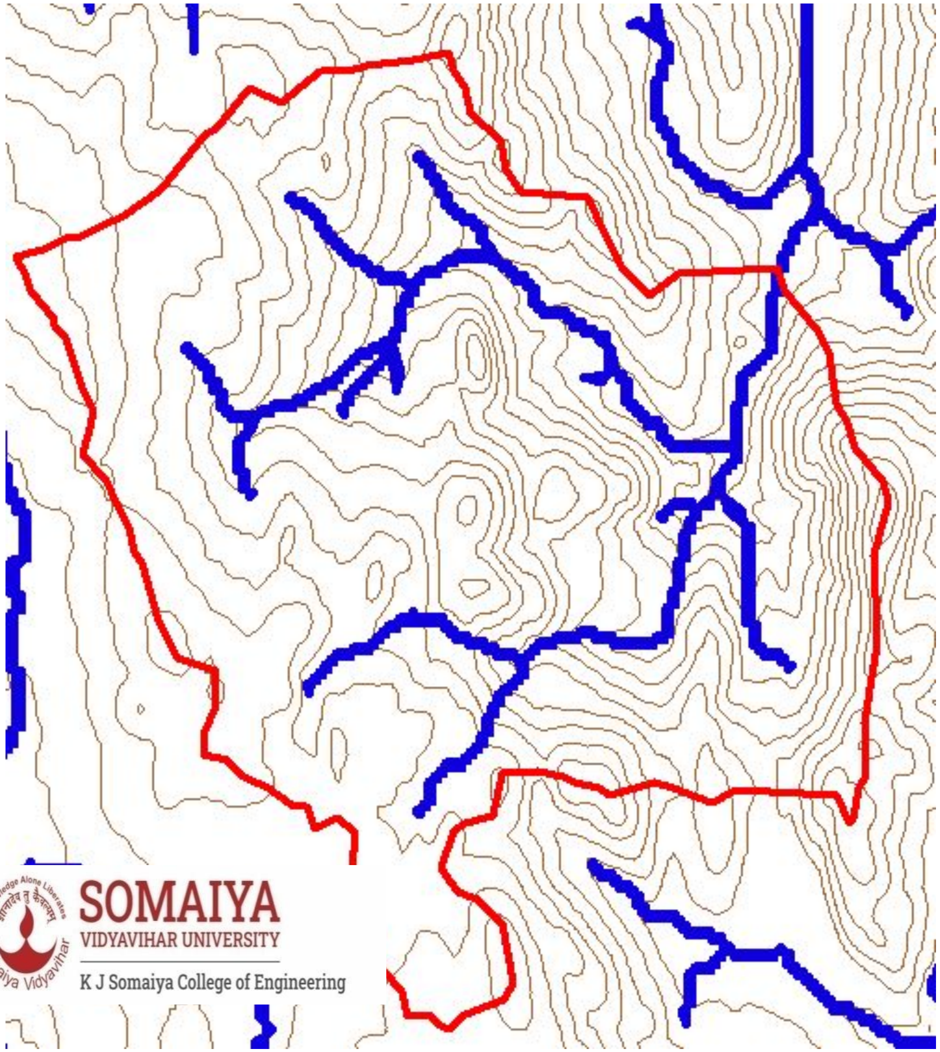
- A set of cells through which surface water flows

- Based on the flow direction surface

Finding watersheds …

- Begin at a source cell of a flow direction database, derived from a DEM
  - Find all cells that flow into the source cell
  - Find all cells that flow into those cells.
    - Repeat …
- All of these cells comprises the watershed
- The resulting watershed is generalized, based on the cell size of the DEM

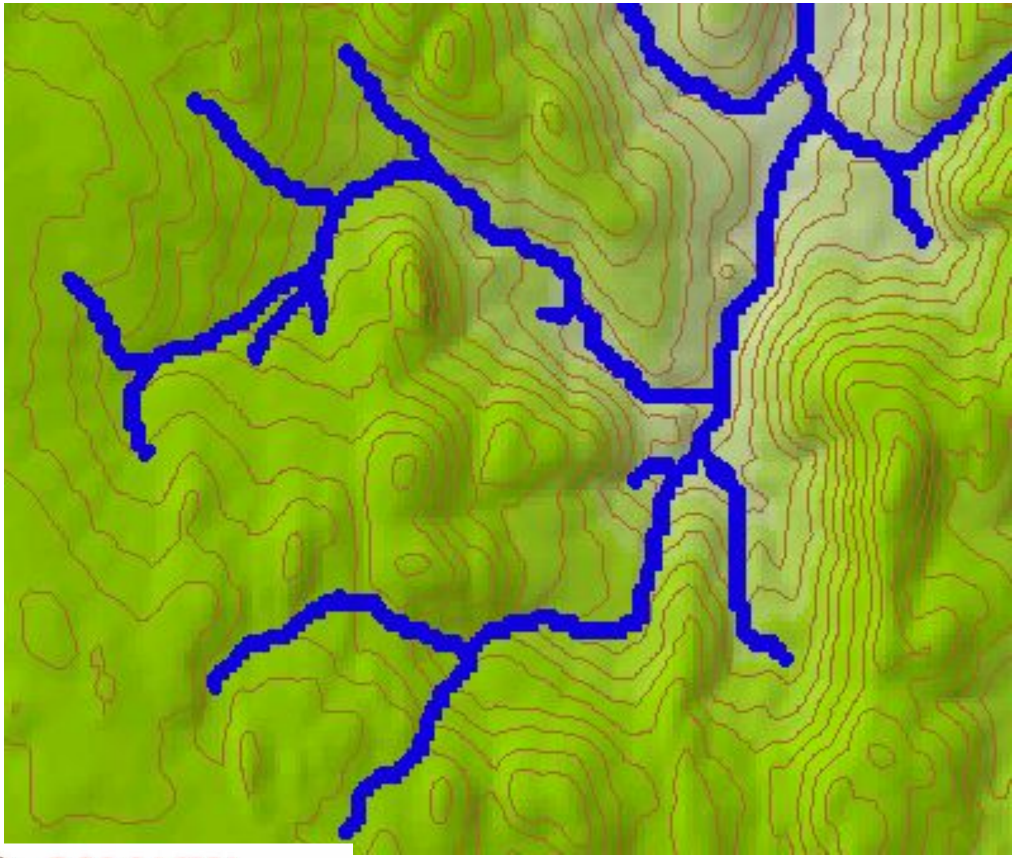
Watersheds …



Once done manually …

Contour lines (brown)
Drainage (blue)
Watershed boundary (red)

Flow accumulation as drainage network



Drainage network as defined by cells above threshold value for region.

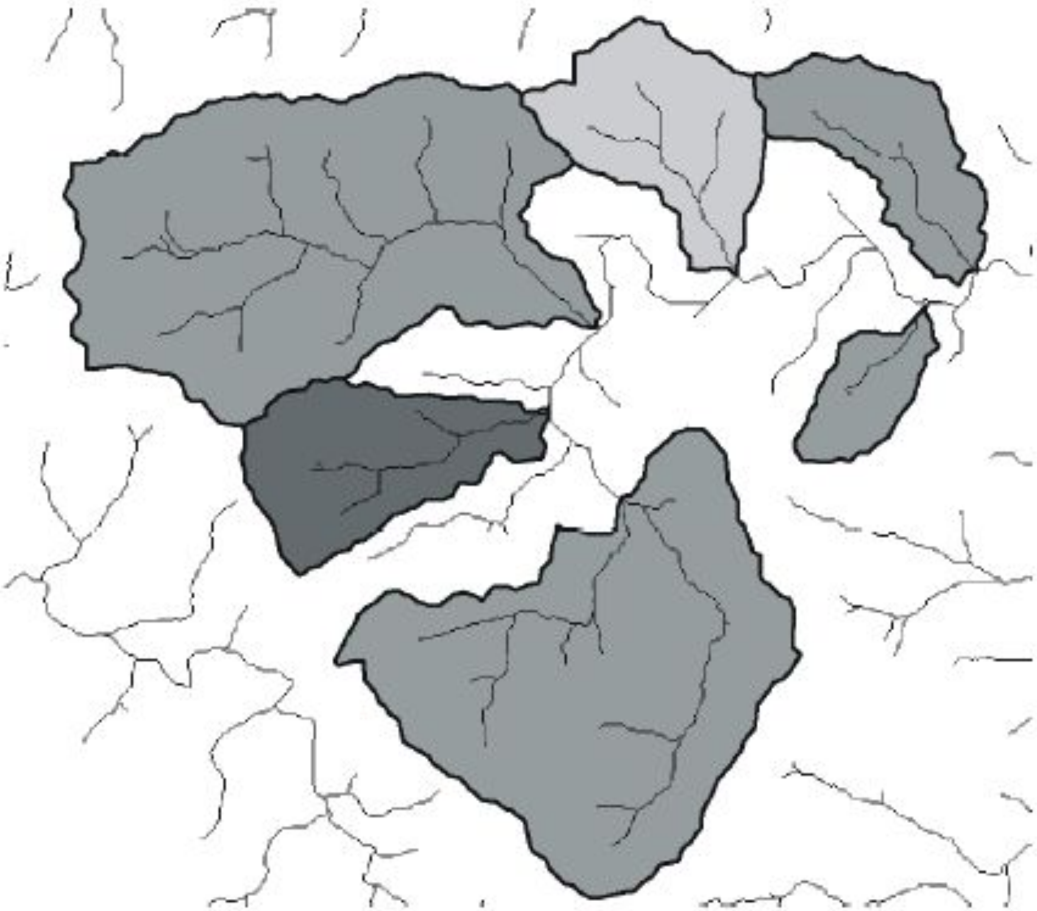


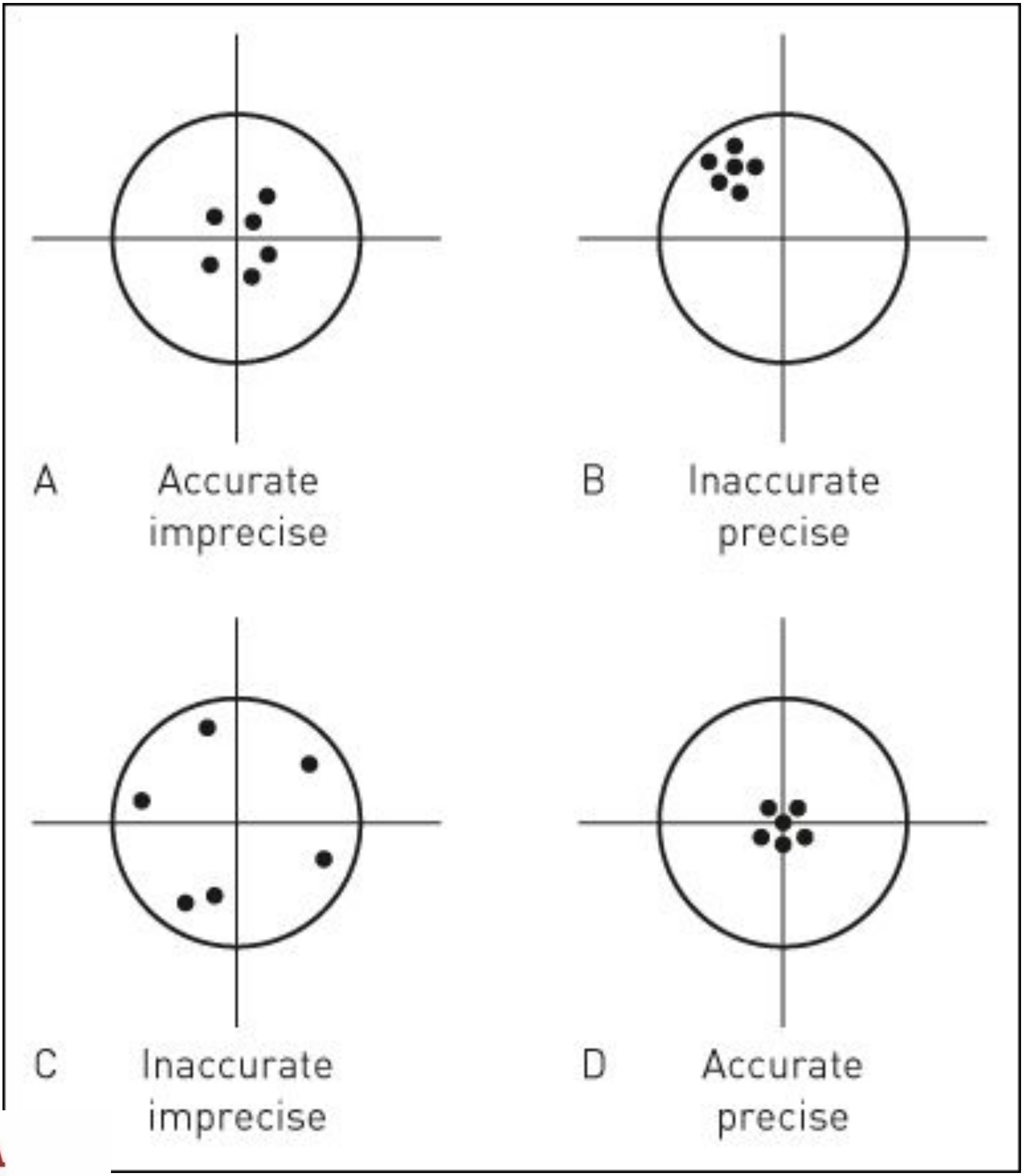
Figure 11-22: Drainage network and watersheds derived from DEM. Flow direction was determined from local aspect. Upstream watersheds and the probably drainage paths were then determined from flow direction.

SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

Concepts Related to Data Quality

- Related to individual data sets:
  - Errors – flaws in data
  - Accuracy – the extent to which an estimated value approaches the true value.
  - Precision – the recorded level of detail of your data.
  - Bias – the systematic variation of the data from reality.


A
Accurate
imprecise
B
Inaccurate
precise
C
Inaccurate
imprecise
D
Accurate
precise

Map scale

- Map scale specifies the amount of reduction between the real world and its graphic representation (usually a paper map).
- It is usually expressed as a ratio (eg 1:20,000), or equivalence (eg 1 mm = 20 m).
- Since a paper map is always the same size, its scale is fixed when it is printed, and cannot change.
- However, a map in a GIS can be shrunk or enlarged at will on the screen or on paper.
- You can zoom in until the screen displays a square metre or less, or zoom out.
- This means that geographic data in a GIS doesn't really have a 'map scale'.

Display scale

- The display scale of a map is the scale at which it 'looks right'.
- Because a paper map is created at certain scale, its 'map scale' and 'display scale' are the same.
- The display scale influences two things about a map :
  - The amount of detail. The map must not be overwhelmed with detail, and become too crowded.
  - The size and placement of text and symbols. These must be sized to be readable at the display scale, and placed so that they do not overlap each other.
- A GIS map's annotation (ie text and symbols) must be designed with a display scale, just like a paper map. There is a range of scale in which it will 'look right', even though it is possible to display it at other scales with the GIS software.

Data accuracy and uncertainty

Data accuracy is a statement of how closely a bit of data represents the real world. It applies to geographical information in all these ways:

- What features have been omitted ?
- What non-existent features are represented ?
- How correct is their classification ?
- How current is the data ?
- How far away is a map feature from its actual location in the world ?

- The last point is also called 'locational' accuracy and is generally stated in terms of uncertainty.
- For example, '95% of the well locations are within 50 metres of their surveyed locations'.
- A rigorous statement of accuracy will also include how and when the information was collected.
- Spatial data accuracy is independent of map scale and display scale, and should be stated in ground measurement units.

Data resolution

- Data resolution is the smallest difference between adjacent positions that can be recorded.
- Since a paper map is always the same size, its data resolution is tied to its scale.
- Resolution also limits the minimum size of feature that can be stored.
- Generally, a line cannot be drawn much narrower than about 1/2 a millimetre.
- Therefore, on a 1:20,000 scale paper map, the minimum distance which can be represented (resolution) is about 10 metres.
- On a 1:250,000 scale paper map, the resolution is 125 metres.
- However, most GIS store locations in ground units (eg UTM coordinates, or Longitude/Latitude) with a resolution of a centimetre or less.
- This resolution is far greater than the uncertainty of any of the map data.

Raster data resolution

- Raster data is stored as (usually square) pixels, which form a grid or mesh over an area of the earth.
- The size of these pixels determines the resolution of the raster, because it is impossible to store anything which falls 'between' the pixels.
- A GIS allows raster pixels to be any size, although they should not be smaller than the uncertainty of the data.
- If a raster coverage is derived from vector linework, its pixels should not be smaller than the uncertainty in the linework.
- If it comes from an air-photo or satellite image, its pixels should not be smaller than the resolution of the camera that recorded it.

Data detail

- Data detail is a measure of how much information is stored for each feature.
- A GIS stores lines (eg, a lake shoreline) as a sequence of point locations, and draws it with the edges that join them.
- There is no limit to how many points can be stored, or how close together they may be.
- The amount of detail on line features should be limited just like data density.
  It does not make sense to store points at intervals which are shorter than the accuracy of their locations.

GIS analysis

- In a GIS, analysis is done at the resolution of the data, not at any display scale.
- For example, the area of a habitat polygon is calculated to the nearest square centimetre.
- The GIS will carry much more resolution through its calculations than are justified by the data's accuracy.
- The results of these calculations should be rounded to a value appropriate to the uncertainty of the data for reporting.
- Some operations may result in features which are smaller than the data uncertainty.
- For example, overlaying rivers and forest polygons may create 'slivers' along the riverbanks which are 10 metres wide, when the uncertainty of the data is 20 metres. These slivers should be ignored, or included with their neighbours before the results of the overlay are used for further analysis.

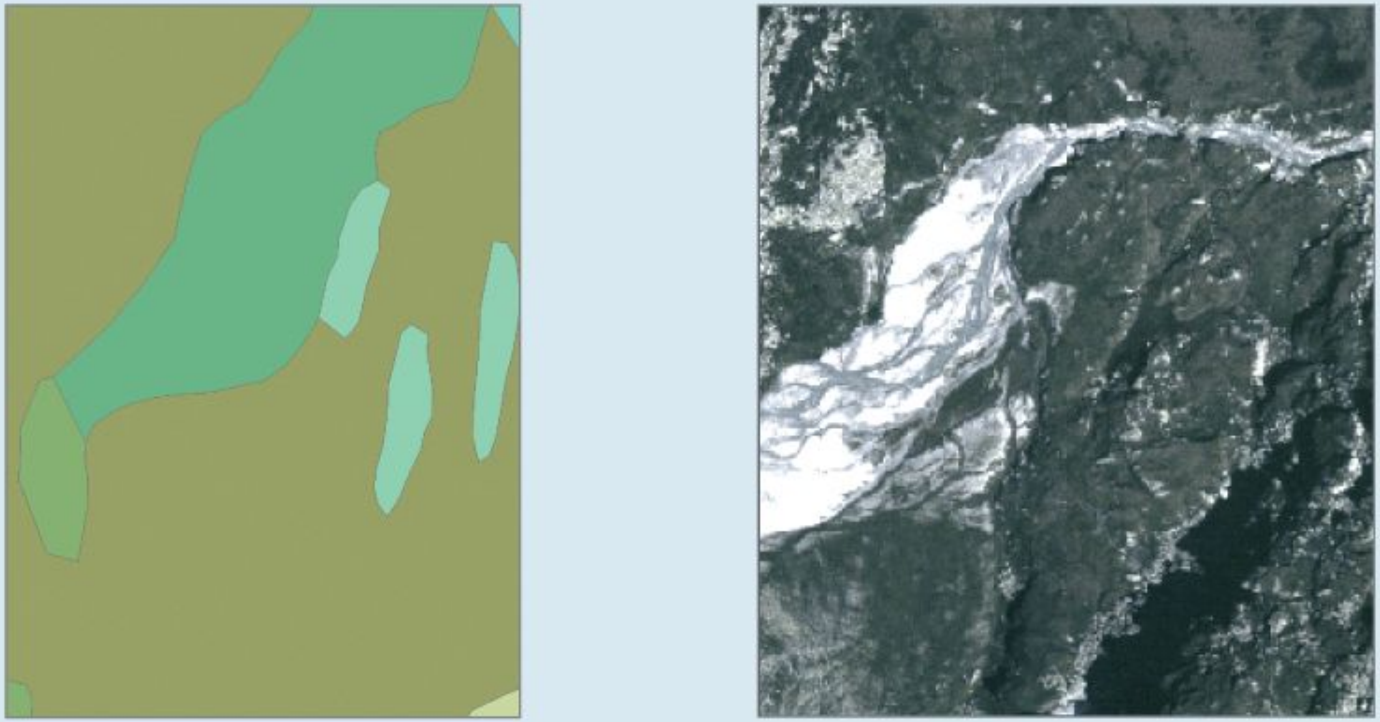

Canada
Sliver
Gap
United States Of America

Generalization

- In a GIS, it is possible to create a new coverage by reducing the amount of detail in existing coverage.
- This 'generalizing' may or may not reduce the number of objects in the coverage.
- For example, a detailed forest cover map may be generalized by combining polygons with similar characteristics.
- This reduces the number of objects in the coverage.
- Conversely, a detailed ecosystem classification map may be generalized by reducing the amount of detail in the boundaries between regions, without reducing the number of regions.
- Generalizing a raster image usually reduces both the number of objects, and the amount of detail.

Concepts Related to Data Quality

- Related to source data:
  - Resolution – the smallest feature in the data set that can be displayed.
  - Generalization- simplification of objects in the real world to produce scale models and maps.



(a) 25 m resolution vegetation map (b) 5 m resolution colour aerial photograph

Resolution and generalization of raster datasets

Data Sets Used for Analysis

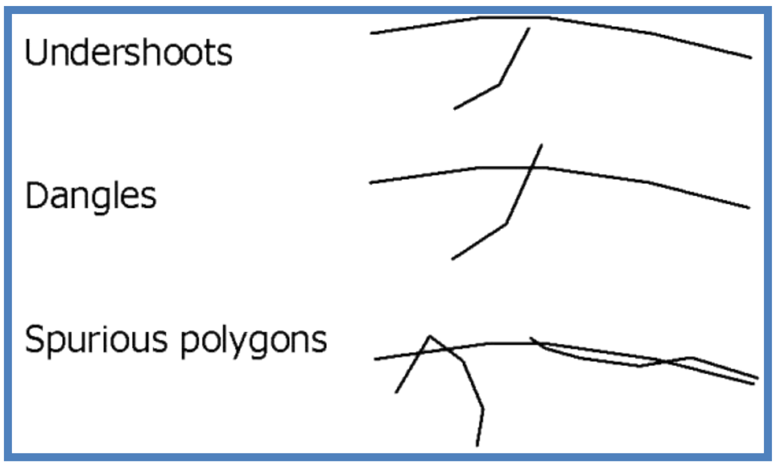
Must be:

- Complete – spatially and temporally
- Compatible – same scale, units of measure, measurement level
- Consistent – both within and between data sets.
- And Applicable for the analysis being performed.

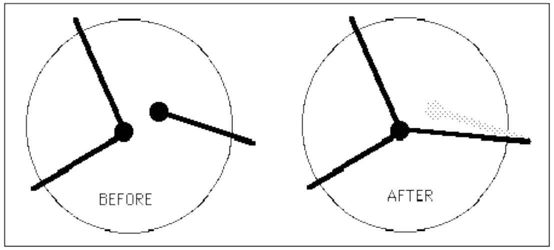
Any digitized map requires:
Considerable post-processing
Check for missing features
Connect lines
Remove spurious polygons
Some of these steps can be automated

Digitizing Error

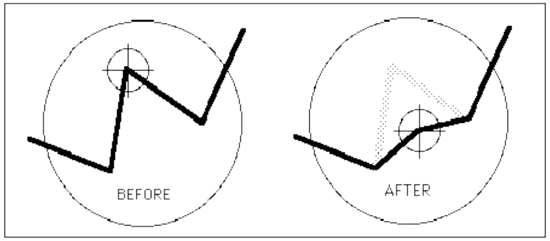
A great deal of spatial data has been digitized from paper maps.



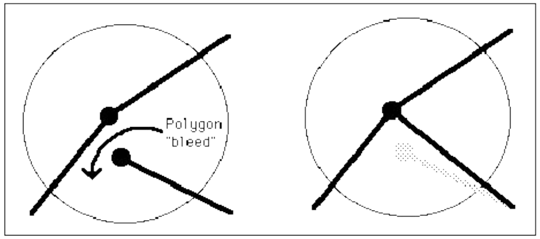
Undershoot error:
Use Move Node tool to snap the lines



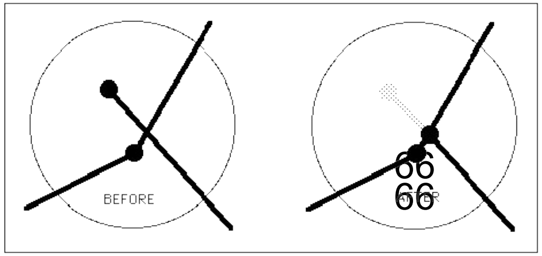
Misplaced points:
Use Move Node tool to interactively select and reposition misplaced points



Polygon closure error:
If polygon boundaries are not closed, areas are not defined and color fills can bleed out into surrounding areas



Overshoot :
Snap overshoots the same way as undershoots, or create a new intersection node for snapping



SOMAIYA
VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

66

Sources of Error in Spatial Data

- Accuracy of Content
- Measurement Errors
- Field Data
- Laboratory Errors
- Locational Accuracy
- Natural Spatial Variation

Uncertainty in the representation of geographic phenomena

- Representation is closely related to measurement.
- Representation is not just an input to analysis, but sometimes also the outcome of it. For this reason, we consider representation separately from measurement.
  - The world is infinitely complex, but computer system are finite.
  - Representation is all about the choices that are made in capturing knowledge about the world

  - Uncertainty in earth model: ellipsoid models, datum, projection types
  - Uncertainty in the raster data model (structure)
  - Uncertainty in the vector data model (structure)

Factors affecting the quality of Spatial data

1. Currency
   Are data up to date?
   Time series
2. Completeness
   Areal coverage—is it partial or complete?
3. Consistency
   Map scale
   Standard descriptions
   Relevance
4. Accessibility
   Format
   Copyright
   Cost

5. Accuracy and Precision.
    - Density of observations
    - Positional accuracy
    - Attribute accuracy—qualitative and quantitative
    - Topological accuracy
    - Lineage—When collected, by whom, how?

6. Sources of errors in data
    - Data entry or output faults
    - Choice of the original data model
    - Natural variation and uncertainty in boundary location and topology
    - Observer bias
    - Processing
        - Numerical errors in the computer
        - Limitations of computer representations of numbers

7. Sources of errors in derived data and in the results of modelling and analysis
    - Problems associated with map overlay
    - Classification and generalization problems
    - Choice of analysis model
    - Misuse of logic
    - Error propagation
    - Method used for interpolation

Factors affecting the reliability of Spatial Data

- Age of Data
- Areal Coverage
- Map scale and Resolution
- Density and Observation
- Relevance
- Data Format, Data Exchange and Interoperability
- Accessibility
- Cost and Copyrighting
- Numerical Errors in the computer
- Rounding Errors
- Geographical Coordinates and Precision

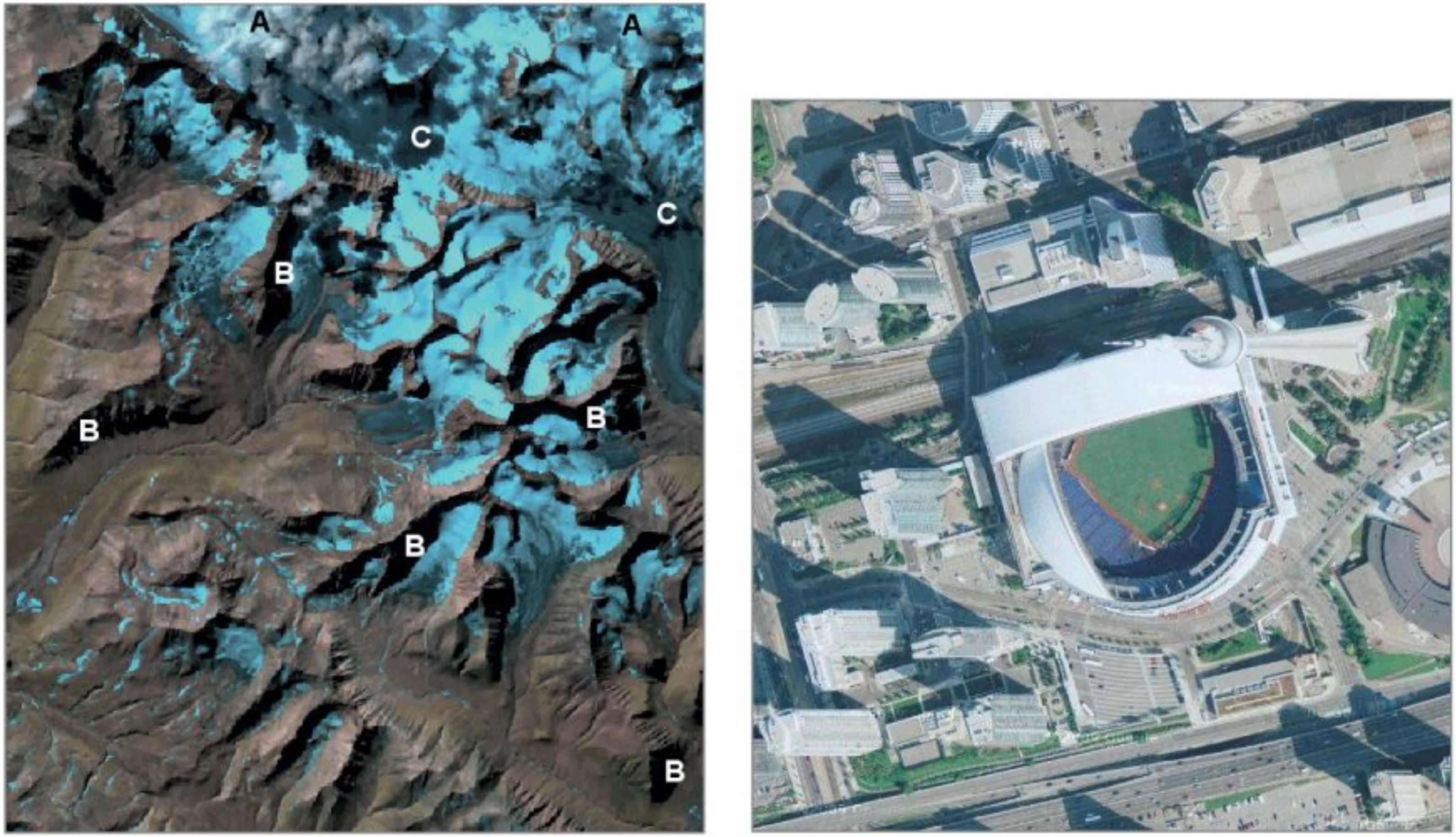


Figure 10.8 Problems with remotely sensed imagery: (left) example of a satellite image with cloud cover (A), shadows from topography (B), *Source*: Ian Bishop (left) and Google UK (right)

Uncertainty in the vector data structure

- Socioeconomic data—facts about people, houses, and households—are often best represented as points.
- For various reasons (to protect privacy, to limit data volume), data are usually aggregated and reported at a zonal level, such as census tracts or ZIP Codes.
- This distorts the data in two ways:
  - First, it gives them a spatially inappropriate representation (polygons instead of points);
  - Second, it forces the data into zones whose boundaries may not respect natural distribution patterns.

SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

Map Representation Error

Map scale	Ground distance, accuracy, or resolution (corresponding to 0.5 mm map distance)
1:1,250	0.625 m
1:2,500	1.25 m
1:5,000	2.5 m
1:10,000	5 m
1:24,000	12 m
1:50,000	25 m
1:100,000	50 m
1:250,000	125 m
1:1,000,000	500 m
1:10,000,000	5 km