

Raster Analysis

Overview

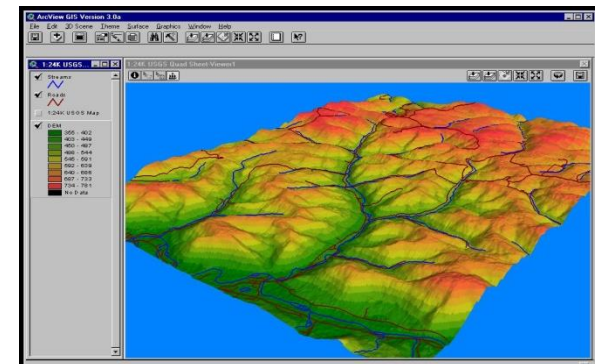
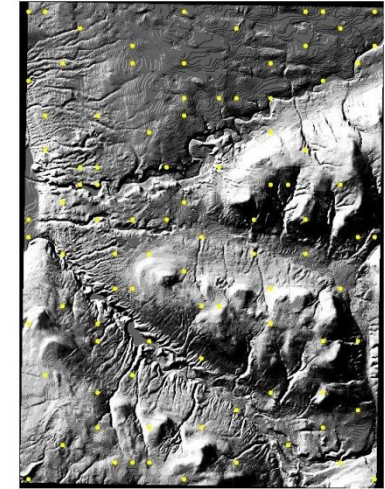
Neighborhood Analysis

Overlay

Cost Surfaces

Why we use Raster GIS

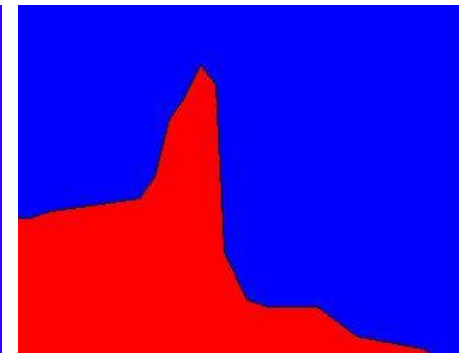
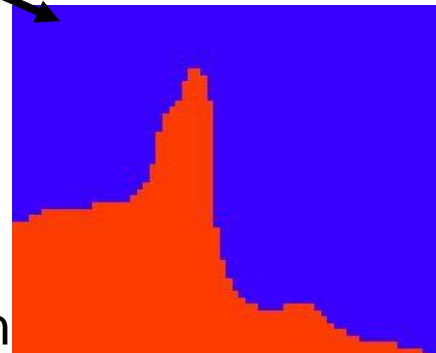
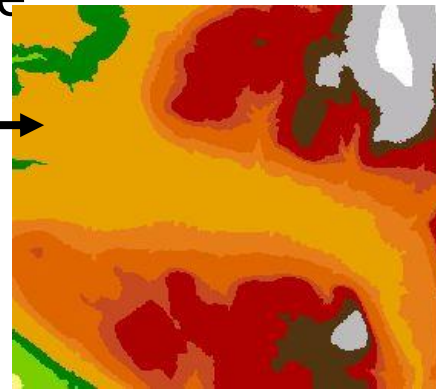
- In our previous discussion of data models, we indicated that Raster GIS is often used because:
 - Raster is better suited for spatially continuous data like elevations
 - Raster is better for creating visualizations and modeling environmental phenomena
 - Other continuous data may include: pH, air pressure, temperature, salinity, etc..
 - Raster data is a simplified realization of the world, and allows for fast and efficient processing
- A raster GIS performs geoprocessing tasks on a grid based realization of the world



Basic Grid Manipulation

Import DEM
Reclassify IthacaDEM
Convert R2V
Mask

- Raster GIS will allow a user to perform basic manipulations on a raster, or grid data set. Some basic manipulations include:
 - Reclassify
 - Convert
 - Preparation for Analysis
 - Set extent
 - Mask (the cousin of what we did with CLIP)
 - Watch out for coordinate systems!!!
 - Also, just because the data lines up doesn't mean they are the same base resolution



Surface Analysis Functions

- Inflection (profile curvature)
 - Curvature of the surface in the direction of slope
 - Can be used to describe the physical characteristics of a drainage basin in an effort to understand erosion and runoff processes
 - **Slope** affects the overall rate of movement down slope.
 - **Aspect** defines the direction of flow.
 - **Profile curvature** affects the acceleration and deceleration of flow, and therefore influences erosion and deposition

Surface Analysis Functions

- Slope

- Rate of change in elevation
 - Calculated as “rise over run”
 - Change in Z over distance
 - Tangent provides slope in degrees

42	45	47		
40	44	49		
44	48	52	dz/dx	-0.19375
			dz/dy	0.48125
a	b	c	rise/run	0.518788
d	e	f	degree slope	27.41973
g	h	i		

- Computed using a 3 X 3 moving window (slope request)

- Subtract 8 neighboring cell elevations from the center cell and divide by the distance

- Choose steepest slope
- Choose average slope:

$$(dz/dx) = ((a + 2d + g) - (c + 2f + i)) / (8 * x_mesh_spacing)$$

$$(dz/dy) = ((a + 2b + c) - (g + 2h + i)) / (8 * y_mesh_spacing)$$

$$rise_run = \text{SQRT}(\text{SQR}(dz/dx) + \text{SQR}(dz/dy))$$

$$degree_slope = \text{ATAN}(rise_run) * 57.29578$$

Example: compute slope and aspect

Calculating Slope

Four nearest cells
elevation values

42	45	47
40	44	49
44	48	52

← C = 10 →

kernal for dZ/dx

Z ₁	Z ₂	Z ₃
0	0	0
Z ₄	Z ₀	Z ₅
-1	0	1
Z ₆	Z ₇	Z ₈
0	0	0

$$dZ/dx = (Z_5 - Z_4)/2C$$

$$dZ/dx = (49 - 40)/20 = 0.45$$

kernal for dZ/dy

Z ₁	Z ₂	Z ₃
0	1	0
Z ₄	Z ₀	Z ₅
0	0	0
Z ₆	Z ₇	Z ₈
0	-1	0

$$dZ/dy = (Z_2 - Z_1)/2C$$

$$dZ/dy = (45 - 48)/20 = -0.15$$

$$\text{slope} = \text{atan}[(0.45)^2 + (-0.15)^2]^{0.5} = 25.3^\circ$$

3rd-order finite difference
elevation values

42	45	47
40	44	49
44	48	52

← C = 10 →

kernal for dZ/dx

Z ₁	Z ₂	Z ₃
-1	0	1
Z ₄	Z ₀	Z ₅
-2	0	2
Z ₆	Z ₇	Z ₈
-1	0	1

$$dZ/dx = [(Z_3 - Z_1) + 2(Z_5 - Z_4) + (Z_8 - Z_6)]/8C$$

$$dZ/dx = [(47 - 42) + 2(49 - 40) + (52 - 44)]/80 = 0.39$$

kernal for dZ/dy

Z ₁	Z ₂	Z ₃
1	2	1
Z ₄	Z ₀	Z ₅
0	0	0
Z ₆	Z ₇	Z ₈
-1	-2	-1

$$dZ/dy = [(Z_1 - Z_6) + 2(Z_2 - Z_7) + (Z_3 - Z_8)]/8C$$

$$dZ/dy = [(47 - 52) + 2(45 - 48) + (42 - 44)]/80 = -0.16$$

$$\text{slope} = \text{atan}[(0.39)^2 + (-0.16)^2]^{0.5} = 22.9^\circ$$

Calculating slope

What is slope?

Slope identifies the steepest downhill slope for a location on a surface. Slope is calculated for each triangle in TINs and for each cell in rasters. For a TIN this is the maximum rate of change in elevation across each triangle. For rasters it is the maximum rate of change in elevation over each cell and its eight neighbors.

The slope command takes an input surface raster and calculates an output raster containing the slope at each cell. The lower the slope value, the flatter the terrain; the higher the slope value, the steeper the terrain. The output slope raster can be calculated as percent slope or degree of slope.

Degree of slope = θ Percent of slope = $\frac{\text{rise}}{\text{run}} * 100$

$\frac{\text{rise}}{\text{run}} = \tan \theta$

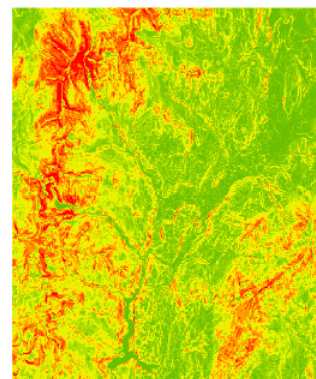
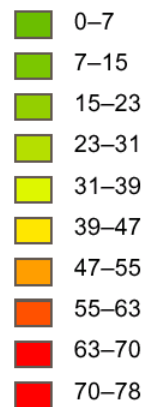
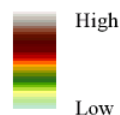
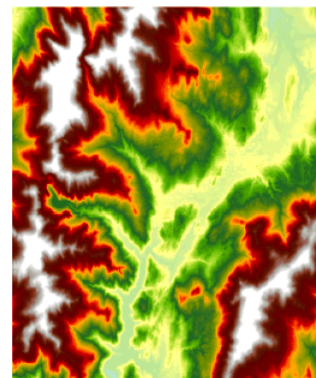


Degree of slope =	30	45	76
Percent of slope =	58	100	375

When the slope angle equals 45 degrees, the rise is equal to the run. Expressed as a percentage, the slope of this angle is 100 percent. Note that as the slope approaches vertical (90°), the percentage slope approaches infinity.

The slope function is most frequently run on an elevation grid, as the following diagrams show. Steeper slopes are shaded red on the output slope map.

Elevation grid



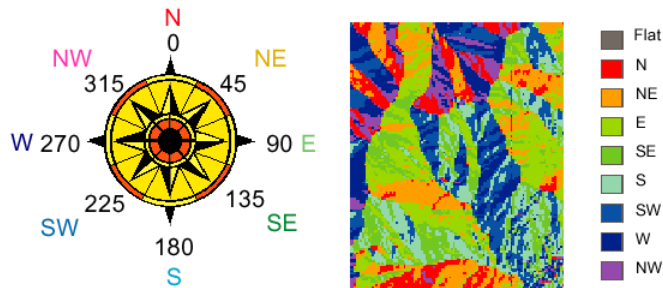
Slope map (in degrees)

Calculating aspect

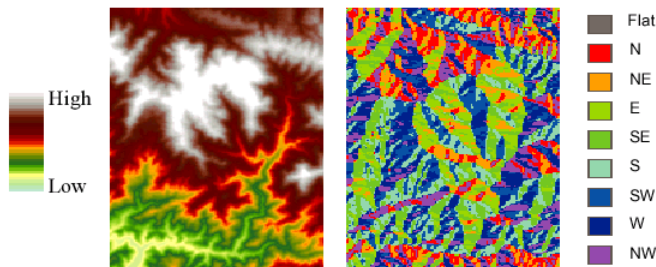
What is the aspect?

Aspect is the direction that a slope faces. It identifies the steepest downslope direction at a location on a surface. It can be thought of as slope direction or the compass direction a hill faces. Aspect is calculated for each triangle in TINs and for each cell in rasters.

Aspect is measured counterclockwise in degrees from 0 (due north) to 360 (again due north, coming full circle). The value of each cell in an aspect grid indicates the direction in which the cell's slope faces. Flat slopes have no direction and are given a value of -1.



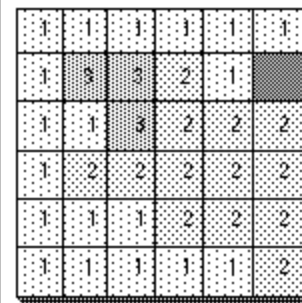
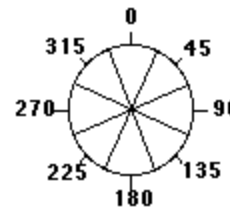
The diagram below shows an input elevation grid and the output aspect grid.



Why use the aspect function?

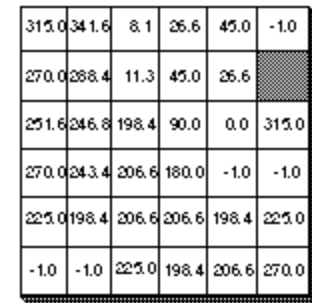
There are many different reasons to use the aspect function. For instance, you may want to:

- Find all north-facing slopes on a mountain as part of a search for the best slopes for ski runs.
- Calculate the solar illumination for each location in a region as part of a study to determine the diversity of life at each site.
- Find all southerly slopes in a mountainous region to identify locations where the snow is likely to melt first, as part of a study to identify those residential locations that are likely to be hit by meltwater first.
- Identify areas of flat land to find an area for a plane to land in case of emergency.



INGRID 1

=



OUTGRID

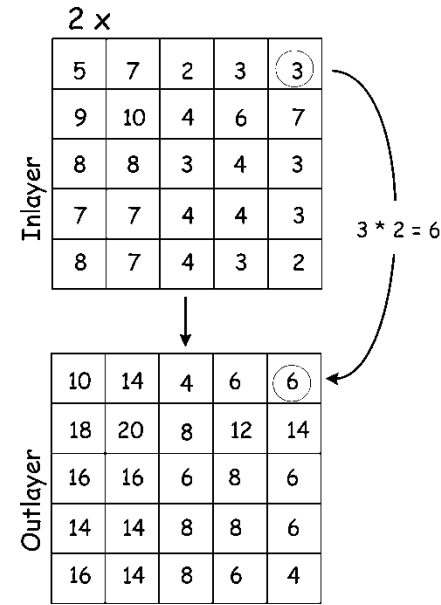
Value = NODATA

White Cells
Value = 0

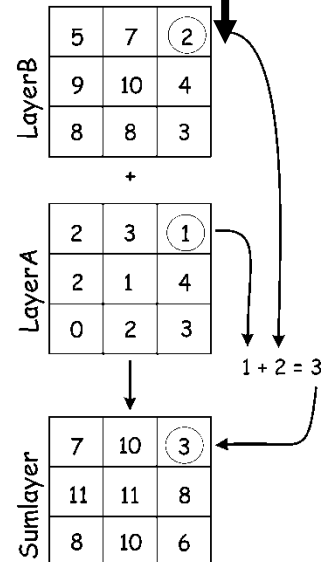
Map Algebra

- Map algebra is a cell by cell combination of raster layers using mathematical operations
 - Unary – one layer
 - Binary – two layers
- Basic Mathematical Operations
 - Addition, subtraction, division, max, min, virtually any mathematical operation you would find in an Excel spreadsheet
- Strong analytical functions

- Some Map Algebra Commands in ARC/INFO
 - `Outgrid = grid1 + grid2`
 - `Outgrid = grid1 * 2`
 - `Outgrid = sin(grid1)`
 - `Outgrid = costallocation(sourcegrid, costgrid, accumgrid, backgrid)`
 - `Outgrid = con(>5 (ingrid1), 0, ingrid1)`
 - `Outgrid = select(grid1, 'VALUE = 10')`



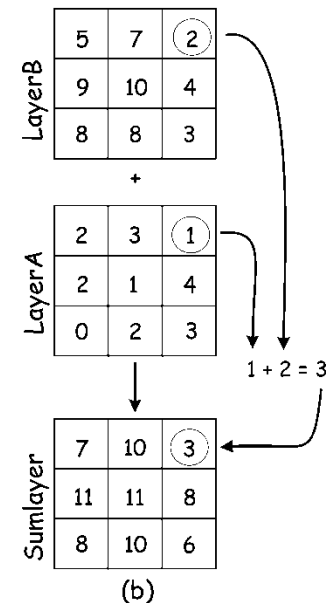
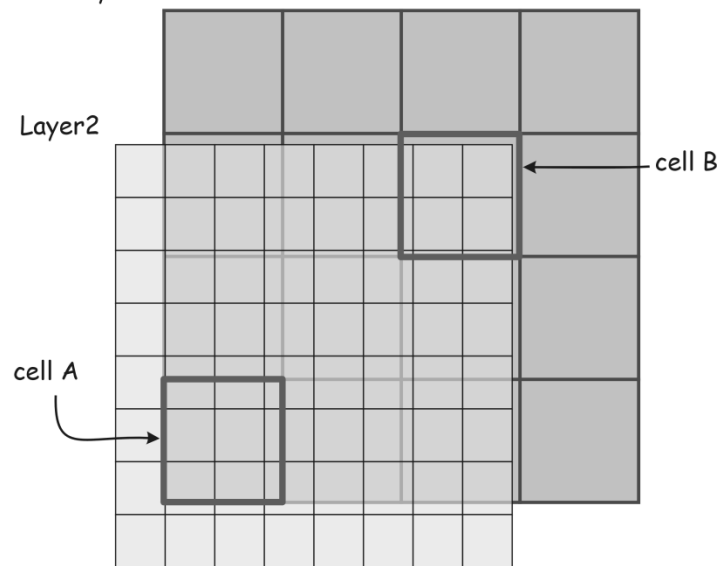
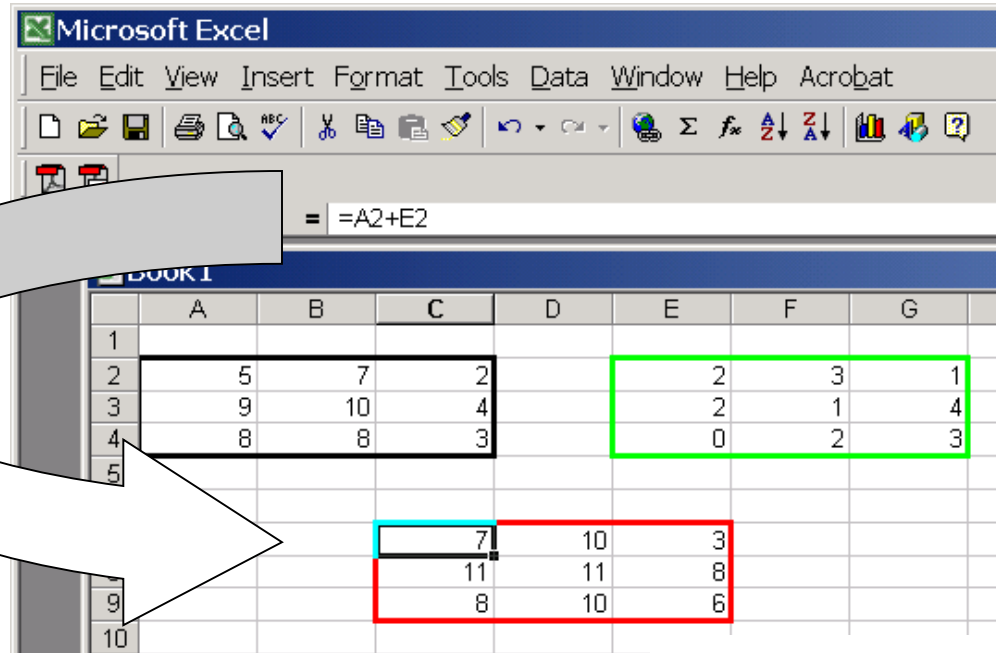
(a)



(b)

Map Algebra

- Map algebra and raster GIS is quite simple to visualize in a spreadsheet. An example of multiplication and addition
- The use of arrays make map algebra and raster GIS very computationally efficient
- But, be careful of:
 - Layers that are not coincident
 - Different cell sizes



© Paul Bolstad, GIS Fundamentals

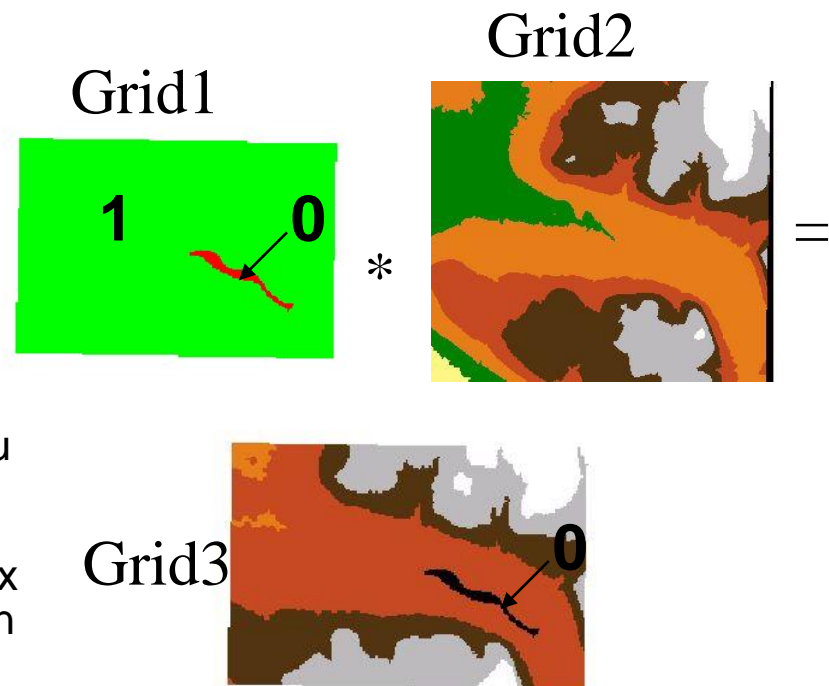
© Arthur J. Lembo, Jr.
Salisbury University

Salisbury
UNIVERSITY

Map Algebra

- Map algebra can be extended to performing a number of mathematical operations.
- The computer will allow you to perform virtually any mathematical calculation
 - beware: some will make sense, others won't.
 - For example, you can create a grid where water features are 0 and land values are 1. Then, you can multiply this grid with an elevation map. The output will include 0's where water existed ($x * 0 = 0$), and the original elevation value where land existed ($x * 1 = x$)
 - Or, you can add the elevations and the grid with 0's and 1's together (but, it would be meaningless!)

$$\text{Grid1} * \text{Grid2} = \text{Grid3}$$



Can't build in the 0 area since there is water

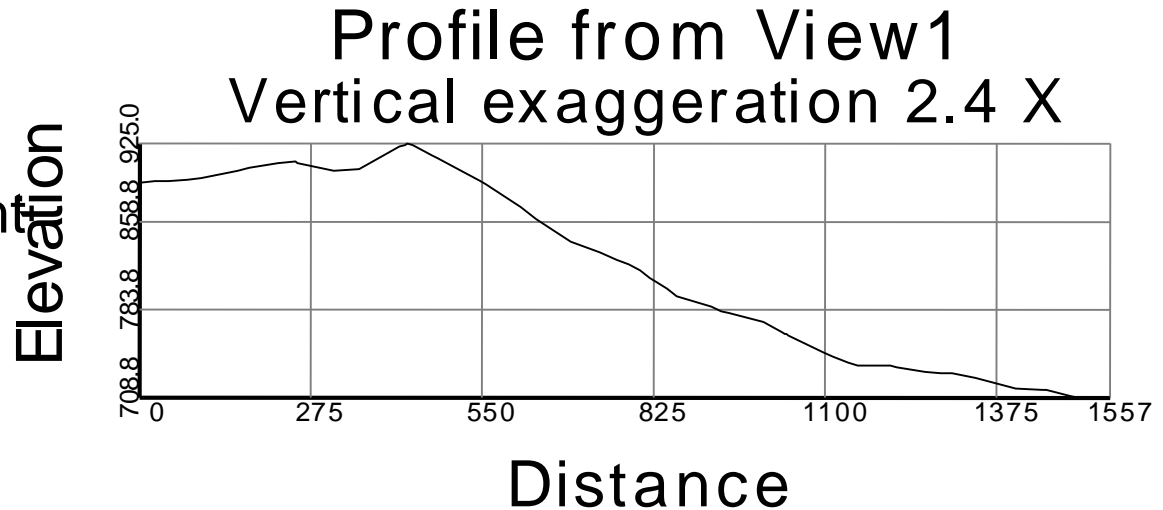
Simple process

K:\GEOG\student_research\CUDA

- Reclassify DEM
- Compute Slope
- Reclassify Slope
- Multiply slope * DEM * Landuse

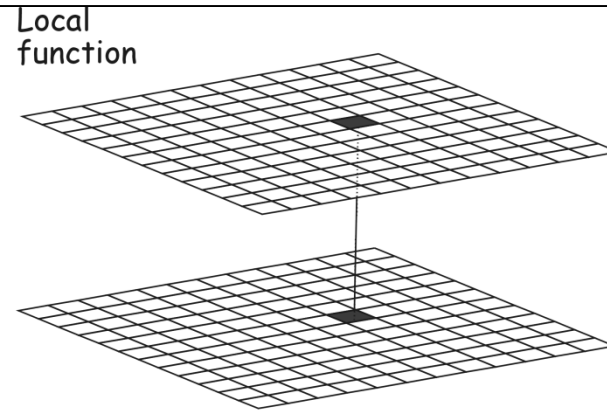
Surface Analysis Functions

- Profile
 - easily see and measure any changes in height along a line
 - evaluating the difficulty of mountain trails
 - or assessing the viability of a corridor for railway lines
 - *Example: a hike through north campus – 3D Analyst (interpolate line)*



Raster Functions

- **Local:** only use data in a single cell to calculate an output value (what we typically think of as map algebra)
- **Neighborhood (Focal):** use data from a set of cells, most often a “kernel”
- **Global:** use all data from a raster layer



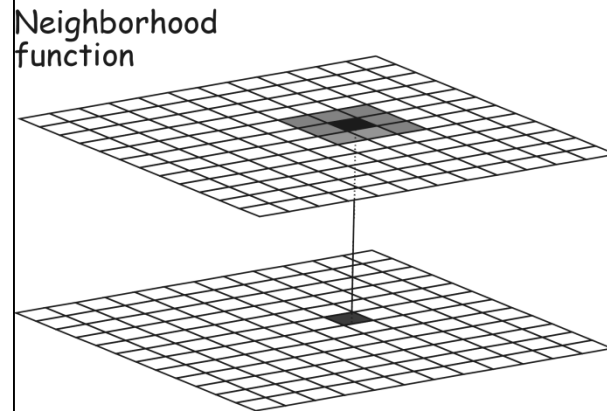
e.g.,

10	12	42
30	9	4
-12	8	15

plus 4

↓

14	16	46
34	13	8
-8	12	19



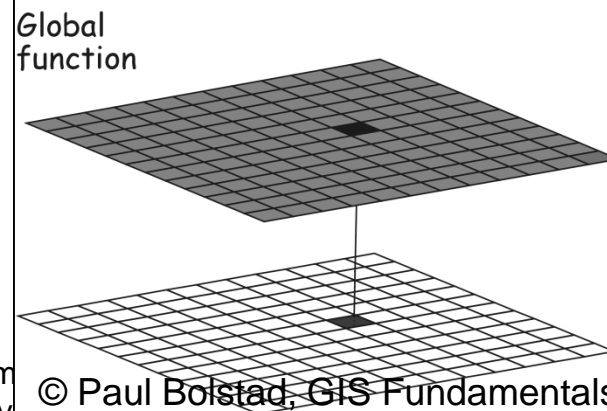
e.g.,

10	12	42
30	9	4
-12	8	15

neighborhood maximum

↓

33	42	42
30	42	42
30	30	17



e.g.,

10	12	42
30	9	4
-12	8	15

global maximum

↓

42	42	42
42	42	42
42	42	42

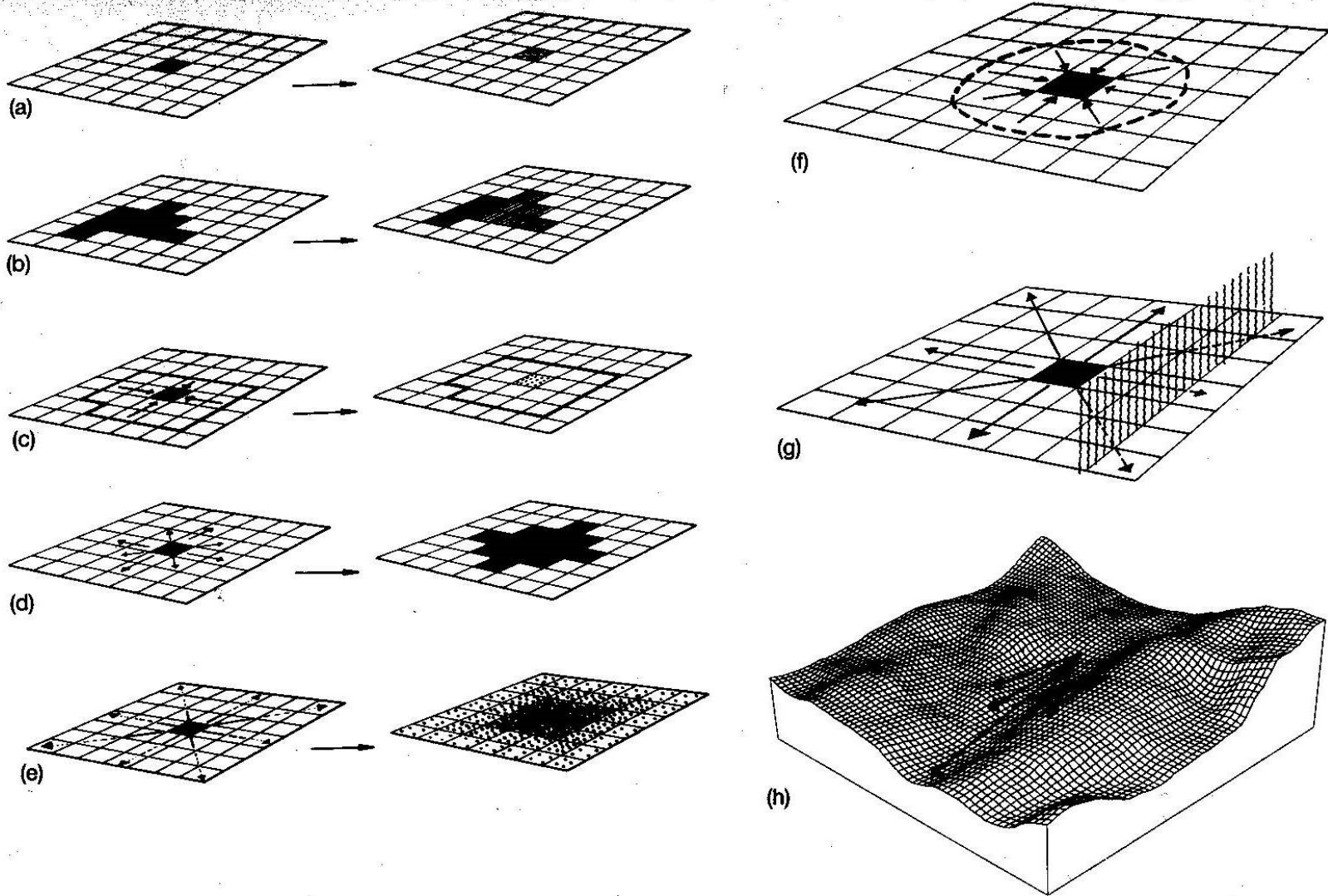


Fig. 5.8 Transformation operations in overlay analysis.

(a) Renumbering or recolouring a point.

(b) Renumbering or recolouring a region.

(c) The point receives a value that reflects a property of the region.

(d) Spreading isotropically from a point.

(e) Spreading with inverse distance weighting from a point.

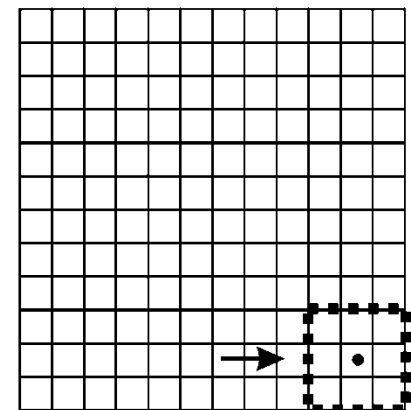
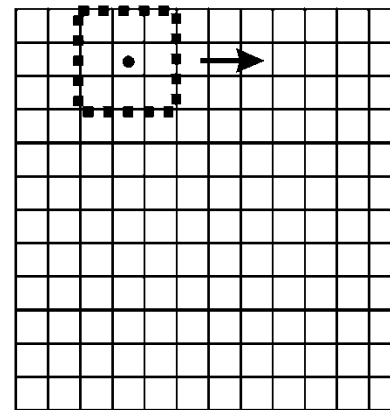
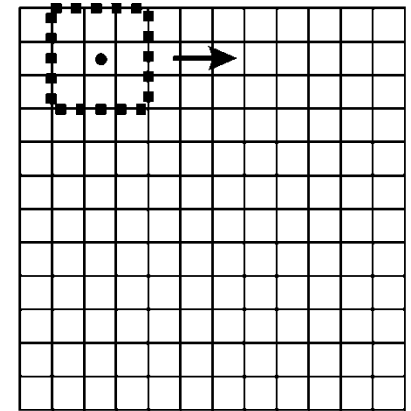
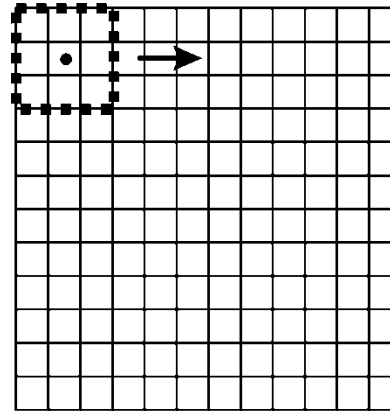
(f) Interpolating a value at a point from its surroundings.

(g) Spreading from a point through a barrier.

(h) Spreading from a point over a surface.

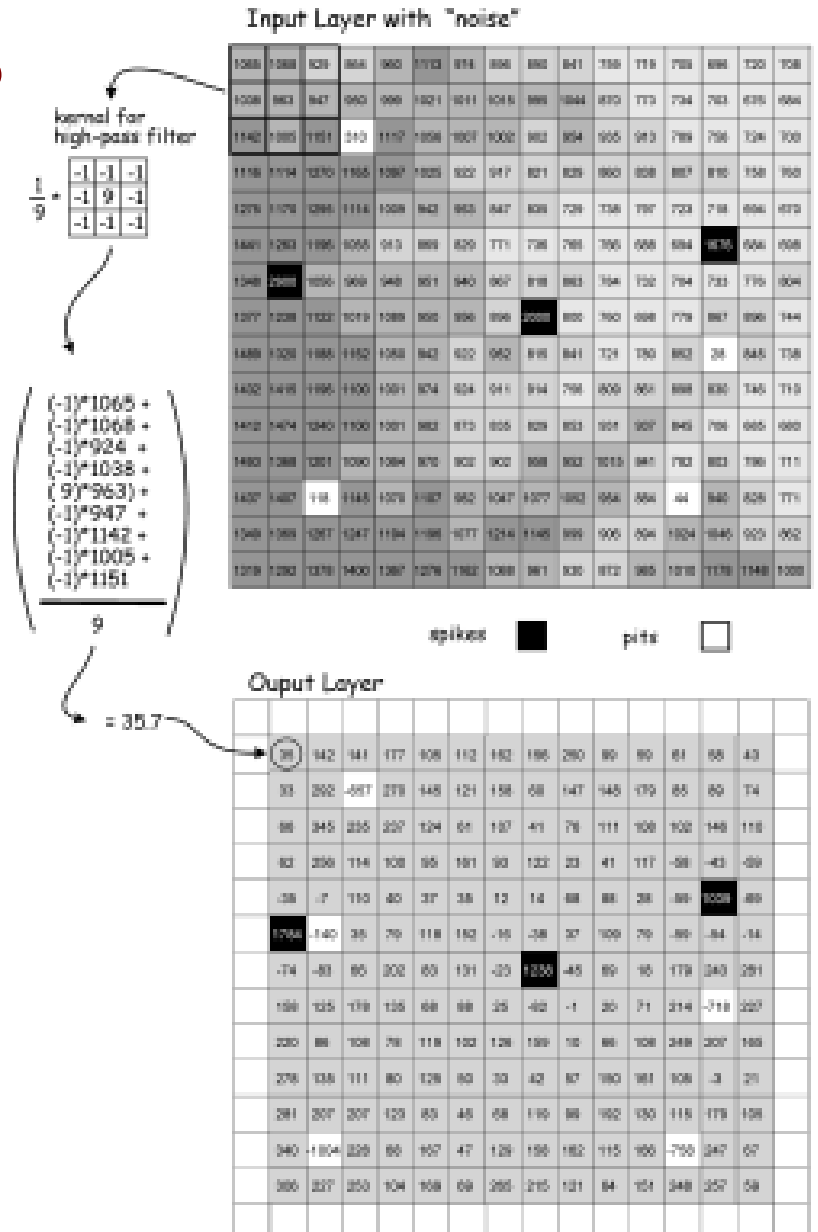
Moving Windows

- Useful for calculating local statistical functions or edge detection
- Kernel: a set of constants applied with a function, such as $1/9$ being the mean of the center cell.
- Other configurations may be used when dealing with diagonal or adjacent cells



Moving Windows

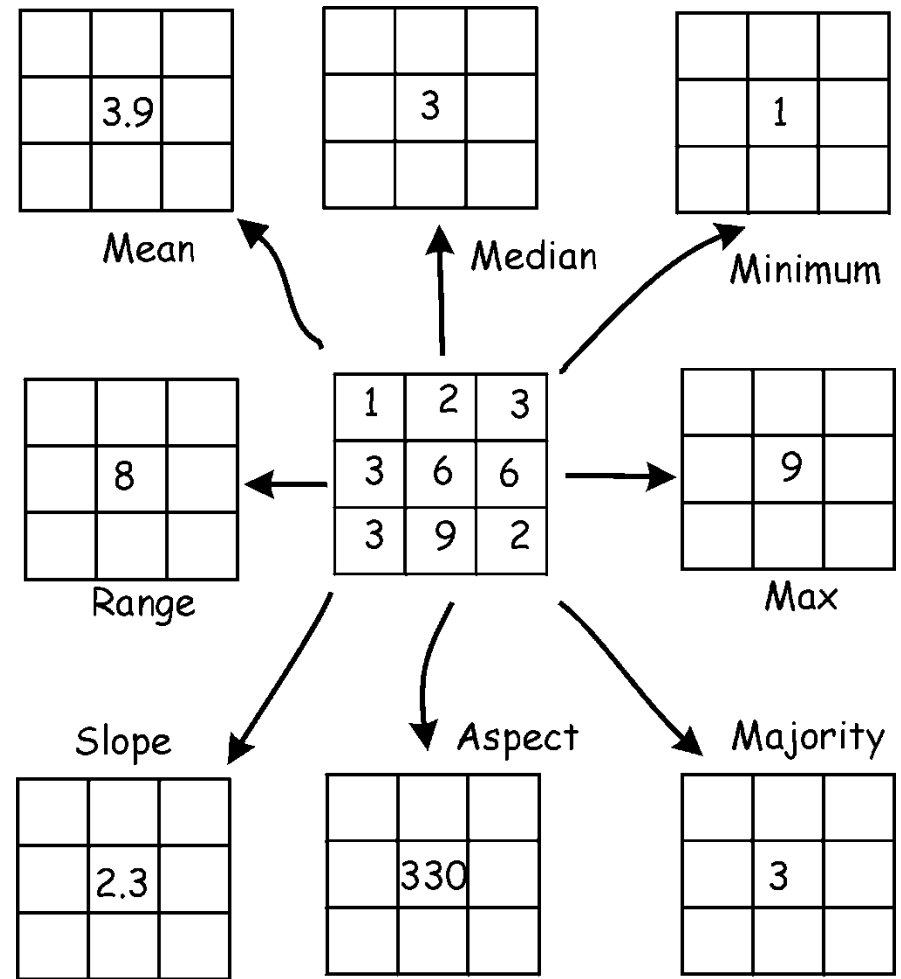
- Noise removal
 - Noise may be erroneous data values, or spikes we wish to remove
 - Gores, or spikes in a DEM may be removed through filtering and smoothing



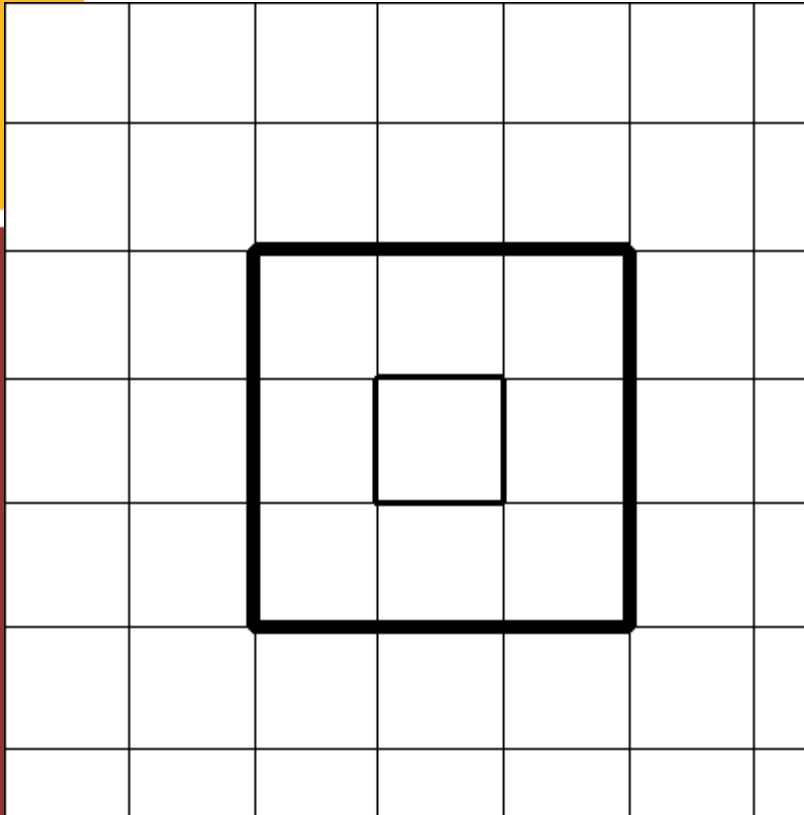
Raster Analysis – overlay and cost surfaces

A look at some raster functions

- Tour of ARC/INFO Grid help
- Example raster analysis in ArcGIS
 - Implemented through the Raster calculator, or the command line in ArcInfo
 - Addition of scripts to automate tasks
 - Docell



```
DOCELL
if (ingrid1 > 5 & ingrid < 50) outgrid = 500
else if (ingrid1 == 50) outgrid = 700
else if (ingrid1 > 50 & ingrid < 100) outgrid = 800
else outgrid = 1000
END
```



-1,-1	0,-1	1,-1
-1, 0	0, 0	1, 0
-1, 1	0, 1	1, 1

$$Y = [(x_{ij} - x_{00})^2]^{1/2}$$

where,

- Y = estimated terrain ruggedness index (“tri”) of cell (0,0)**
- x_{ij} = elevation of neighbor cell to cell (0,0)**

Terrain Ruggedness Index

docell

```
ssdiff := ((sqr (el (0, 0) - el (-1, -1))) + (sqr (el (0, 0) - el (0, -1)))  
           + (sqr (el (0, 0) - el (1, -1))) + (sqr (el (0, 0) - el (-1, 0)))  
           + (sqr (el (0, 0) - el (1, 0))) + (sqr (el (0,0) - el (-1,1)))  
           + (sqr (el (0, 0) - el (0, 1))) + (sqr (el (0, 0) - el (1, 1))))
```

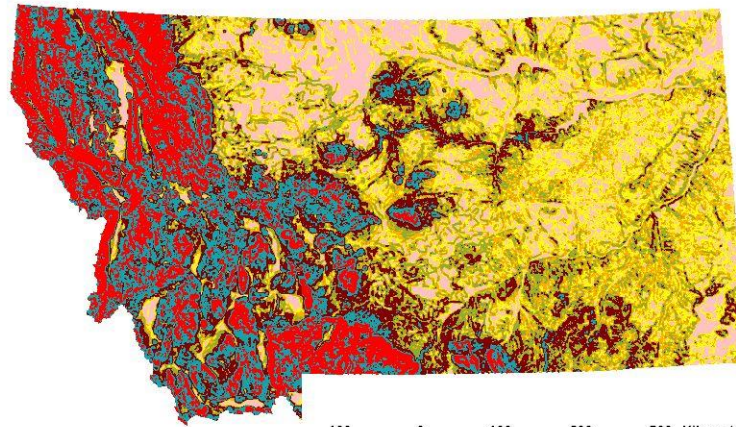
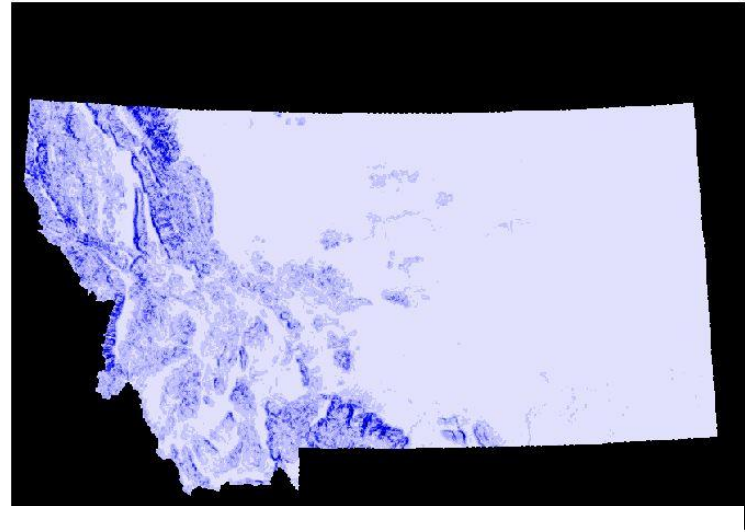
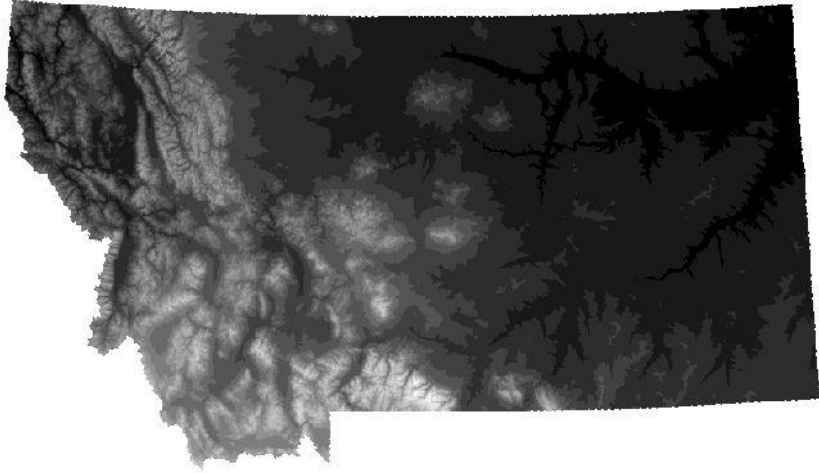
```
tri = sqrt (ssdiff)
```

end

ssdiff = temporary scalar, “sum squared difference” (square meters).

tri = terrain roughness index (meters).

el = name of elevation grid (meters).



- Terrain Ruggedness Index**
- Level
 - Nearly Level
 - Slightly Rugged
 - Intermediately Rugged
 - Moderately Rugged
 - Highly Rugged
 - Extremely Rugged



100 0 100 200 300 Kilometers

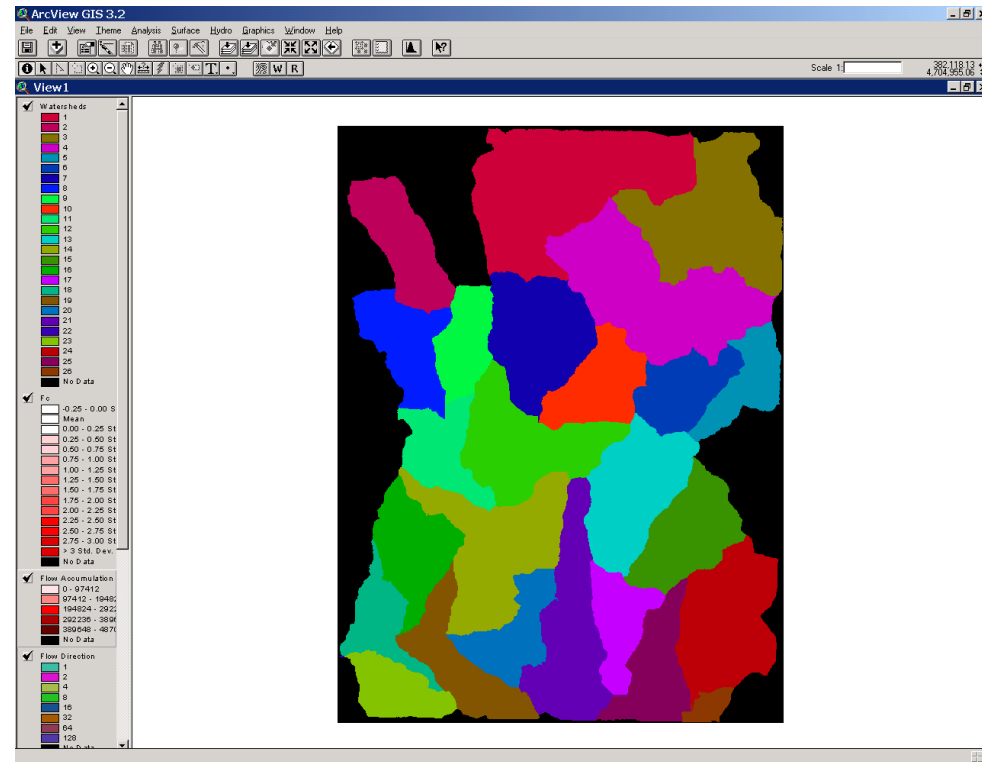
A tour of raster functions in ArcGIS

- Spatial Analyst Tools
 - Conditional
 - Density
 - Distance
 - Corridor
 - Cost Allocation
 - Cost Back Link
 - Cost Distance
 - Cost Path**
 - Euclidean Allocation
 - Euclidean Direction
 - Euclidean Distance
 - Extraction
 - Generalize
 - Groundwater
 - Hydrology
 - Basin
 - Fill
 - Flow Accumulation**
 - Flow Direction
 - Flow Length
 - Sink
 - Snap Pour Point
 - Stream Link
 - Stream Order
 - Stream to Feature
 - Watershed
 - Interpolation
 - IDW
 - Krige
 - Natural Neighbor
 - Spline
 - TopoToRaster
 - TopoToRasterByFile
 - Trend

- Local
 - Map Algebra**
 - Math
 - Multivariate
 - Neighborhood
 - Block Statistics
 - Filter
 - Focal Flow
 - Focal Statistics
 - Line Statistics
 - Point Statistics
 - Overlay
 - Weighted Overlay
 - Raster Creation
 - Reclass
 - Surface
 - Aspect
 - Contour
 - Contour List
 - Curvature
 - Cut/Fill
 - Hillshade
 - Observer Points
 - Slope
 - Viewshed
 - Zonal
 - Tabulate Area**
 - Zonal Fill
 - Zonal Geometry
 - Zonal Geometry as Table
 - Zonal Statistics
 - Zonal Statistics as Table

Creating Watersheds

- Fill Sinks in DEM
- Compute Flow Direction
- Compute Flow Accumulation
- Generate Watershed



Surface Analysis Functions

- **Line-of-sight**

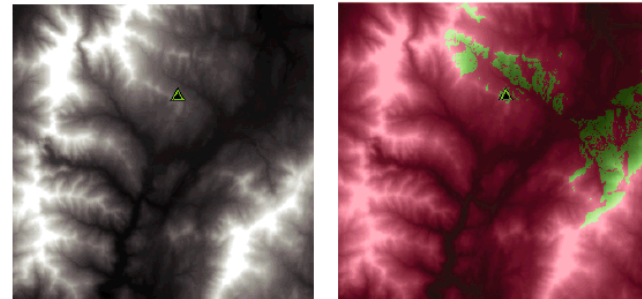
- Shows the intervisibility of landscape features
 - **Viewshed** – land area seen from a specified point
 - **Visual Impact** – all observation points from which a tall object can be seen
- Useful for radio tower siting, scenic area creation

- **Illumination**

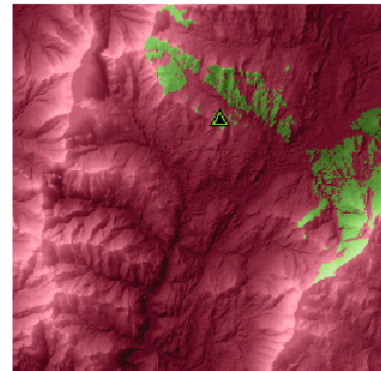
- Calculates the reflectance of a surface, given the slope and aspect of the data, and the sun angle

*Example: hillshade Ithaca
visualization in New
Zealand*

In the example below, the viewshed from an observation point is identified. The elevation grid displays the height of the land (darker locations represent lower elevations), and the observation point is marked as a green triangle. Cells in green are visible from the observation point, while cells in red are not visible.



Displaying a hillshade underneath your elevation and the output from the Viewshed function gives a very realistic impression of the landscape and clearly indicates the locations that an observer can see from the observation point.



Surface Analysis Functions

- Surface Water Flow (run-off)
 - Water (among other things) flows downhill
- Volume and Cut/Fill Determination
 - Surface area is measured along the slope of a surface, taking height into consideration.
 - The calculated area will always be greater than the area measured by simply using the 2D planimetric extent of a model (pyramid example)
 - Cut/Fill analysis determines how much material has been lost or gained in a study area by comparing two surface models of the area -- one before a change and one after

Surface Analysis Application

- LIRR Grade Crossing Accident
- Routing to Nearest Main

Watershed Delineation

- Watershed – the area of land that contributes precipitation runoff to a waterbody.
- A watershed is an area that drains water and other substances to a common outlet as concentrated drainage.
 - basin, catchment , or contributing area.
- the total area flowing to a given outlet, or pour point.
- The boundary between two watersheds is referred to as a watershed boundary or drainage divide.

Watershed Process

- Fill depressions in an elevation grid
- Calculate Flow Direction
- Calculate Flow Accumulation
- Calculate Flow Length
- Delineate Watershed

Flow Direction

- Calculates the direction of flow out of each cell into one of its eight neighbors (using the steepest drop)
- If a cell is lower than its 8 neighbors, that cell is given the value of its lowest neighbor and flow is defined towards this cell
- Downstream tracing can be done by using the flow direction Grid
- If the descent to all adjacent cells is the same, the neighborhood is enlarged until a steepest descent is found

78	72	69	71	58	49
74	67	56	49	46	50
69	53	44	37	38	48
64	58	55	22	31	24
68	61	47	21	16	19
74	53	34	12	11	12

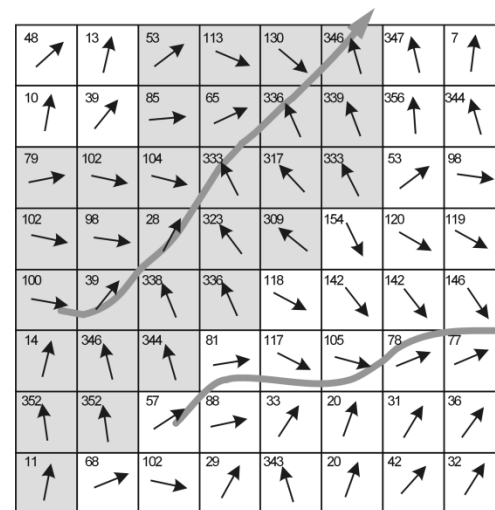
elevGrid

=

2	2	2	4	4	8
2	2	2	4	4	8
1	1	2	4	8	4
128	128	1	2	4	8
2	2	1	4	4	4
1	1	1	1	4	16

flowGrid

© ESRI, ArcView Help



© Paul Bolstad, GIS Fundamentals

Flow Accumulation

- Flow accumulation can be used to create a stream network
- The accumulated flow is based upon the number of cells flowing into each cell
- Output cells with a flow accumulation of 0 are local topographic highs and may be used to identify ridges
- Assumes that all rain became runoff and there was no interception, evapotranspiration, or loss to groundwater.

32 64 128
16 X 1
8 4 2

2	2	2	4	4	8
2	2	2	4	4	8
1	1	2	4	8	4
128	128	1	2	4	8
2	2	1	4	4	4
1	1	1	1	4	16

flowGrid

=

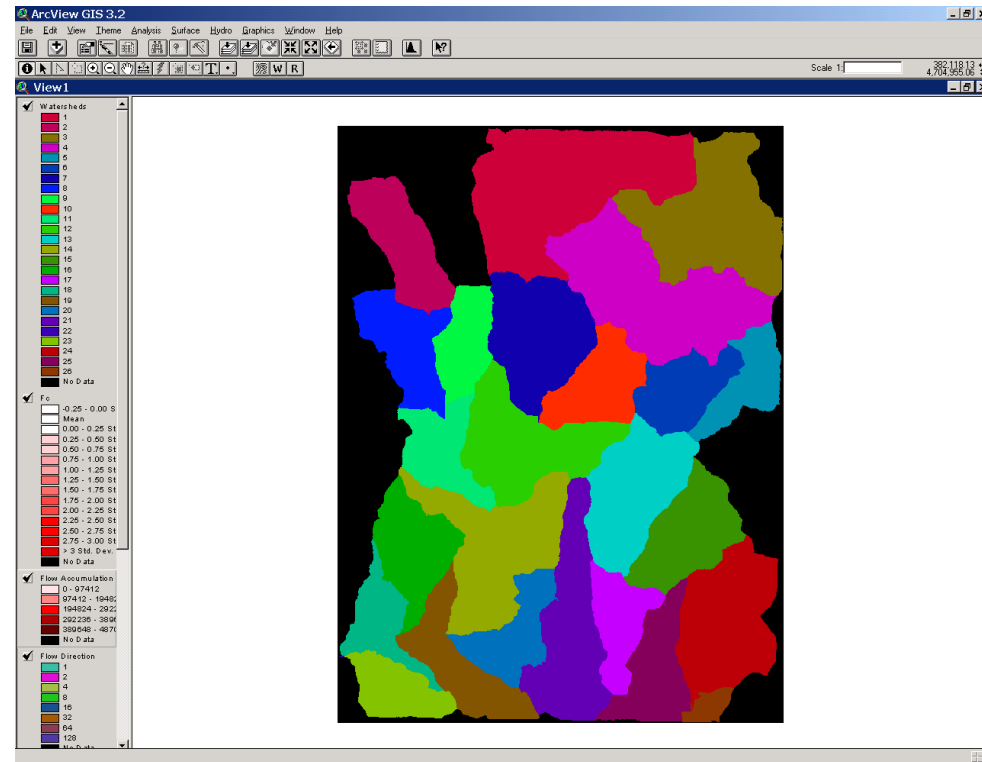
0	0	0	0	0	0
0	1	1	2	2	0
0	3	7	5	4	0
0	0	0	20	0	1
0	0	0	1	24	0
0	2	4	7	35	2

accumGrid

© ESRI, ArcView Help

Creating Watersheds

- Fill Sinks in DEM
- Compute Flow Direction
- Compute Flow Accumulation
- Generate Watershed
- Distance
- Distance with cost



Topographic Index - To determine likelihood of runoff

$$TI_i = \ln (a_i / \tan B_i)$$

Where, TI = Topographic Index,

i = land pixel

a = land pixel i 's upslope contributing area per contour length

B = land pixel i 's slope angle.

(Endreny & Wood 2003)

ArcGIS 10

```
from arcpy.sa import *
```

```
TI =
```

```
Ln(Divide(FlowAccumulation(FlowDirection("alt")),Slope("alt","PERCENT_RISE"  
)))
```

Manifold:

$$\frac{\log(\text{flowaccum}(\text{flowdir}([\text{Alt}])))}{(100 * (\tan(\text{Deg2Rad}(\text{slope}([\text{Alt}])))) / 100)}$$

Watershed:

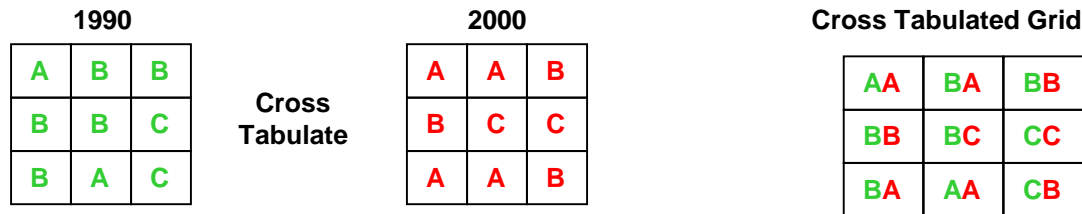
```
flowareas(flowdir([Alt]),flowstreams(flowdir([Alt]),f  
lowaccum(flowdir([Alt]),1000))
```

Comparing Two Maps

- Map comparison is easily facilitated using the **Tabulate Area** function in ArcGIS (or any decent raster based GIS)
 - Determines the cross tabulation between two grid themes on a cell-by-cell basis
- Once the tabulations are made, the data is displayed in a simple matrix
 - Map one is the X axis and Map two is the Y axis

Cross Tabulation

- Assume we have a 9 cell land cover map from 1990 with three categories: A, B, and C. We also have another map from 2000.



- You can see that the resulting cross tabulation provides a pixel by pixel comparison of the interpreted land cover types with the reference land cover. So, for the upper right hand cell, the the 1990 data was B, and the 2000 data was also B. Therefore, this is a match between the two data sets. However, in the lower right cell you can see that the 1990 data indicated a value of C and the 2000 data set had a value of B
- We can now quantify the results into a matrix as shown below.

	A	B	C
A	2	2	0
B	0	2	1
C	0	1	1

- The matrix on the right shows the comparison of two hypothetical data sets. The **1990** data set represents the the land use in 1990, while the **2000** data set represents the land use in 2000
- As an example, geographic features that were **A** in 2000, and were A in 1990, represent the upper left hand matrix with the value 2 (there were two pixels that met this criteria). This means that 2 units in the overall map that was **A** in 1990, was also A in 2000. Similarly, the same exists for **B** and **C**. The diagonal represents areas that have not changed.
- But, there may have been times where in 1990 the value was **A** but in 2000, the value was **B**. In this case, the **2** represented in the top row of the matrix says that there are **2** units of something we said was **A** in 1990 but is **B** in 2000
- We can begin to add these number up, by adding an additional row and column. But what do these numbers tell us?

2000 Data

	A	B	C
A	2	2	0
B	0	2	1
C	0	1	1

1990 Data

2000 Data

	A	B	C	
A	2	2	0	4
B	0	2	1	3
C	0	1	1	2
	2	5	2	5

1990 Data

- The bottom row tells us that there were two cells that were A in both 1990 and 2000, five cells that were B, and two cells that were C. The rightmost column tells us that there were four cells that were A in 2000, three cells as B, and 2 cells as C. Adding up the *Diagonal* cells says there were 5 cells where we actually got it right.

- So, the similarities is really a function of:

- Total cells on the diagonal / total number of cells.
 - $(2 + 2 + 1) / (2 + 2 + 0 + 0 + 2 + 1 + 0 + 1 + 1) = 5/9 = .55\%$ agreement
 - The change is 45%.

- We can also dig deeper and look at **A**. In 1990 there were 2 A's. But, by the time 2000 came around there were 4 A's. The A's doubled. But, at what cost:

- Well, **A** grew by replacing 2 **B**'s. Therefore, we can see that **B** lost some ground to **A**. Imagine a real world example of being able to say that as Developed land grew, it actually grew by replacing Forested land: *that's the power of evaluating a cross tabulation.*

1990

	A	B	C	
A	2	2	0	4
B	0	2	1	3
C	0	1	1	2
	2	5	2	5

2000