Geographic Information Systems 2018/19
Week 10, Topic 1

Mark Foley
School of Computing
DIT, Kevin Street

mark.foley@dit.ie

In this session...

Raster Data Analysis

- Data Analysis Environment
- Local Operations
- Neighbourhood Operations
- Zonal Operations
- Physical Distance Measure Operations
- Other Raster Data Operations
- Comparison of Vector- and Raster-Based Data Analysis

Raster Data Analysis

Raster data analysis is based on cells and rasters.

Raster data analysis can be performed at the level of individual cells, or groups of cells, or cells within an entire raster.

Some raster data operations use a single raster, others use two or more rasters.

Raster data analysis also depends on the type of cell value (numeric or categorical values).

Raster Analysis Environment

The analysis environment refers to the area for analysis and the output cell size.

Local Operations: Single Raster

Given a single raster as the input, a local operation computes each cell value in the output raster as a mathematical function of the cell value in the input raster.

Arithmetic	+, -, /, *, absolute,integer,floating-point
Logarithmic	exponentials, logarithms
Trigonometric	sin,cos, tan, arcsin, arccos, arctan
Power	square, square root, power

Arithmetic, logarithmic, trigonometric, and power functions for local operations.

Local Operations: Single Raster

15.2	16.0	18.5
17.8	18.3	19.6
18.0	19.1	20.2

(a)

8.64	9.09	10.48
10.09	10.37	11.09
10.20	10.81	11.42

(b)

A local operation can convert a slope raster from percent (a) to degrees (b).

Local Operations: Multiple Rasters

A common term for local operations with multiple input rasters is map algebra, a term that refers to algebraic operations with raster map layers.

Besides mathematical functions that can be used on individual rasters, other measures that are based on the cell values or their frequencies in the input rasters can also be derived and stored on the output raster of a local operation with multiple rasters.

Local Operations: Multiple Rasters

(a)

5	2	3
	2	2
3	1	1

(b)

1	3	2
4	7	5
1	1	

(c)

3	4	1
4	3	2
2	1	1

(d)

3	3	2
	4	3
2	1	

The cell value in (d) is the mean calculated from three input rasters (a, b, and c) in a local operation. The shaded cells have no data.

Local Operations: Multiple Rasters

(a)

3	2	1
2	1	2
1	2	3

(b)

3	2	4
3	2	4
2	4	1

(c)

1	3	6
2	4	5
4	5	7

(d)

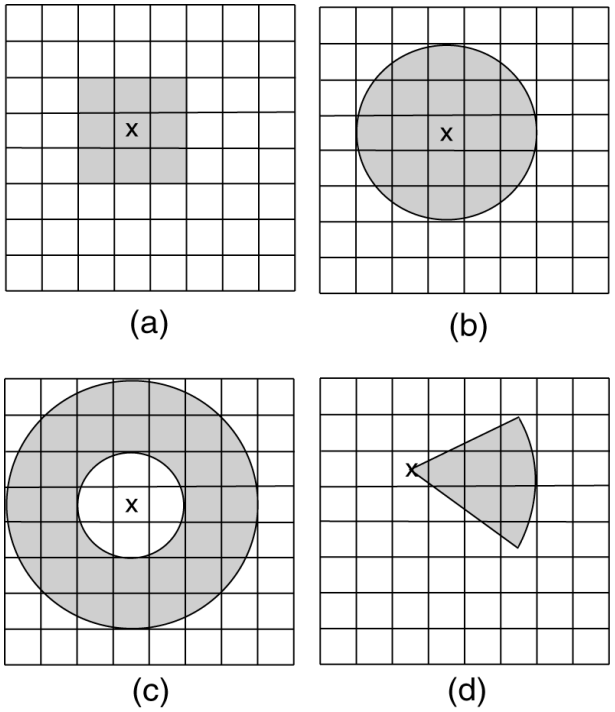
Combine code	1	2	3	4	5	6	7
(slope, aspect)	(3,3)	(2,3)	(2,2)	(1,2)	(2,4)	(1,4)	(3,1)

Each cell value in (c) represents a unique combination of cell values in (a) and (b). The combination codes and their representations are shown in (d).

Neighbourhood Operations

A **neighbourhood operation** involves a focal cell and a set of its surrounding cells. The surrounding cells are chosen for their distance and/or directional relationship to the focal cell.

Common neighborhoods include rectangles, circles, annuluses, and wedges.



Four common neighborhood types: rectangle (a), circle (b), annulus (c), and wedge (d). The cell marked with an x is the focal cell.

Neighbourhood Operations

(a)

1	2	2	2	2
1	2	2	2	3
1	2	1	3	3
2	2	2	3	3
2	2	2	2	3

(b)

1.56	2.00	2.22
1.67	2.11	2.44
1.67	2.11	2.44

The cell values in (b) are the neighborhood means of the shaded cells in (a) using a 3 x 3 neighborhood. For example, 1.56 in the output raster is calculated from (1 +2 +2 +1 +2 +2 +1 +2 +1) / 9.

Neighbourhood Operations

(a)

1	2	2	2	2
1	2	2	2	3
1	2	1	3	3
2	2	2	3	3
2	2	2	2	3

(b)

2	2	2
2	2	3
2	2	3

The cell values in (b) are the neighbourhood majority statistics of the shaded cells in (a) using a 3 x 3 neighborhood. For example, the upper-left cell in the output raster has a cell value of 2 because there are five 2s and four 1s in its neighborhood.

Zonal Operations

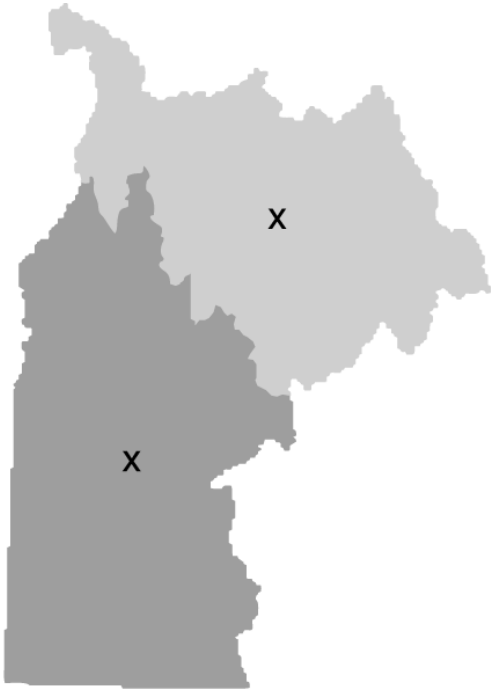
A zonal operation works with groups of cells of same values or like features. These groups are called zones. Zones may be contiguous or noncontiguous.

A zonal operation may work with a single raster or two rasters.

Given a single input raster, zonal operations measure the geometry of each zone in the raster, such as area, perimeter, thickness, and centroid.

Given two rasters in a zonal operation, one input raster and one zonal raster, a zonal operation produces an output raster, which summarizes the cell values in the input raster for each zone in the zonal raster.

Zonal Operations



Zone	Area	Perimeter	Thickness
1	36,224	1,708	77.6
2	48,268	1,464	77.4

Thickness and centroid for two large watersheds (zones). Area is measured in square kilometers, and perimeter and thickness are measured in kilometers. The centroid of each zone is marked with an x.

Zonal Operations

1	2	2	1
1	4	5	1
2	3	7	6
1	3	4	4

(a)

1	1	2	2
1	1	2	2
1	1	3	3
3	3	3	3

(b)

2.17	2.17	2.25	2.25
2.17	2.17	2.25	2.25
2.17	2.17	4.17	4.17
4.17	4.17	4.17	4.17

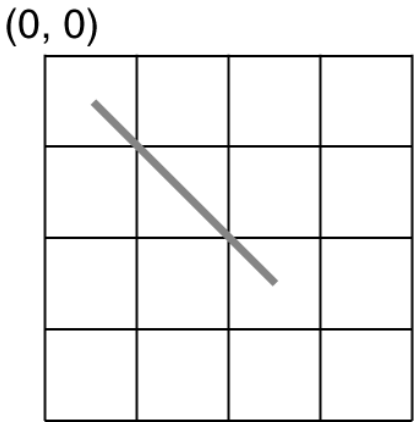
(c)

The cell values in (c) are the zonal means derived from an input raster (a) and a zonal raster (b). For example, 2.17 is the mean of {1, 1, 2, 2, 4, 3} for zone 1.

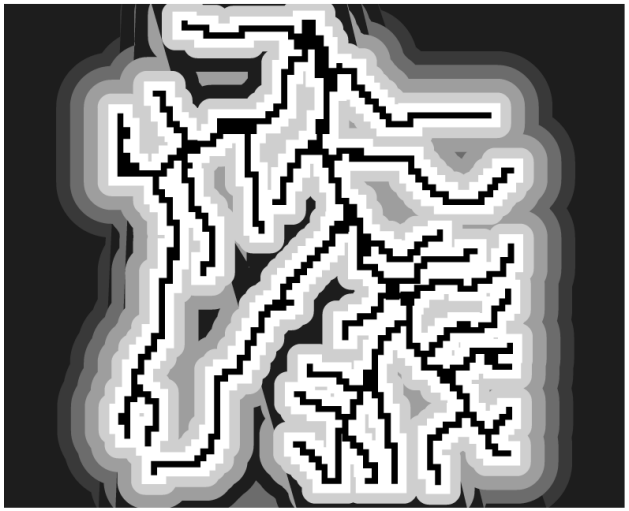
Physical Distance Measure Operations

The physical distance measures the straight-line or euclidean distance.

Physical distance measure operations calculate straight-line distances away from cells designated as the source cells.



A straight-line distance is measured from a cell center to another cell center. This illustration shows the straight-line distance between cell (1,1) and cell (3,3).

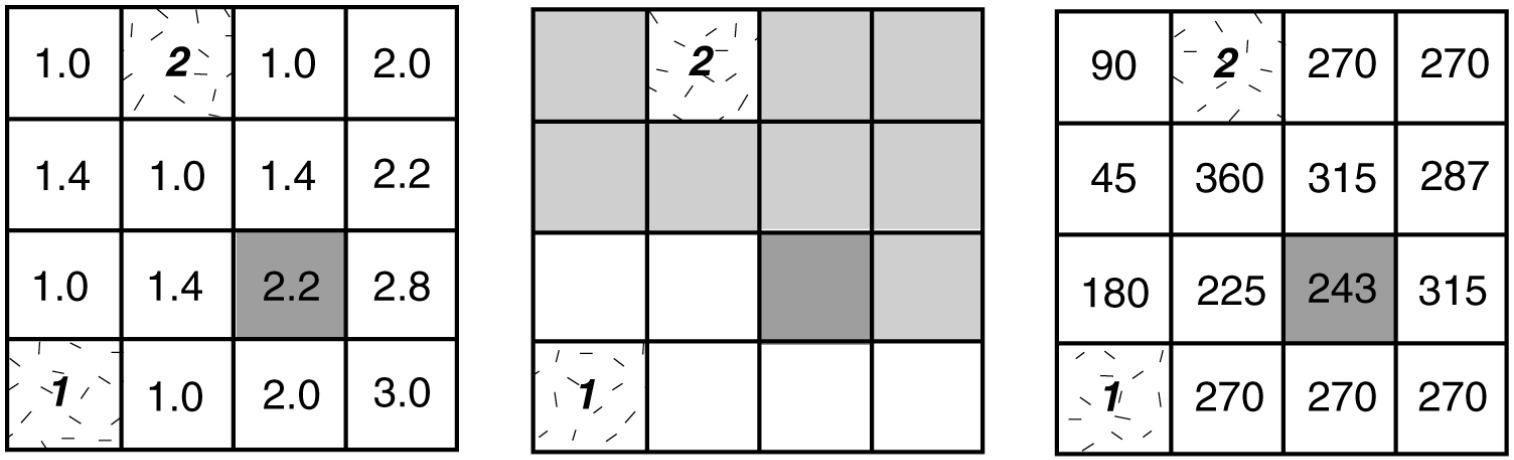


Continuous distance measures from a stream network.

Allocation and Direction

Allocation produces a raster in which the cell value corresponds to the closest source cell for the cell.

Direction produces a raster in which the cell value corresponds to the direction in degrees that the cell is from the closest source cell.



Based on the source cells denoted as 1 and 2, (a) shows the physical distance measures in cell units from each cell to the closest source cell; (b) shows the allocation of each cell to the closest source cell; and (c) shows the direction in degrees from each cell to the closest source cell. The cell in a dark shade (row 3, column 3) has the same distance to both source cells. Therefore, the cell can be allocated to either source cell. The direction of 243^0 is to the source cell 1.

Other Raster Data Operations

Operations for raster data **management** include Clip and Mosaic.

Operations for raster data **extraction** include use of a data set, a graphic object, or a query expression to create a new raster by extracting data from an existing raster.

Operations for raster data **generalization** include Aggregate and RegionGroup.

Clip

(a) (b) (c)

An analysis mask (b) is used to clip an input raster (a). The output raster is (c), which has the same area extent as the analysis mask.

Extract

(a) (b)

A circle, shown in white, is used to extract cell values from the input raster (a). The output (b) has the same area extent as the input raster but has no data outside the circular area.

Aggregate

1	3	2	5
1	3	2	7
1	1	2	5
2	4	3	2

(a)

2	4
2	3

(b)

An Aggregate operation creates a lower-resolution raster (b) from the input (a). The operation uses the mean statistic and a factor of 2 (i.e., a cell in b covers 2 x2 cells in a). For example, the cell value of 4 in (b) is the mean of {2, 2, 5, 7} in (a).

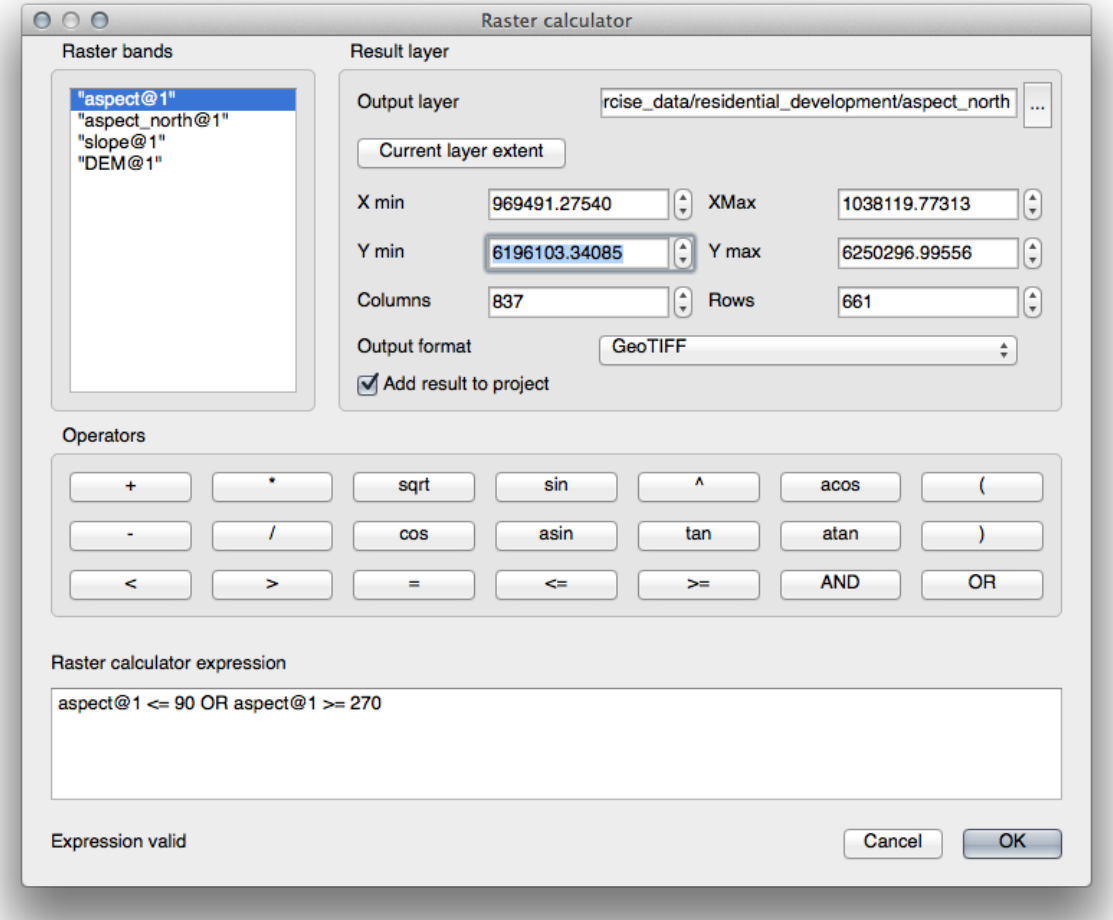
Map Algebra

Informal language with syntax similar to algebra, which can be used to facilitate manipulation and analysis of raster data.

Uses expression to link input to output.

Example from *QGIS* Raster Calculator

aspect@1 <= 90 OR aspect@1 >= 270



Coming next...

Terrain Mapping and Analysis