Lecture 4:
Distance, Buffers, and Map Algebra

Paul A. Raschky,

University of St Gallen, October 2017

Today's Lecture

1. Distance
2. Buffer
3. Map Algebra

1. Distance

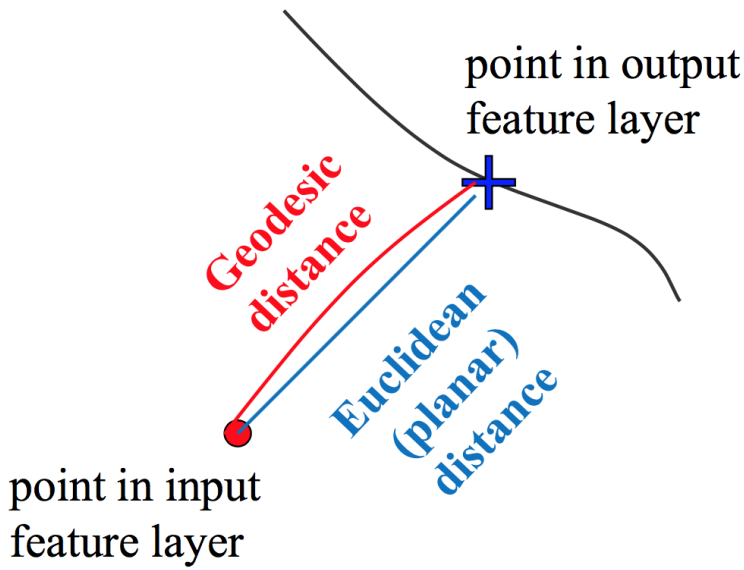
Measuring Distance in GIS

- Computes the distance from each point in the input feature layer to the nearest point, line or polygon in the near feature layer (within the maximum search radius).
- **Geodesic**, great-circle distance or orthodromic distance is the shortest distance between two points on the surface of a sphere.
- **Euclidean** distance is the straight-line distance between two points in Euclidean space.

1. Distance

Measuring Distance in GIS

- On a global / continental / national scale, distance measures based on an Euclidean geometry will not be correct.
- Map projections distort distance measures



1.1. Point to Point - Distance

Hjort & Poulson (2017) "The Arrival of Fast Internet and Employment in Africa"

- Estimate the effect of fast internet on employment outcomes in Africa.
- Identification Strategy:
  - Two Groups Individuals and firms in locations in Africa that are on the terrestrial network of Internet cables to those that are not.
  - Compare these two groups during the gradual arrival in coastal cities of submarine cables from Europe that greatly increase speed and capacity on the terrestrial network.

1.1. Point to Point - Distance

Hjort & Poulson (2017)

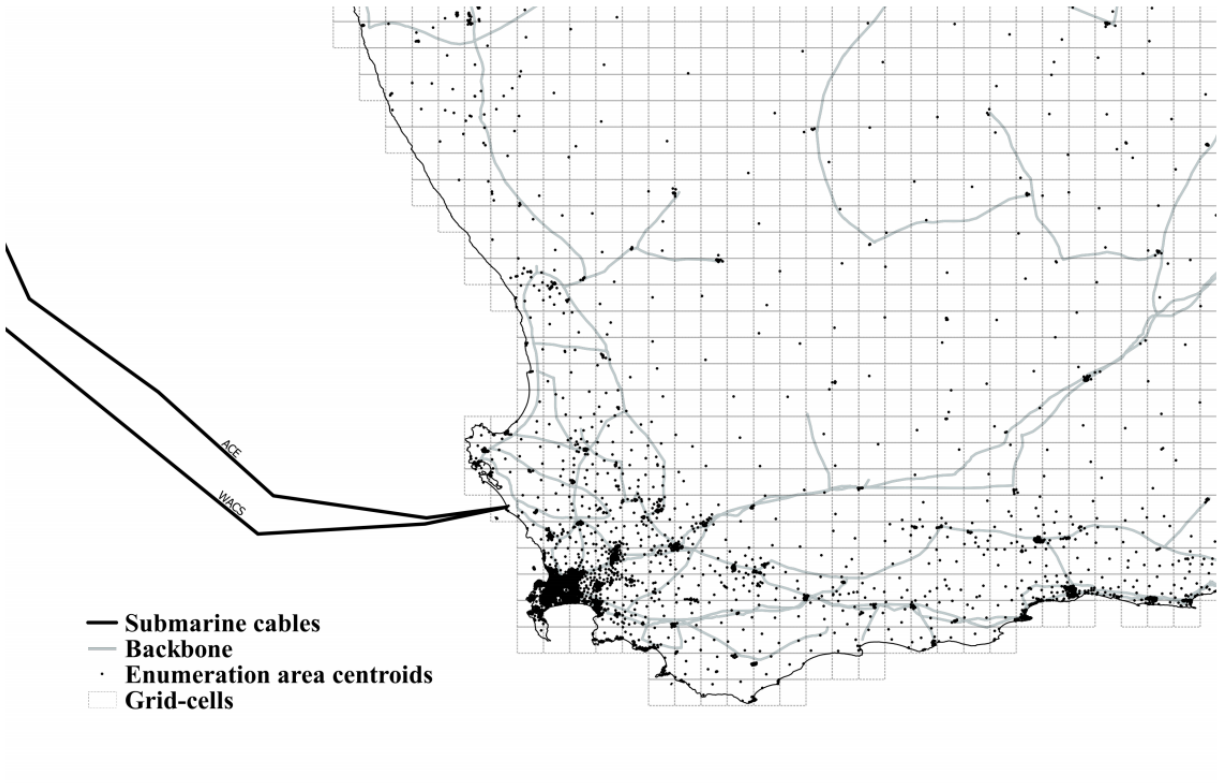


Image sources: Hjort & Poulson (2017)

1.1. Point to Point - Distance

Hjort & Poulson (2017) - Empirical Strategy

- Difference-in-Differences Model

$$y_{ij(i)t} = \alpha + \beta \text{SubmarineCables}_{it} \times \text{Connected}_i + \delta_{j(i)} + \eta_t + \epsilon_{ij(i)t}$$

1.1. Point to Point - Distance

Hjort & Poulson (2017) - Results 1

Outcome:	**Internet speed (asinh)**					**Internet use** **Daily (0/1)**			**Weekly (0/1)**		
Unit of analysis:	Location					Individual					
Sample:	**Akamai**					**Afrobarometer**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
SubmarineCables × Connected	0.245**	0.252**	0.409***	0.366**	0.452**	0.054**	0.068**	0.103*	0.093***	0.103***	0.104**
	(0.105)	(0.126)	(0.149)	(0.167)	(0.206)	(0.025)	(0.028)	(0.052)	(0.029)	(0.032)	(0.049)
Observations	8562	7293	2373	2373	2373	4928	4928	4928	4928	4928	4928
Mean of Outcome						0.087	0.087	0.087	0.172	0.172	0.172
Time FE	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# unique IP's > 10	No	No	Yes	Yes	Yes	No	No	No	No	No	No
Country×Time FE	No	No	No	Yes	No	No	Yes	No	No	Yes	No
Connected×Time FE	No	No	No	No	Yes	No	No	Yes	No	No	Yes
Including biggest cities	Yes	No	No	No	No	No	No	No	No	No	No

1.1. Point to Point - Distance

Hjort & Poulson (2017) - Results 2

Panel A: Employment			
Outcome:	**Employment (0/1)**		
Unit of analysis:	Individual		
Sample:	**DHS**	**Afro-barometer**	**SA-QLFS**
	(1)	(2)	(3)
SubmarineCables × Connected	0.031***	0.058***	0.030***
	(0.010)	(0.021)	(0.010)
Observations	186434	13176	322944
Mean of Outcome	0.697	0.580	0.711
Time FE	Yes	Yes	Yes
Grid-cell FE	Yes	Yes	Yes

Panel B: Work-related outcomes from SA-QLFS				
Outcome:	**Hours worked (asinh)**	**Wants to work more (0/1)**	**Formal employment (0/1)**	**Informal employment (0/1)**
Unit of analysis:	Individual			
	(1)	(2)	(3)	(4)
SubmarineCables × Connected	0.141***	-0.022***	0.040***	0.001
	(0.043)	(0.008)	(0.012)	(0.005)
Observations	321556	547476	322944	322944
Mean of Outcome		0.666	0.476	0.121
Time FE	Yes	Yes	Yes	Yes
Grid-cell FE	Yes	Yes	Yes	Yes

1.1. Point to Point - Distance

Exercise 4a - Point to nearest point

1.1. Point to Point - Distance

Exercise 4b - DIY Exercise Distance Matrix

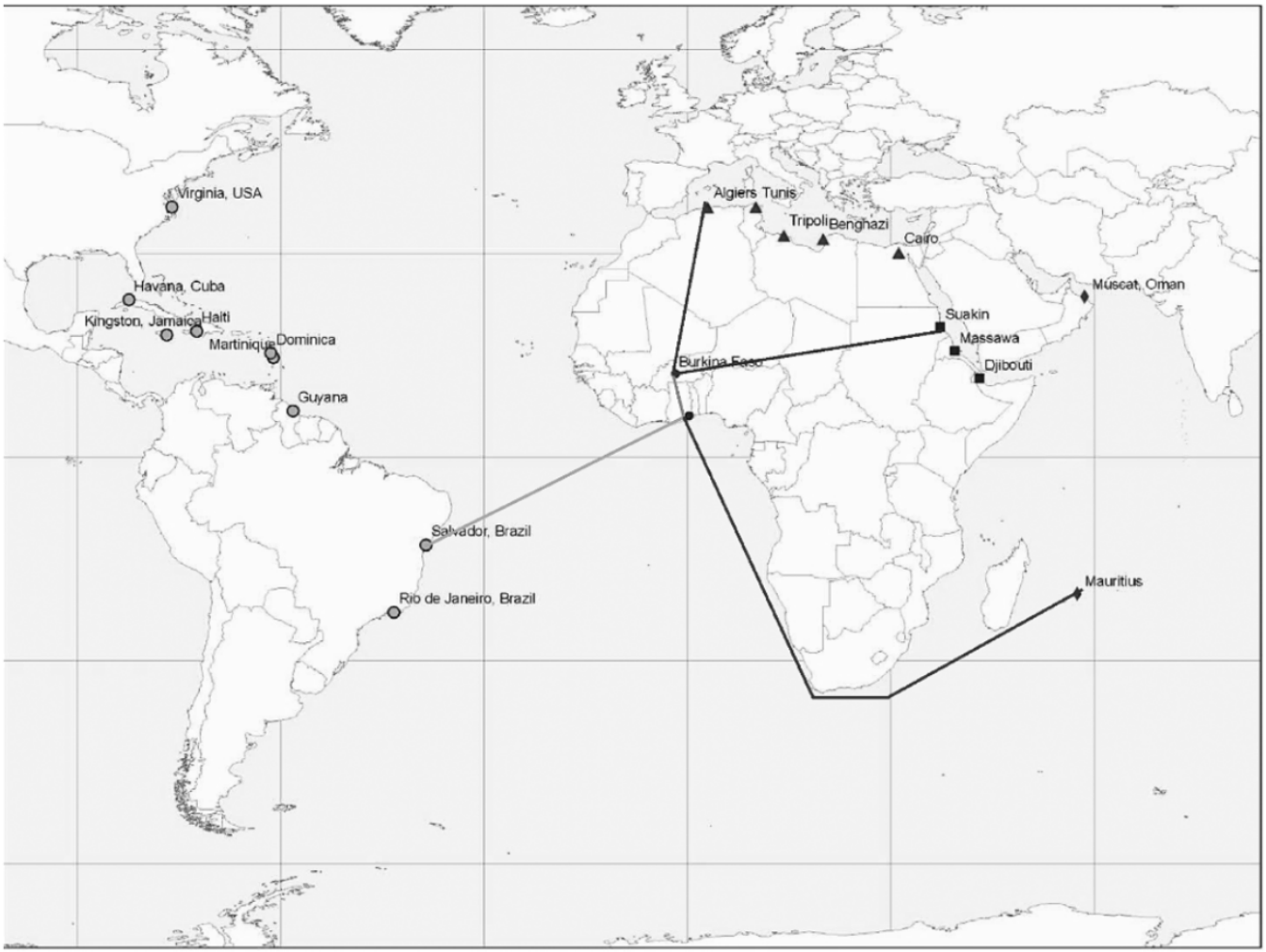
1.2. Point to Line - Distance

Nunn (2008) "The Long-term Effects of Africa's Slave Trades"

- Estimate the effect of slave trade on contemporary African underdevelopment.
- Identification Strategy:
  - Number of slaves exported from each port in Africa.
  - IV: Distance to the to nearest slave trade centers

1.2. Point to Line - Distance

Nunn (2008)



1.2. Point to Line - Distance

Nunn (2008) - Empirical Strategy

- Second Stage

$$y_i = \alpha + \beta \ln\left(\frac{exports_i}{area_i}\right) + \boldsymbol{X}_i'\boldsymbol{\gamma} + \varepsilon_i$$

1.2. Point to Line - Distance

Nunn (2008) - Empirical Strategy

- First Stage

$$\ln\left(\frac{exports_i}{area_i}\right) = \delta + \boldsymbol{D}_i'\boldsymbol{\Omega} + \boldsymbol{X}_i'\boldsymbol{\eta} + \mu_i$$

1.2. Point to Line - Distance

Nunn (2008) - Results 1

First Stage. Dependent variable is slave exports, ln(exports/area)

Atlantic distance	−1.31***	−1.74***	−1.32*	−1.69**
	(0.357)	(0.425)	(0.761)	(0.680)
Indian distance	−1.10***	−1.43***	−1.08	−1.57*
	(0.380)	(0.531)	(0.697)	(0.801)
Saharan distance	−2.43***	−3.00***	−1.14	−4.08**
	(0.823)	(1.05)	(1.59)	(1.55)
Red Sea distance	−0.002	−0.152	−1.22	2.13
	(0.710)	(0.813)	(1.82)	(2.40)
F-stat	4.55	2.38	1.82	4.01
Colonizer fixed effects	No	Yes	Yes	Yes
Geography controls	No	No	Yes	Yes
Restricted sample	No	No	No	Yes
Hausman test (p-value)	.02	.01	.02	.04
Sargan test (p-value)	.18	.30	.65	.51

1.2. Point to Line - Distance

Nunn (2008) - Results 2

Second Stage. Dependent variable is log income in 2000, $\ln y$				
ln(exports/area)	-0.208^{***}	-0.201^{***}	-0.286^{*}	-0.248^{***}
	(0.053)	(0.047)	(0.153)	(0.071)
	$[-0.51, -0.14]$	$[-0.42, -0.13]$	$[-\infty, +\infty]$	$[-0.62, -0.12]$
Colonizer fixed effects	No	Yes	Yes	Yes
Geography controls	No	No	Yes	Yes
Restricted sample	No	No	No	Yes
F-stat	15.4	4.32	1.73	2.17
Number of obs.	52	52	52	42

1.2. Point to Line - Distance

Exercise 4c - Point to nearest polyline

2. Buffer

Create a polygon of an input feature's neighborhood

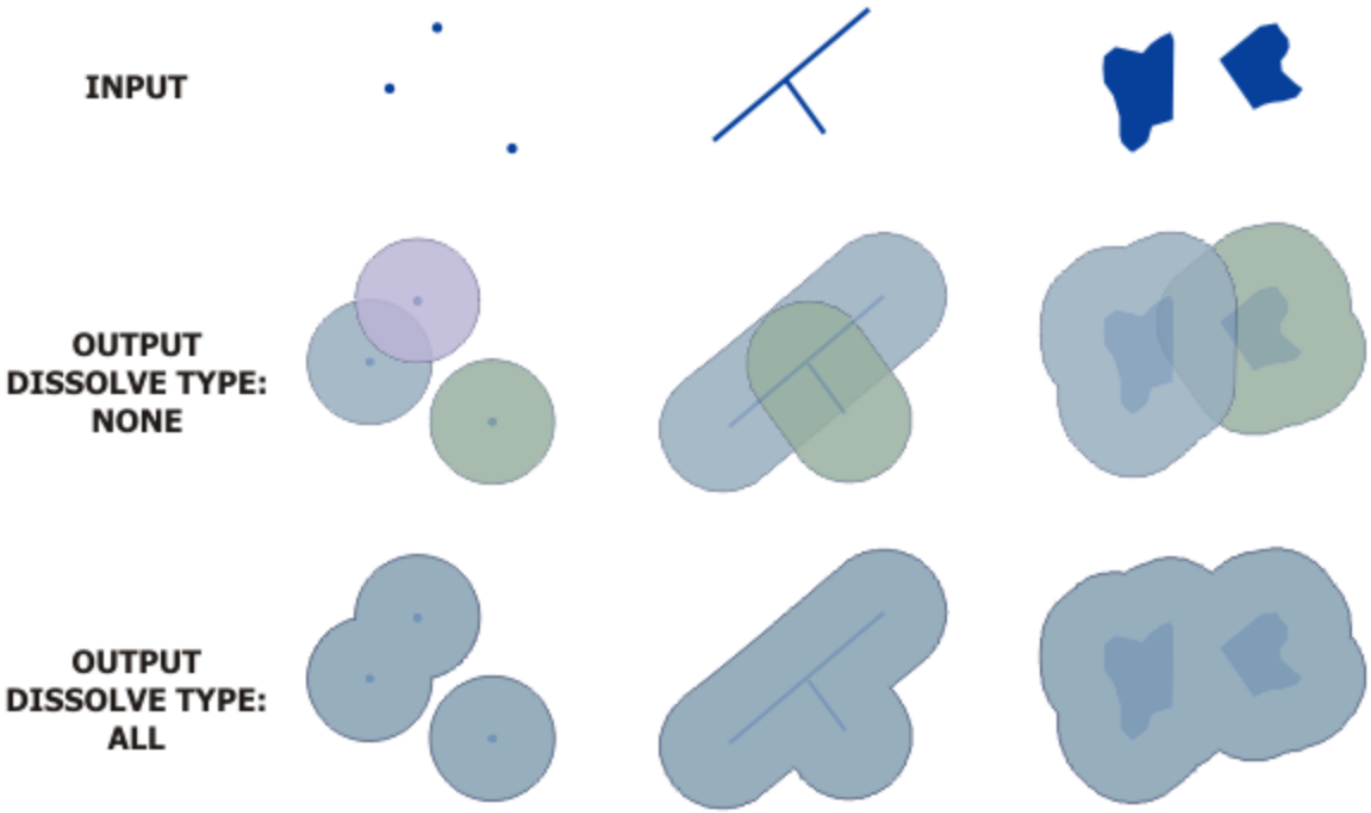


Image sources: desktop.arcgis.co

2. Buffer

Miguel & Kremer (2009) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities"

- Estimate the effect of slave trade on contemporary African underdevelopment.
- RCT evaluating school-based mass treatment with deworming drugs were randomly phased into schools
- $\rightarrow$ estimation of overall program effects.
  - Need to identify the surrounding area of treated schools

2. Buffer

Miguel & Kremer (2009) - Empirical Strategy

$$Y_{ijt} = a + \beta_1 \cdot T_{1it} + b_1 \cdot D_{1ij} + b_2 \cdot (T_{1it} * D_{1ij}) + X'_{ijt}\delta + \sum_d (\gamma_d \cdot N^T_{dit}) + \sum_d (\phi_d \cdot N_{dit}) + u_i + e_{ijt}.$$

2. Buffer

Exercise 4d - Buffer around schools

3. Map Algebra

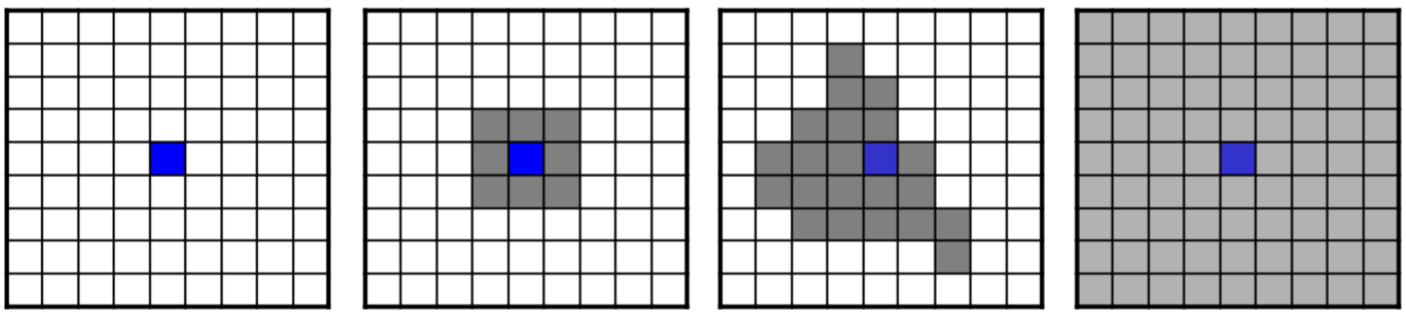
Cell-by-cell calculation across multiple raster datasets

- Arithmetic operations
  - with numbers (e.g., raster*2)
  - across several raster datasets (e.g., ras1 + ras2)
- Functions

3. Map Algebra

Functions

- Tomlin(1990) defines and organizes operations as local, focal, zonal, and global according to the spatial scope of the operations



3. Map Algebra

Local Functions - Arithmetic

438	450	468
455	481	473
476	498	502

*** 0.3048**

133,5	137,2	142,6
138.7	146.6	144.2
145.1	151.8	153.0

3. Map Algebra

Local Functions - Arithmetic

5	4	1
2	1	2
4	2	1

+

3	2	1
1	4	5
2	7	3

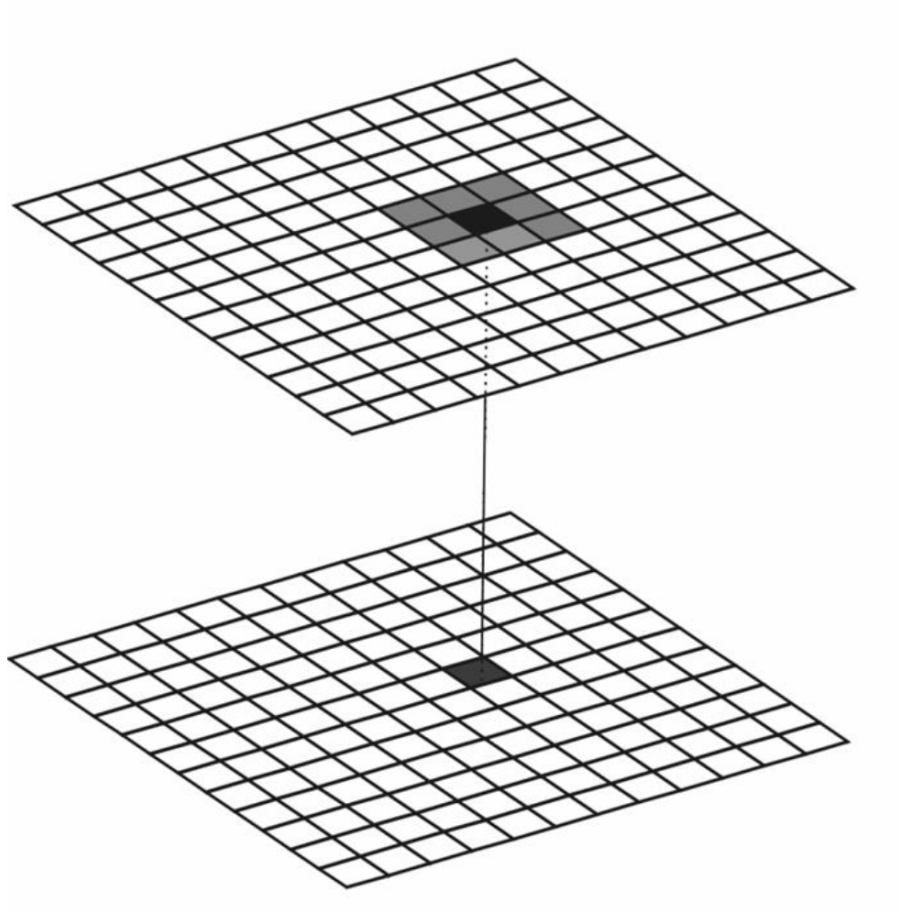
=

8	6	2
3	5	7
6	9	4

3. Map Algebra

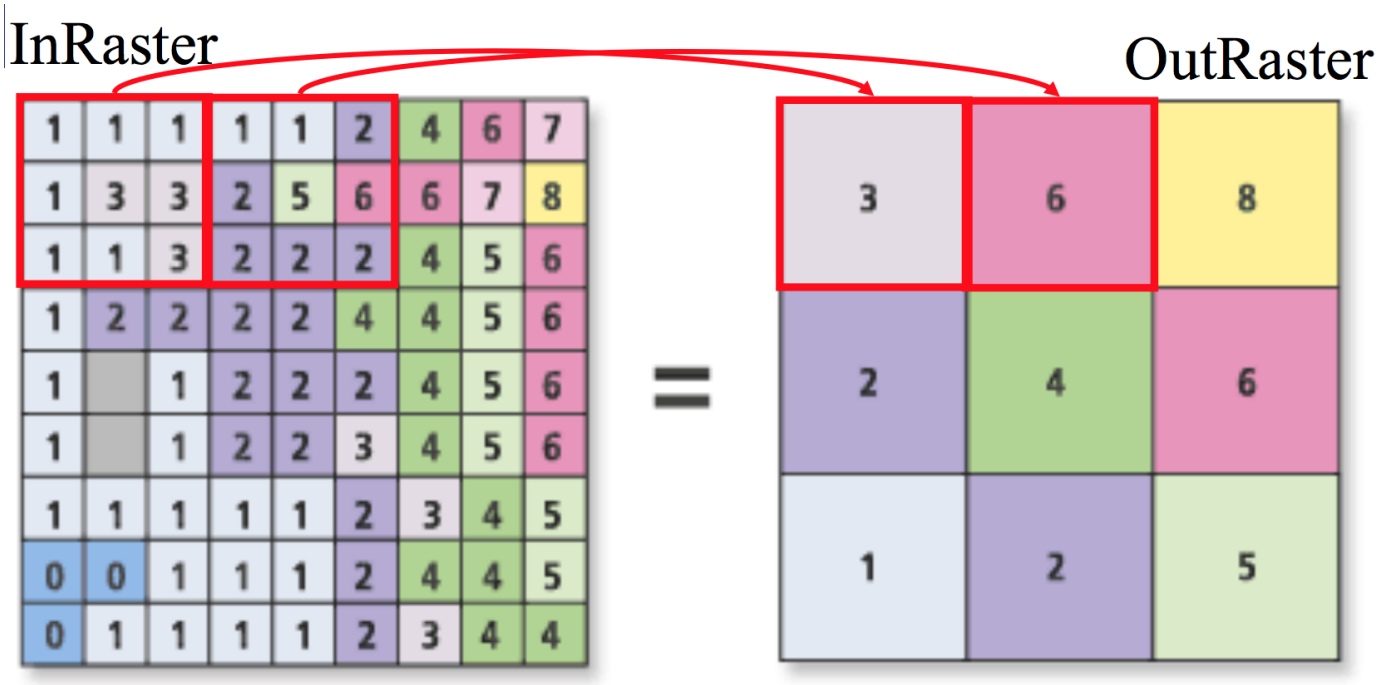
Focal Functions:

- Focal functions process cell data depending on the values of neighboring cells.
- We define a kernel, window, or rover to use as the neighbourhood (3x3, 5x5 cells).



3. Map Algebra

Focal Functions - Spatial Aggregation:



3. Map Algebra

Focal Functions - Application in Economics:

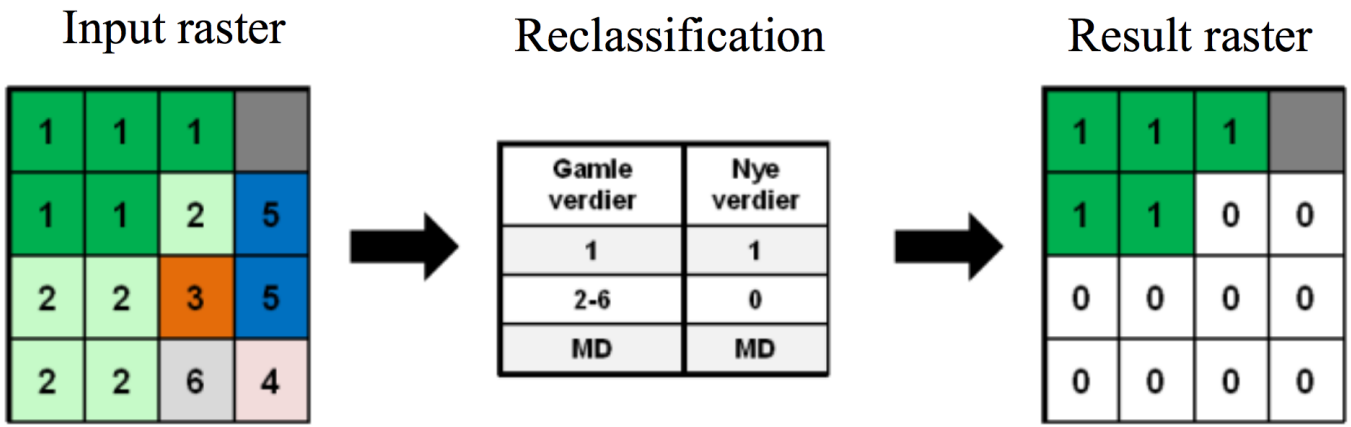
- Nunn & Puga (2011) "Ruggedness: The Blessing of Bad Geography in Africa"
  - Ruggedness $\rightarrow$ History $\rightarrow$ Current Development
  - Negative effects on transport costs and trade.
  - Protection to those being raided during the slave trades.

3. Map Algebra

Exercise 4e - Calculating Ruggedness

3. Map Algebra

Reclassify



3. Map Algebra

Reclassify

Old values	New values
1	1
2	1
3	0
4	0
5	0
6	0
NoData	NoData

3. Map Algebra

Reclassify

- In ArcGIS: Spatial Analyst Tools — Reclass — Reclassify
- In QGIS: r.reclass