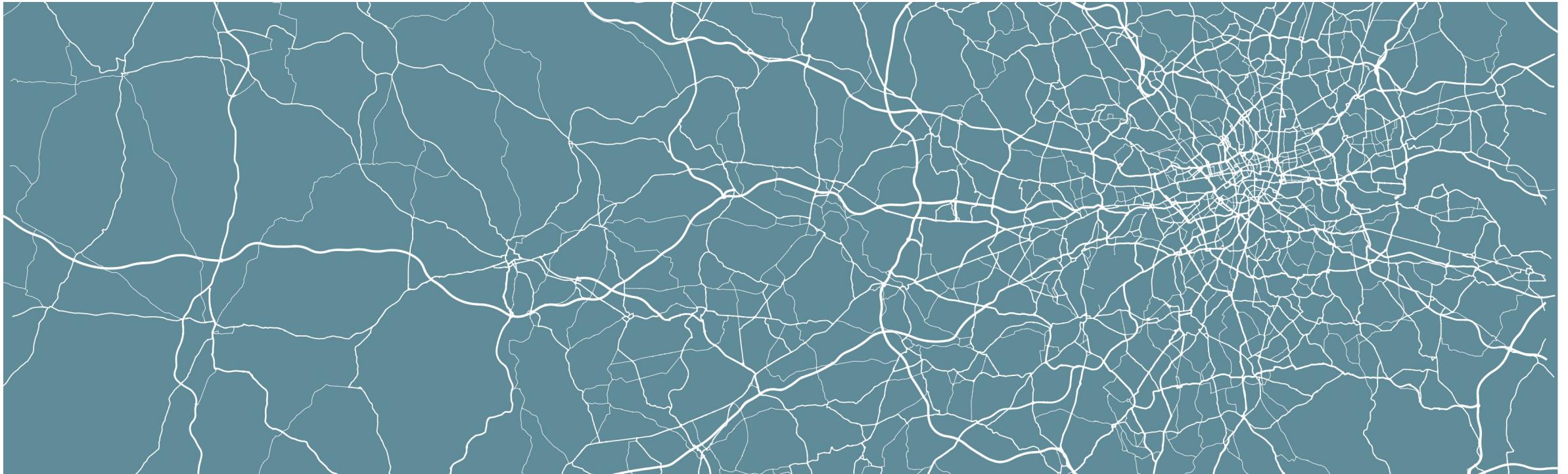


Geocomputation

W9 – Rasters, Zonal Statistics and Interpolation



Where are we at?

Part I: Foundational Concepts

W1 Geocomputation: An Introduction

W2 GIScience and GIS software

W3 Cartography and Visualisation



QGIS

W4 Programming for Data Analysis

W5 Programming for Spatial Analysis



R

Where are we at?

Part II: Core Spatial Analysis

W6 Geometric Operations and Spatial Queries

W7 Spatial Autocorrelation

W8 Point pattern analysis



R

Part III: Advanced Spatial Analysis

W9 **Rasters, Zonal Statistics and Interpolation**

W10 Transport Network Analysis



R

This week

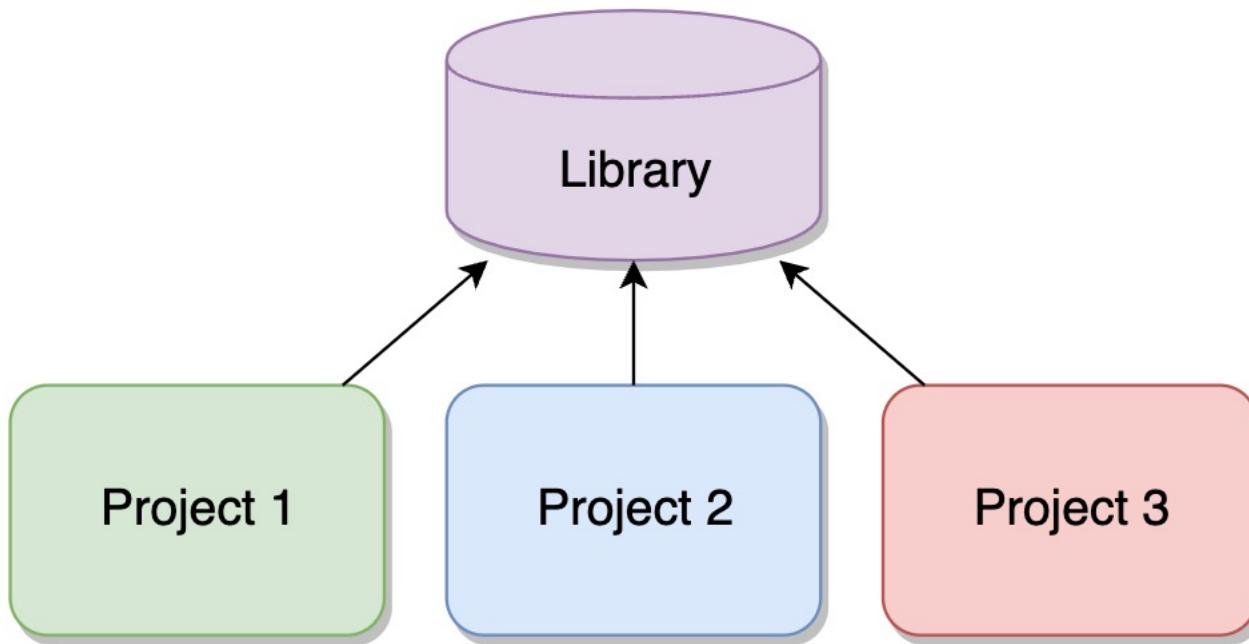
- A brief note on package management.
- Recap on vector versus raster GIS data model.
- Raster-specific functions and operations.
- Data interpolation and spatial data interpolation.

Package management

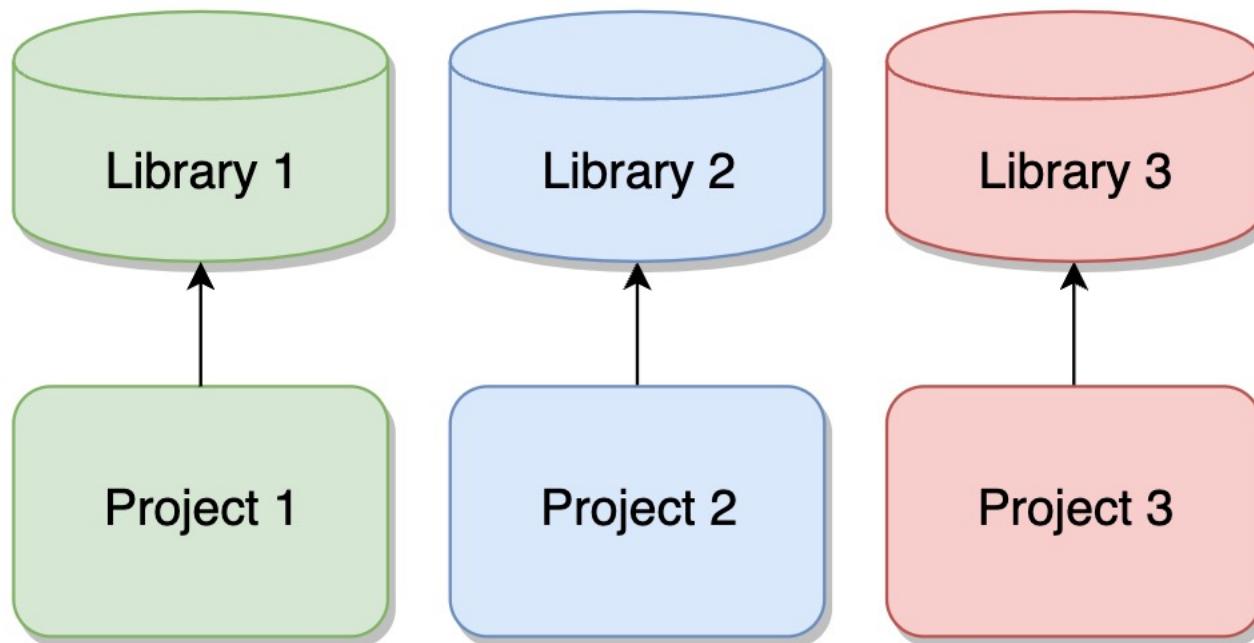
- Package management is the process of handling the many and varied dependencies and artifacts for your servers, applications, and developers.
- Toolkit used to manage project-local libraries.
- Combination with version management through Git.



Package management



Package management



GIS data models

- GIScience requires spatial information to be represented in a digital format
- Traditionally, geographic information is represented in two ways:

vector: a finite set of discrete geometric objects

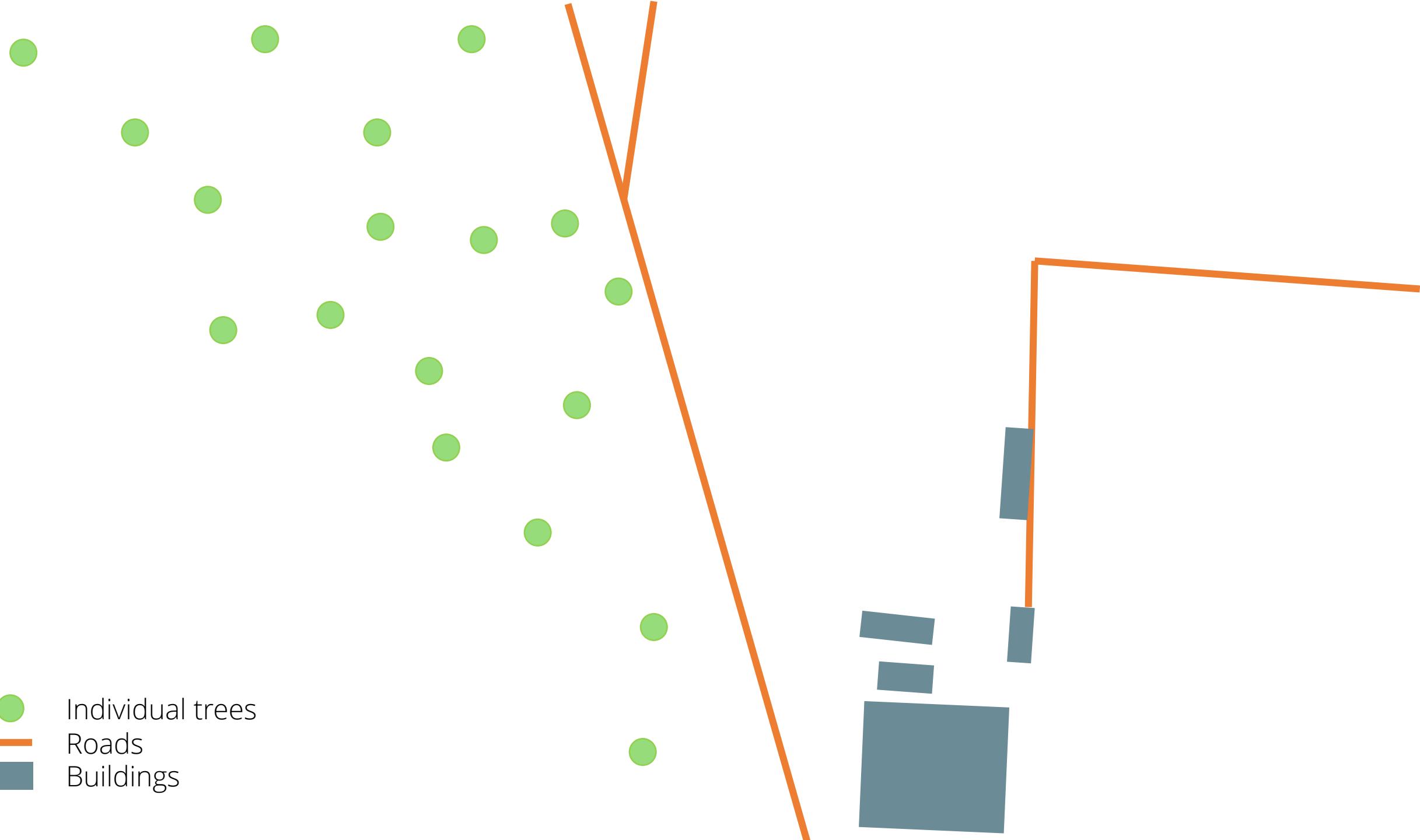
raster: images representing a surface (values, colours)











- Individual trees
- Roads
- Buildings





8	9	9	10	0	10	10	10	0	0	0	7	5	3	0	0	0	0	1
8	9	9	10	10	0	10	9	9	0	0	5	3	0	0	0	0	0	0
8	8	9	9	10	0	0	9	8	7	5	0	0	0	1	0	0	0	0
5	8	8	9	10	10	0	9	7	5	0	0	5	5	5	0	0	0	1
3	5	8	9	9	10	0	0	3	0	0	0	5	0	0	1	0	0	2
2	5	8	8	9	9	10	0	0	0	1	5	0	0	0	0	0	0	1
2	4	6	8	8	9	0	0	0	1	5	0	0	5	5	5	0	0	1
0	3	6	8	8	0	0	0	0	5	0	5	5	5	5	5	0	0	0
2	2	5	8	0	0	0	0	0	0	5	5	5	5	5	5	3	0	0
0	2	5	0	0	1	2	3	4	4	4	4	4	4	4	5	0	0	0
0	0	0	0	1	1	1	1	4	4	4	4	4	4	4	5	0	0	0
0	0	1	1	2	2	2	2	3	3	3	3	3	3	3	4	0	3	0
1	1	1	1	2	2	3	3	3	3	1	1	1	1	1	2	3	4	3

Raster properties

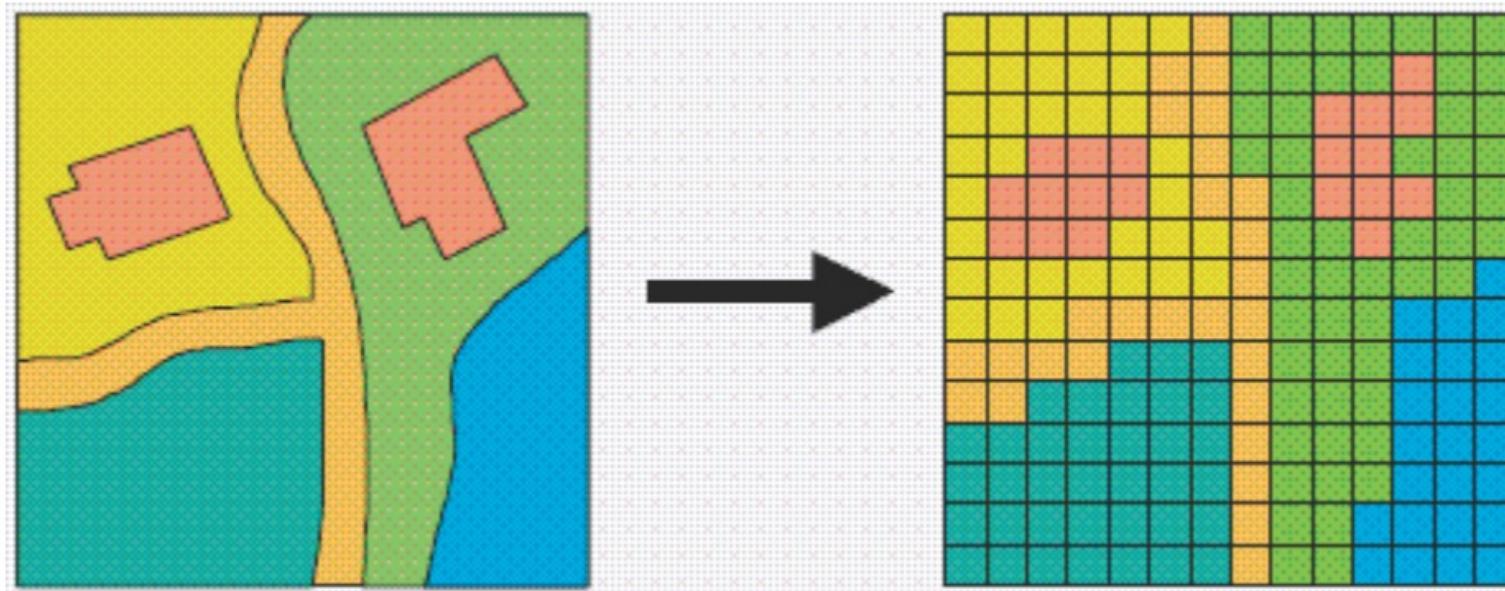
- Raster data are commonly found in applications where every location has one value.
- Variable can be continuous (temperature, elevation, air pollution), categorical (land use), or imagery (reflectance values).
- Cell size is constant across the raster surface.
- When do we encounter raster data:
 - remote sensing
 - rasterised vector map (tiling)
 - outcome of spatial analysis

Raster data



Continuous (numerical value of data)

Raster data



Categorical (number represent category)

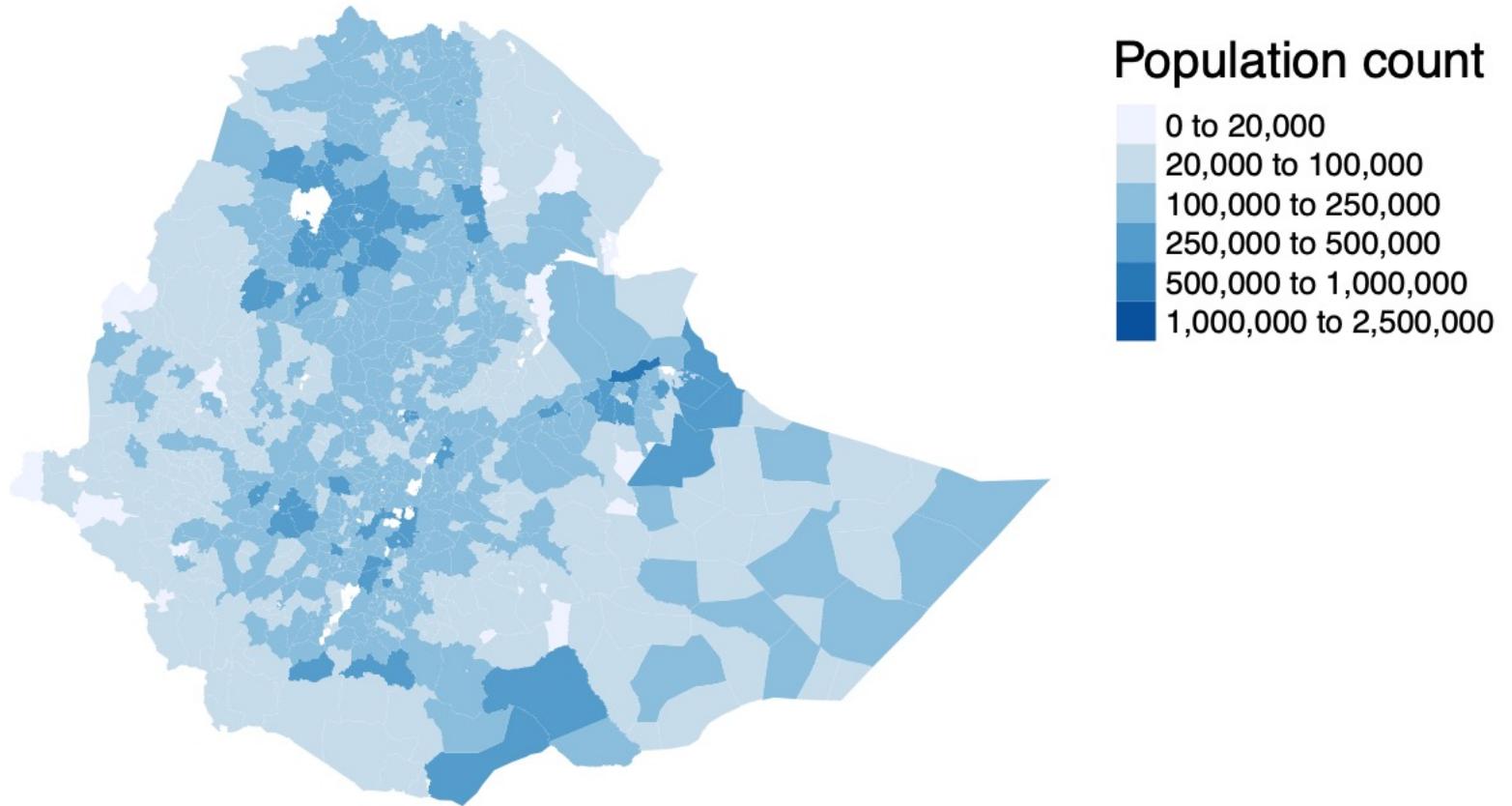
Raster data

The screenshot shows the homepage of the WorldPop website, which focuses on open spatial demographic data and research. The page features a navigation bar with links to About, Methods & Tools, News, Data, and Contact. The main title "Open Spatial Demographic Data and Research" is prominently displayed in large blue text. Below the title, a subtitle states: "WorldPop develops peer-reviewed research and methods for the construction of open and high-resolution geospatial data on population distributions, demographic and dynamics, with a focus on low and middle income countries." The central part of the page is a grid of six images illustrating different research areas:

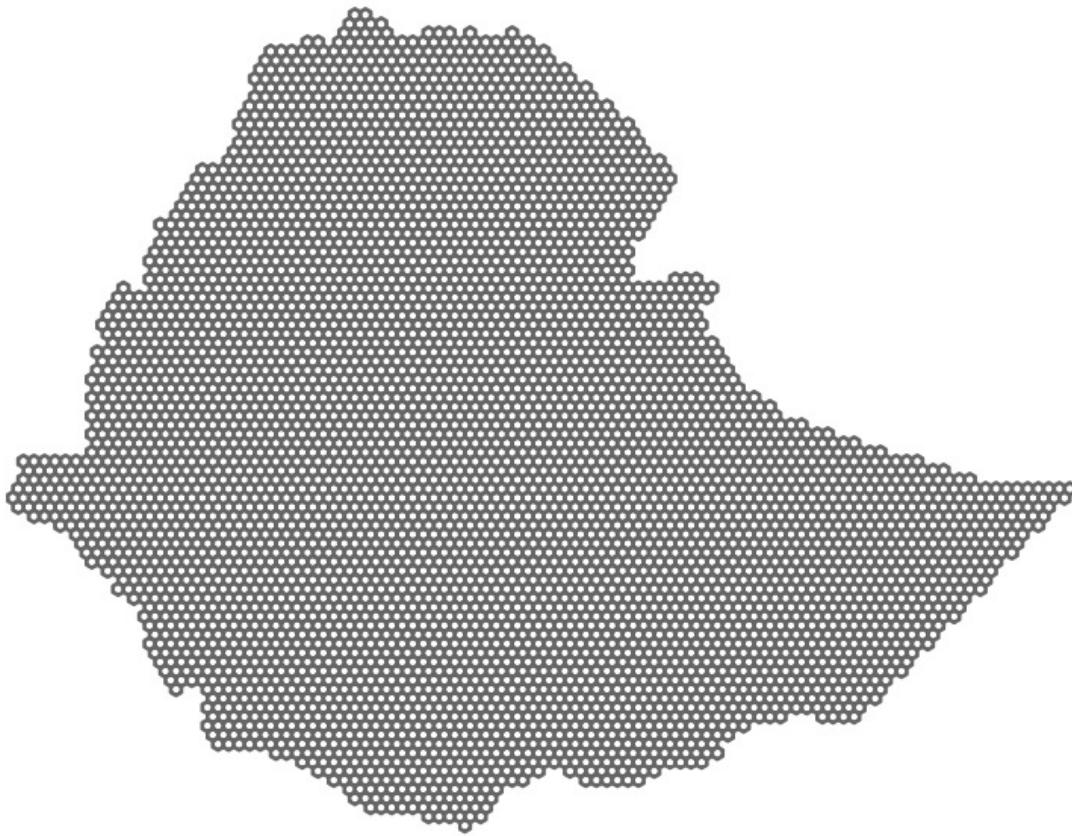
- WorldPop COVID-19 research:** Shows a purple-toned map of COVID-19 virus particles.
- Mapping populations:** Shows a map with green and blue density patterns.
- Spatial demographics:** Shows a map with a red and green heatmap overlay.
- Mapping development indicators:** Shows a map with blue and yellow shaded regions.
- Maternal and child health:** Shows a map with purple and pink shaded regions.
- Population dynamics:** Shows a map with blue lines representing movement patterns between cities like Niamey and Sokoto.

At the bottom of the page, there are two buttons: "What is WorldPop?" and "WorldPop Country Datasets".

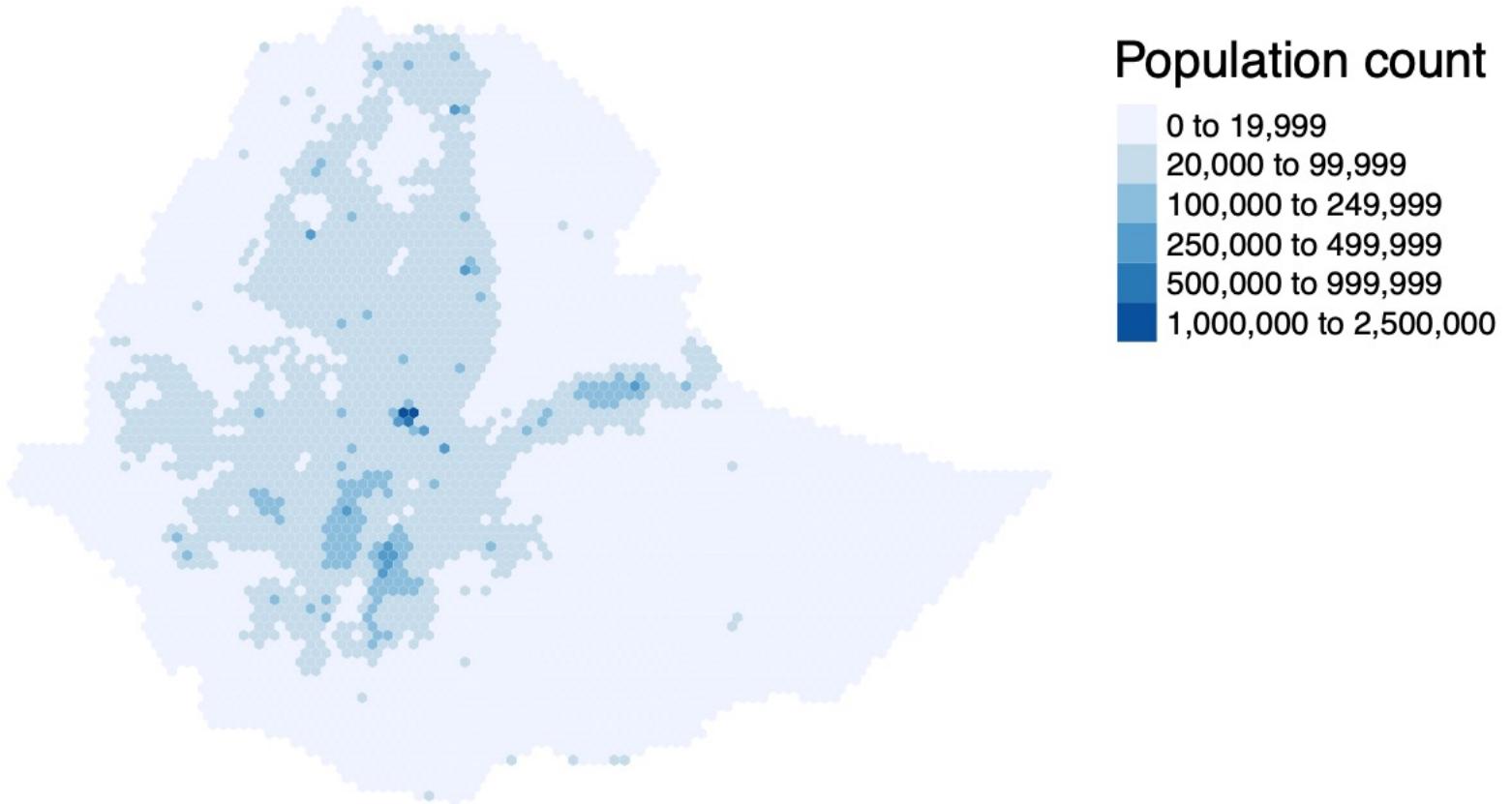
Raster data



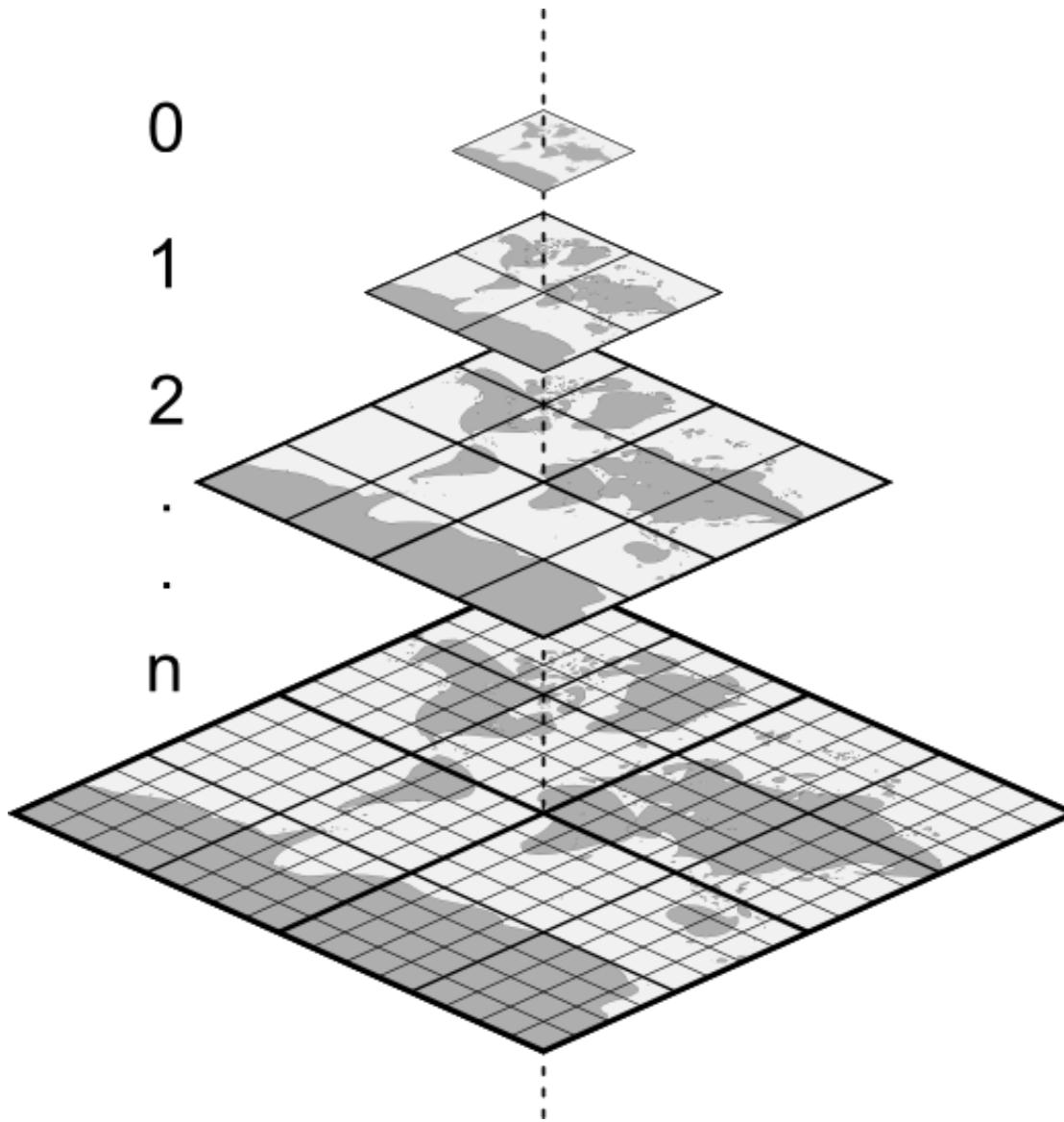
Raster data



Raster data

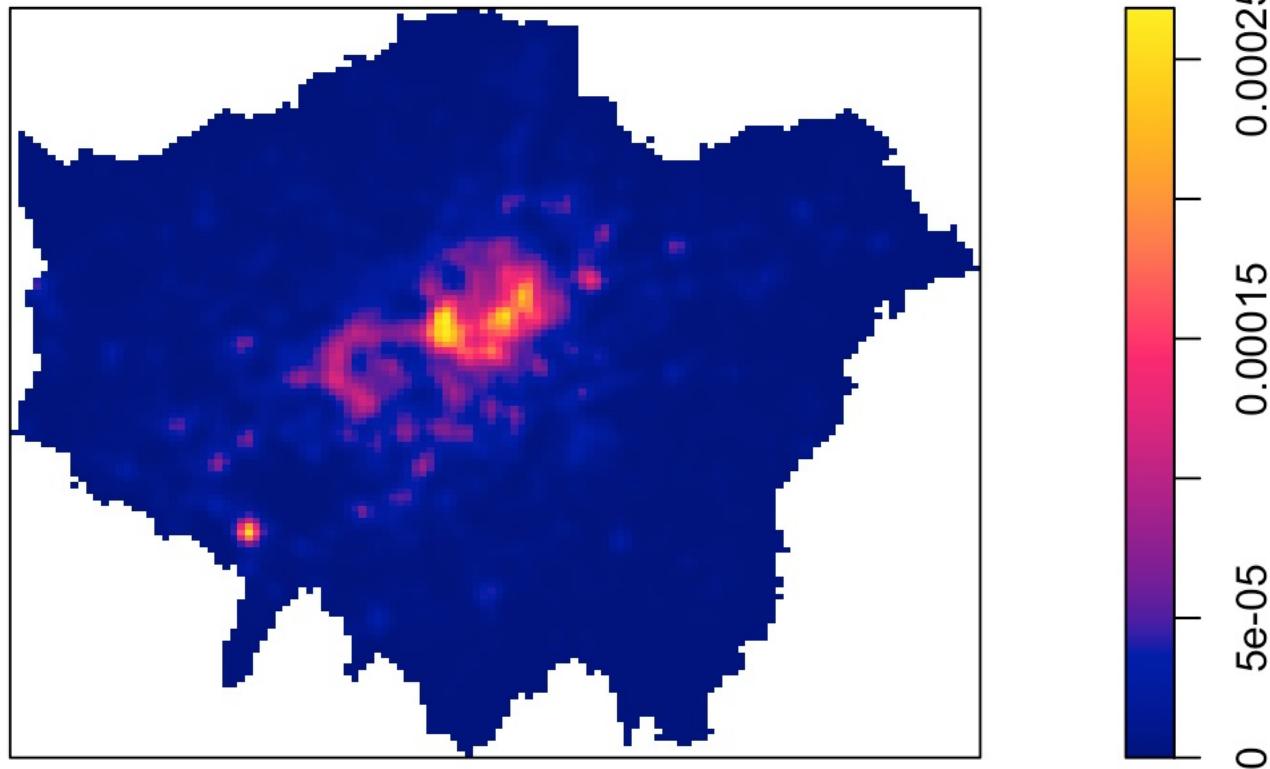


Raster data



Raster data

Gaussian



Raster data

- Multiple data formats, including GeoTIFF, ASCII Grid, Esri Grid and GIF.
- Must be georeferenced (i.e. have a CRS) to plot properly.
- Using raster data also means we should use raster-specific functions and operations.

Raster data

- Local operations applied to each individual cell and only involve those cells sharing the same location
- Focal operations assign to the output cells some summary value of the neighbouring cells
- Zonal operations computes a new summary value from aggregated cells
- Global operations make use of some or all input cells when computing an output cell

Raster data

1	0	0	0	0	0
1	1	0	0	0	0
1	1	1	1	0	1
0	1	1	0	0	0
0	0	0	0	1	0

basic raster

Local operations and functions

1	0	0	0	0	0
1	1	0	0	0	0
1	1	1	1	0	1
0	1	1	0	0	0
0	0	0	0	1	0

multiplication * 2

2	0	0	0	0	0
2	2	0	0	0	0
2	2	2	2	0	2
0	2	2	0	0	0
0	0	0	0	2	0

1	0	0	0	0	0
1	1	0	0	0	0
1	1	1	1	0	1
0	1	1	0	0	0
0	0	0	0	1	0

1	0	0	0	0	0
1	2	0	0	0	0
1	2	1	1	0	1
0	2	1	0	0	0
0	0	0	0	1	0

addition

2	0	0	0	0	0
2	3	0	0	0	0
2	3	2	2	0	2
0	3	2	0	0	0
0	0	0	0	2	0

1	0	0	0	0	0
1	2	0	0	0	0
1	2	1	1	0	1
0	2	1	0	0	0
0	0	0	0	1	0

logical

0	0	0	0	0	0
0	2	0	0	0	0
0	2	0	0	0	0
0	2	0	0	0	0
0	0	0	0	0	0

1	0	0	0	0	0
1	1	0	0	0	0
1	1	1	1	0	1
0	1	1	0	0	0
0	0	0	0	1	0

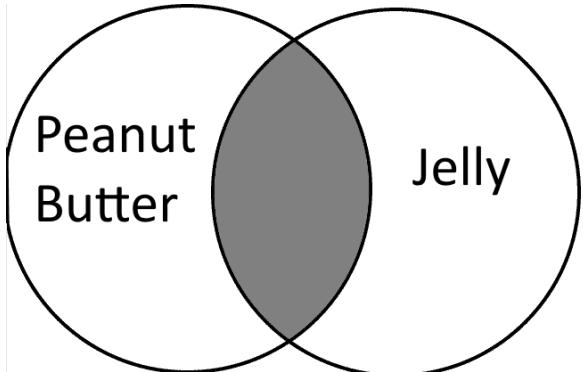
1	0	0	0	0	0
1	2	0	0	0	0
1	2	1	1	0	1
0	2	1	0	0	0
0	0	0	0	1	0

Boolean AND

1	1	1	1	1	1
1	0	1	1	1	1
1	0	1	1	1	1
1	0	1	1	1	1
1	1	1	1	1	1

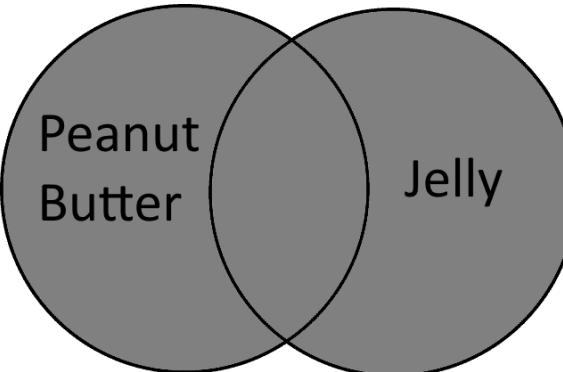
Boolean operators

Truth evaluation using Boolean operators



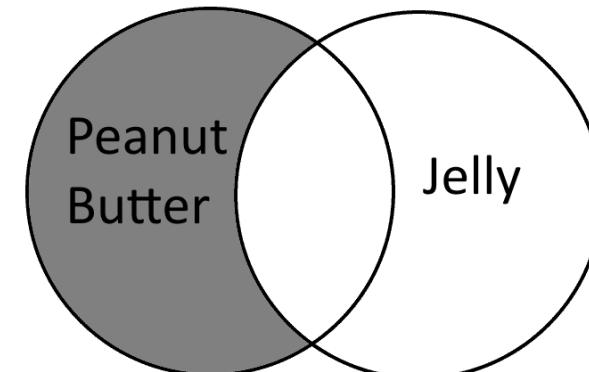
AND

Using AND, this search would only retrieve results with Peanut Butter and Jelly.



OR

Using OR, this search would retrieve results with peanut butter, with jelly, and with both.



NOT

Using NOT, this search would retrieve results with peanut butter, and exclude those with jelly or PB with jelly.

Focal operations and functions

1	0	0	0	0	0
1	2	0	0	0	0
1	2	1	1	0	1
0	2	1	0	0	0
0	0	0	0	1	0

focal (sum)

3	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

1	0	0	0	0	0
1	2	0	0	0	0
1	2	1	1	0	1
0	2	1	0	0	0
0	0	0	0	1	0

focal (sum)

3	0	0	0	0	0
6	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

1	0	0	0	0	0	0
1	2	0	0	0	0	0
1	2	1	1	0	0	1
0	2	1	0	0	0	0
0	0	0	0	1	0	0

focal (sum)

3	0	0	0	0	0
6	0	0	0	0	0
0	0	8	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Zonal operations and functions

1	0	0	0	0	0
1	1	0	0	0	0
1	1	1	1	0	1
0	1	1	0	0	0
0	0	0	0	1	0

1	0	0	0	0	0
1	2	0	0	0	0
1	2	1	1	0	1
0	2	1	0	0	0
0	0	0	0	1	0

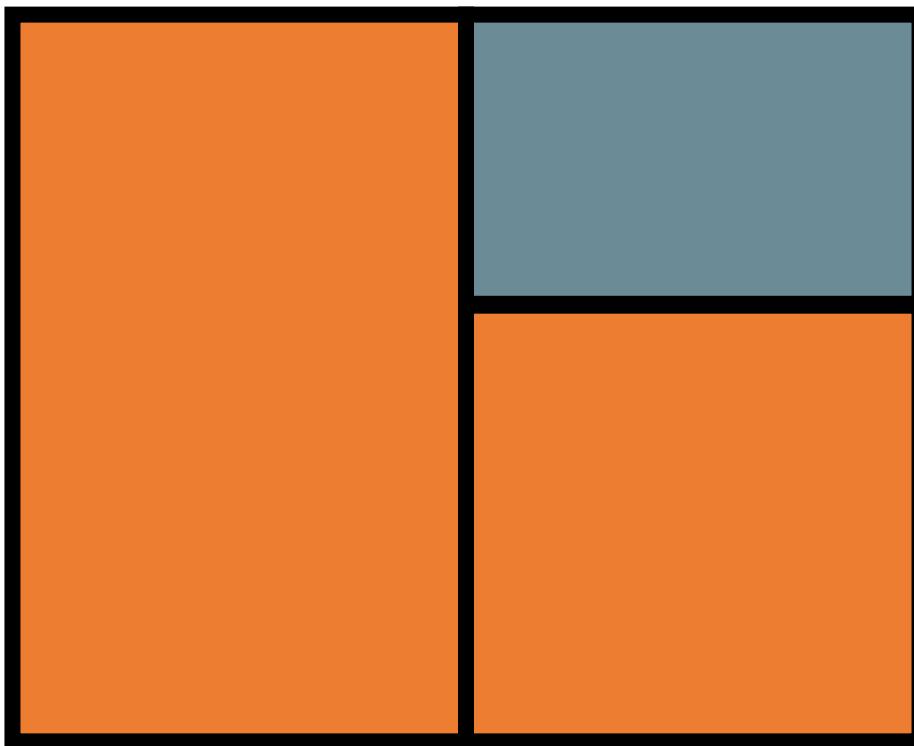
zonal

11	11	11	0	0	0
11	11	11	0	0	0
11	11	11	3	3	3
11	11	11	3	3	3
11	11	11	3	3	3

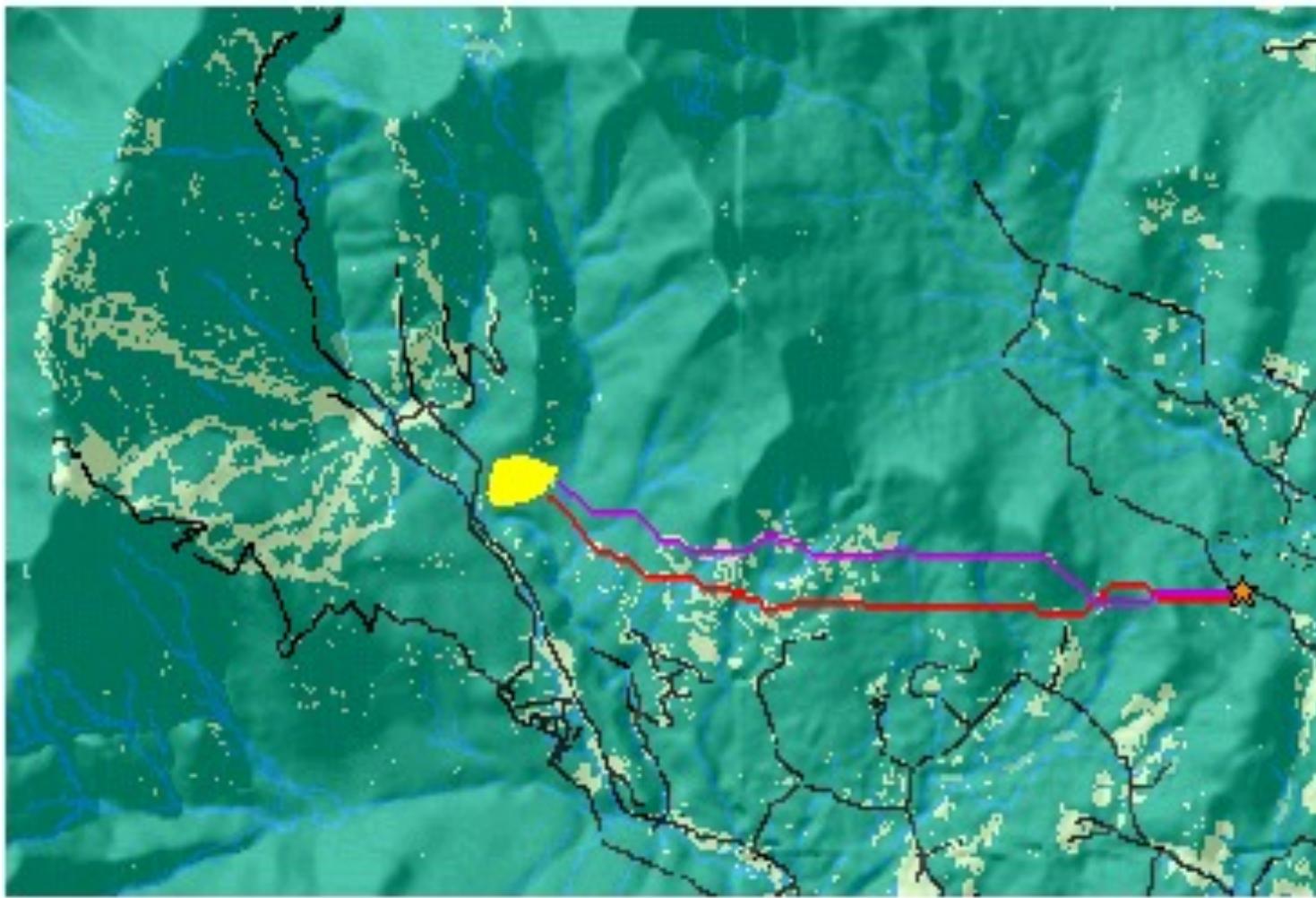
1	0	0	0	0	0
1	1	0	0	0	0
1	1	1	1	0	1
0	1	1	0	0	0
0	0	0	0	1	0

1	0	0	0	0	0
1	2	0	0	0	0
1	2	1	1	0	1
0	2	1	0	0	0
0	0	0	0	1	0

zonal



Global operations and functions



Least cost path (e.g. height/slope for modeling water flow)

Multi-criteria decision support system for wind farm site selection using GIS

Contents lists available at ScienceDirect

Sustainable Energy Technologies and Assessments

journal homepage: www.elsevier.com/locate/seta

Original Research Article

Multi-criteria decision support system for wind farm site selection using GIS

Younes Noorollahi*, Hossein Yousefi, Mohammad Mohammadi

Department of Renewable Energies and Environmental Eng., Faculty of New Sciences and Technologies, University of Tehran, Iran

ARTICLE INFO

Article history:
Received 10 May 2015
Revised 19 November 2015
Accepted 19 November 2015

Keywords:
Wind energy
GIS
Site selection
Multi criteria
Wind farm
Iran

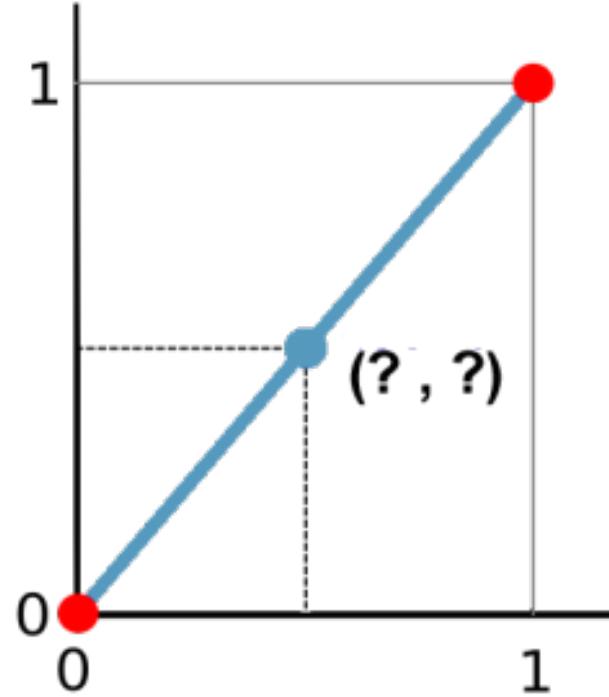
ABSTRACT

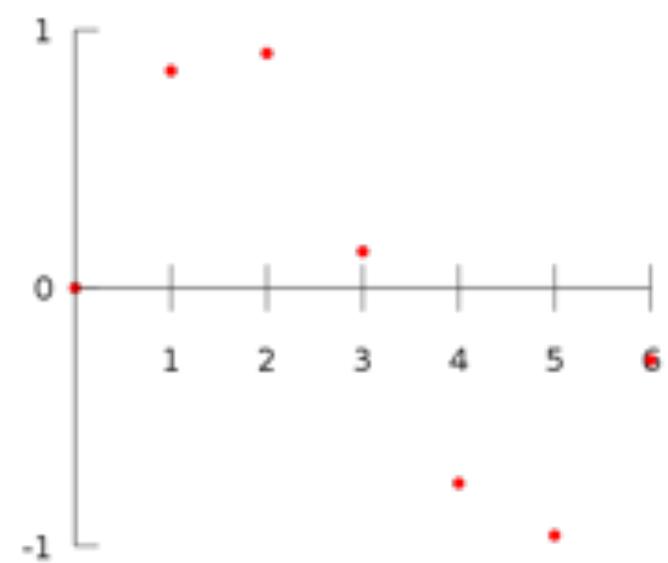
The present study analyzed a multi-criteria decision support system to define wind energy resources in western Iran. Clean domestic renewable energy can be the best option in consideration of intense economic development and its accompanying increase in energy consumption. The most important barrier to wider deployment of renewable resources in Iran is the price of fossil fuels, which is the lowest in the world. The government has recently decided to remove subsidies for fossil fuel, meaning that its price will increase and will make the cost of green energies more attractive.

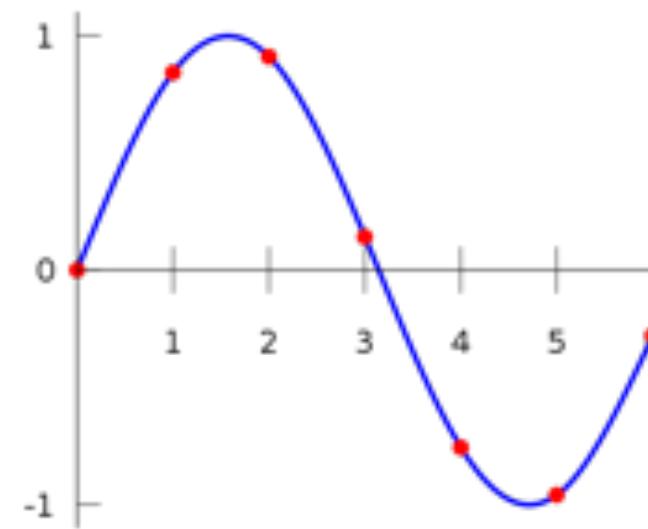
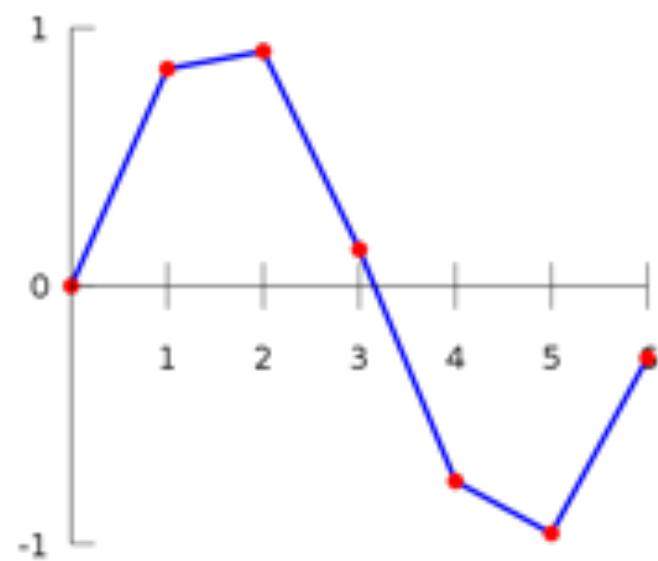
Wind power is an option for improved economic conditions in the region and low environmental impacts. This study applied geographic information system to determine the potential of wind energy in Markazi province in western Iran. The multiple criteria decision making method and site selection criterion for wind resources assessment is explained and developed for the study area. Criteria of equal importance were investigated, including technical, environmental, economic and geographic standards. The results were favorable for electricity production in accordance with international standards from wind in western Iran. The results show that 28% of the study area has capacity for installing large wind farms.

© 2015 Elsevier Ltd. All rights reserved.

Data interpolation



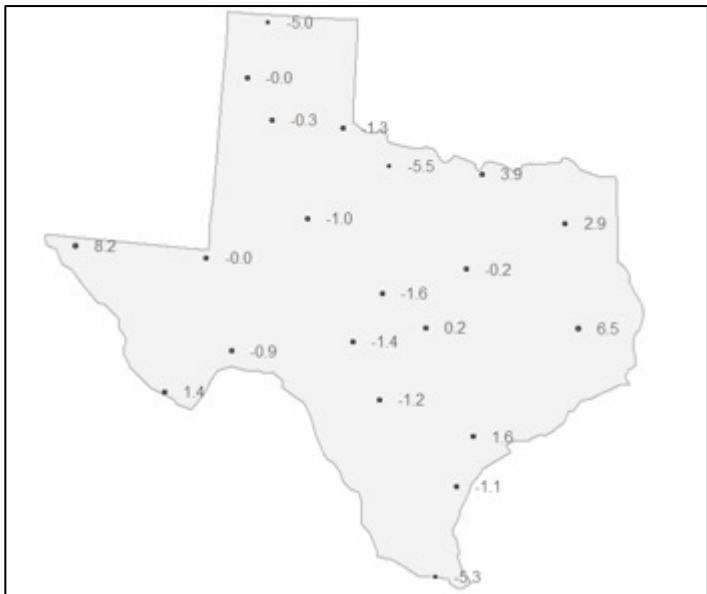




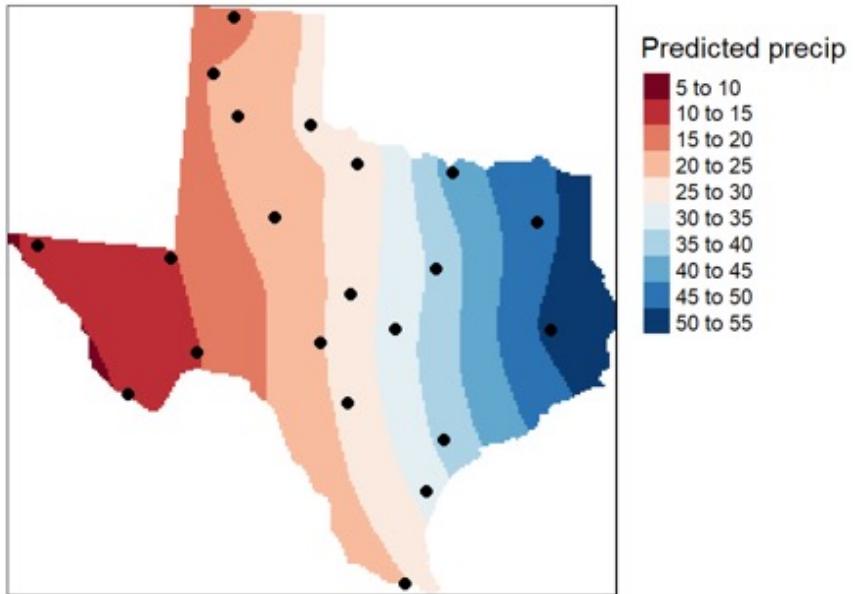
Spatial data interpolation

- Same idea but we need to account for the importance of space.
- Creation of a continuous (or prediction) surface from sampled point values.
- Common for difficult to measure phenomena (e.g. concentrations over space).
- Using regularly distributed or sampled points.
- Multiple approaches to derive a prediction.

Spatial data interpolation



Gimond, M. 2020. *Geodesic geometry*. [online] <https://mgimond.github.io/>



Spatial data interpolation

- Difficult to take measures for everything – instead measure the phenomenon at strategically dispersed sample locations and predicted values can be assigned to all other locations.
- Points can be randomly or regularly spaced or even based on a sampling scheme.
- Multiple approaches to derive a prediction: each method is a type of model, as a result, different assumptions made of the data.

Spatial data interpolation

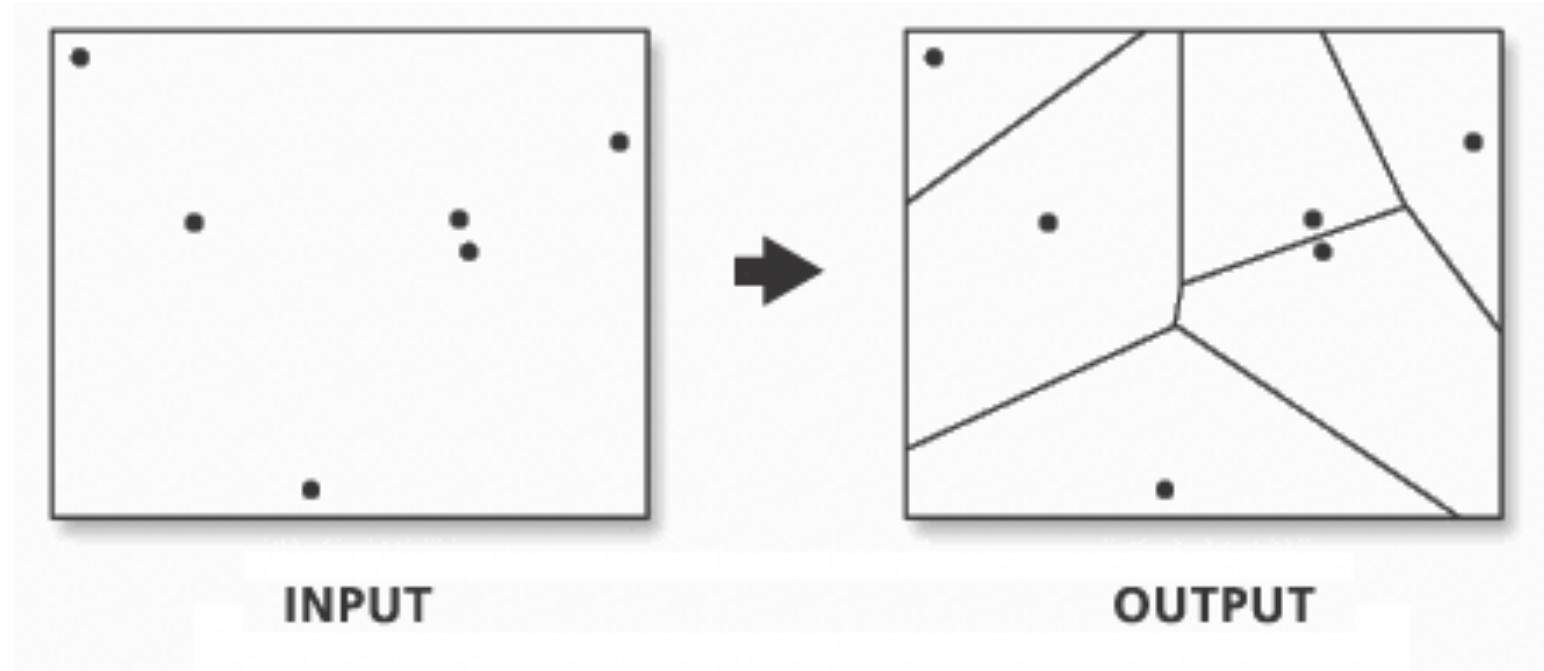
Deterministic methods

These type of models have parameter values that are typically arbitrarily defined.

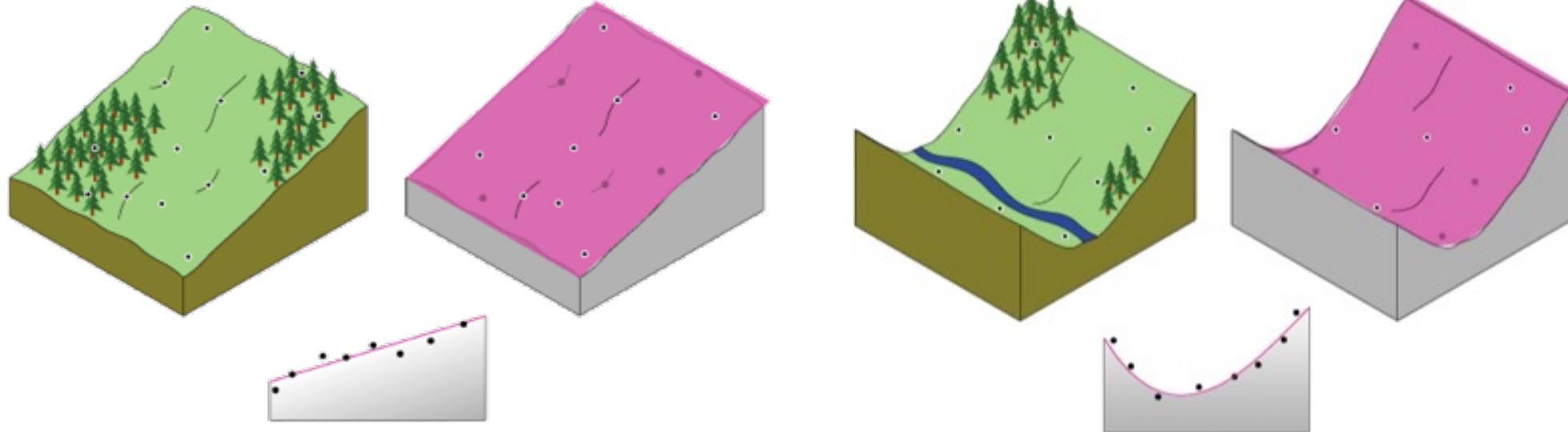
Geostatistical methods

The parameter values for the these set of models have to be estimated.

Spatial data interpolation: nearest neighbour



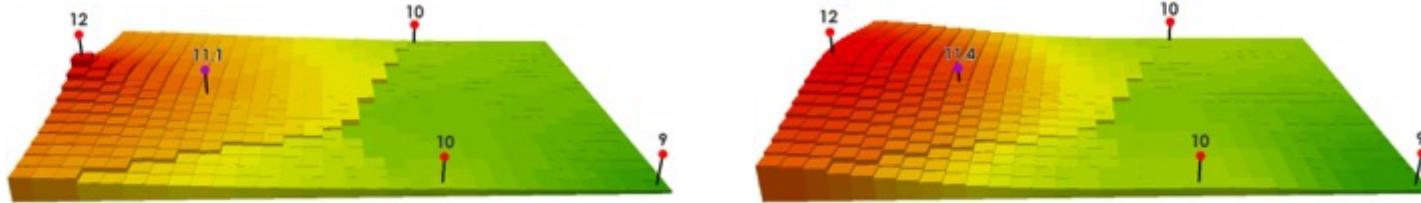
Spatial data interpolation: trend



Spatial Interpolation: IDW

- IDW: Inverse Distance Weighting (“Tobler’s Law”)
- Assumes near points are more alike than far points (distance decay)
- Spatial autocorrelation is the underlying assumption of IDW

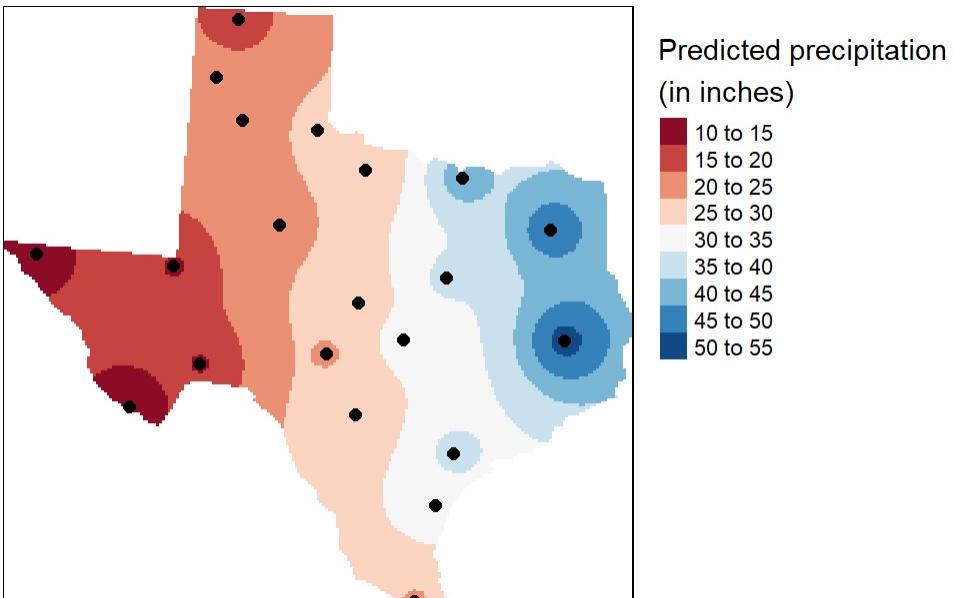
Spatial data interpolation: IDW



IDW with different power settings

$$z_p = \frac{\sum_{i=1}^n \left(\frac{z_i}{d_i^p} \right)}{\sum_{i=1}^n \left(\frac{1}{d_i^p} \right)}$$

Spatial data interpolation: IDW

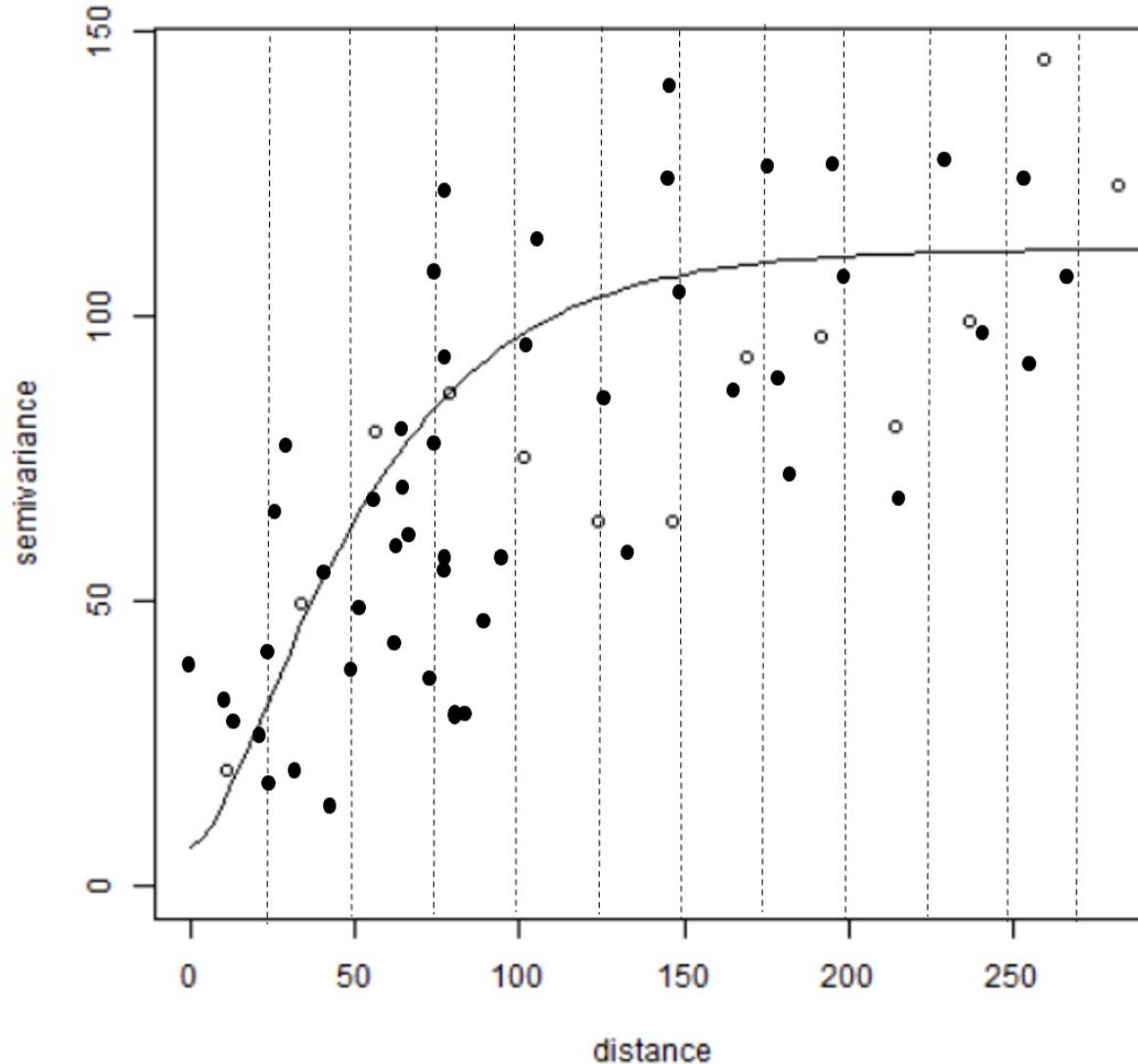


Gimond, M. 2020. *Geodesic geometry*. [online] <https://mgimond.github.io/>

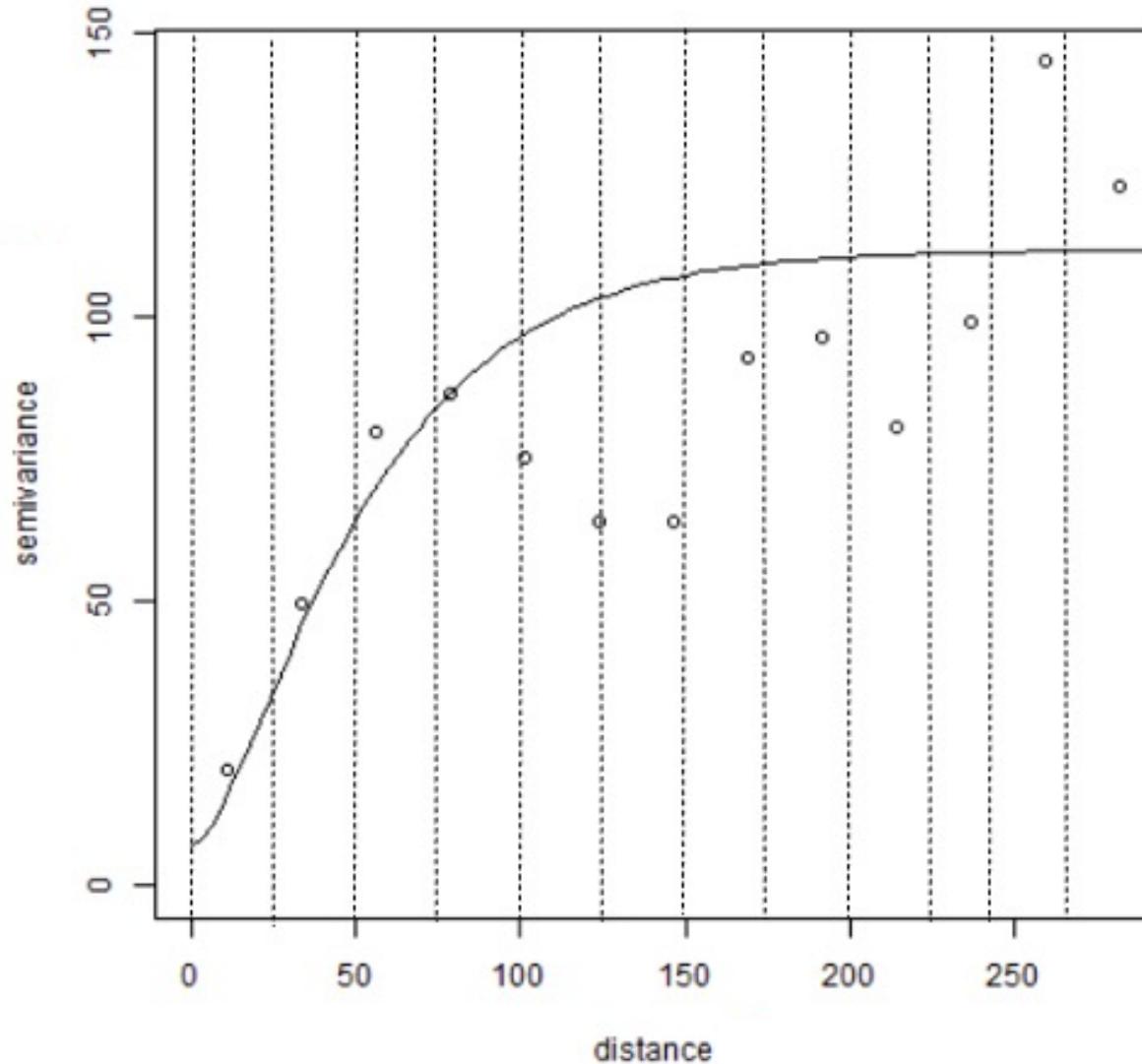
Spatial data interpolation: Kriging

- Considers both distance and degree of variation between known data points when estimating values in unknown areas.
- Assumption also here: closer things are more predictable and have less variability, while distant things are less predictable and less related.
- Kriging is a complex, multistep process starting with a semi-variogram.

Spatial data interpolation: Kriging

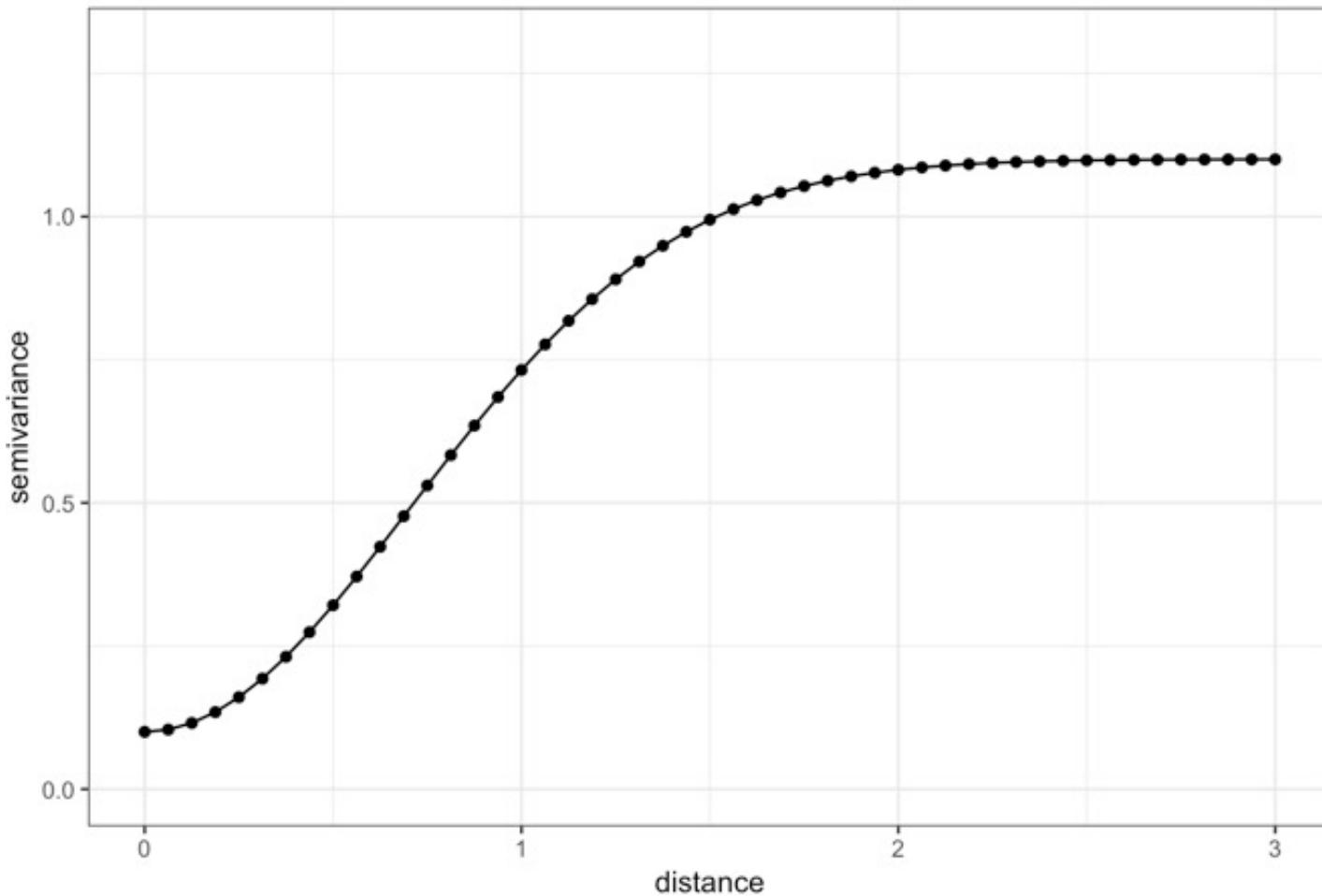


Spatial data interpolation: Kriging



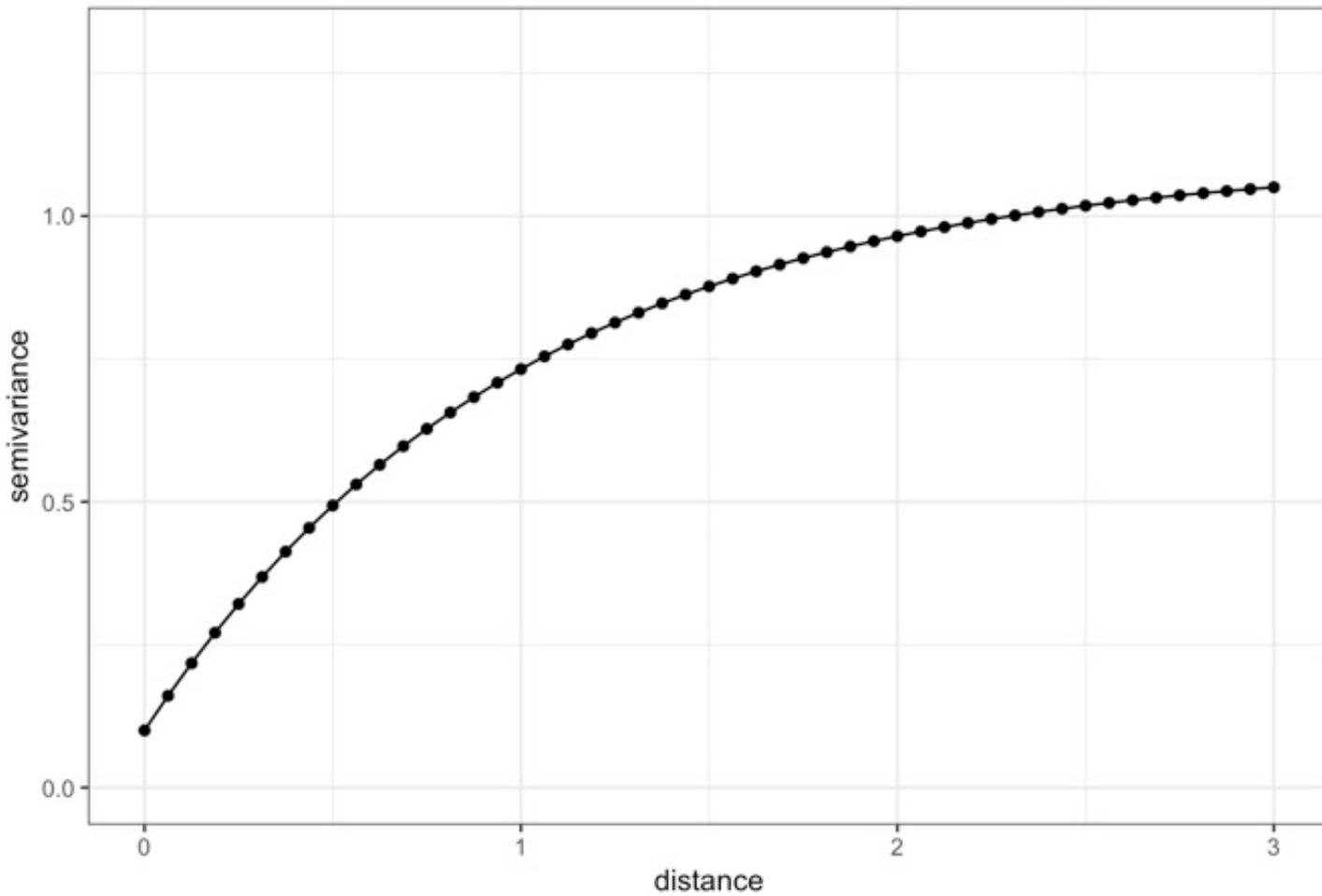
Spatial data interpolation: Kriging

Gaussian variogram model; Nugget = 0.1



Spatial data interpolation: Kriging

Exponential variogram model; Nugget = 0.1



IDW versus Kriging

- IDW assumes that spatial autocorrelation between neighboring points is proportional to the distance (and that it can be defined by distance reverse function).
- Kriging assumes that distance (mainly) or directionality between sampling points reflects the spatial autocorrelation, and functions can be fitted to describe the correlation between points (and explain the variation on the surface).

Conclusion

- Raster data comes with raster specific functions and operations.
- Raster data become important in the context of certain types of categorical data as well as continuous data.
- Spatial data interpolation is the idea to estimate unknown values of a phenomenon – spatial autocorrelation is the underlying process.
- Because spatial data interpolation tries to create a continuous surface we need to use the raster data model.
- Spatial interpolation is different than Point Pattern Analysis which tries to predict densities.

Computer tutorial

- Today: two types of raster datasets to analyse population change and air pollution in London, including some interpolation.
- Carefully read the instructions although there *may* be small deviations in the syntax.
- Assignment: no need to hand-in but if you want to leave before the end of the computer tutorial you should be able to show your results.

Geocomputation Extra Help Session

- Monday April 11 from 14h00-16h00
- Torrington Place 113.
- Not mandatory.

Questions

Justin van Dijk

j.t.vandijk@ucl.ac.uk

