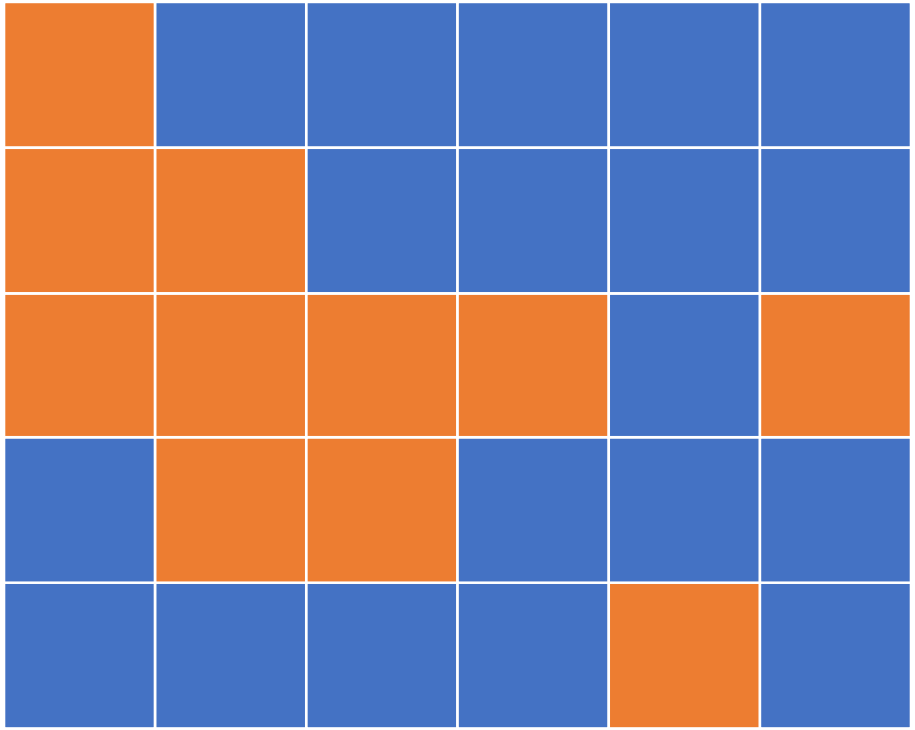
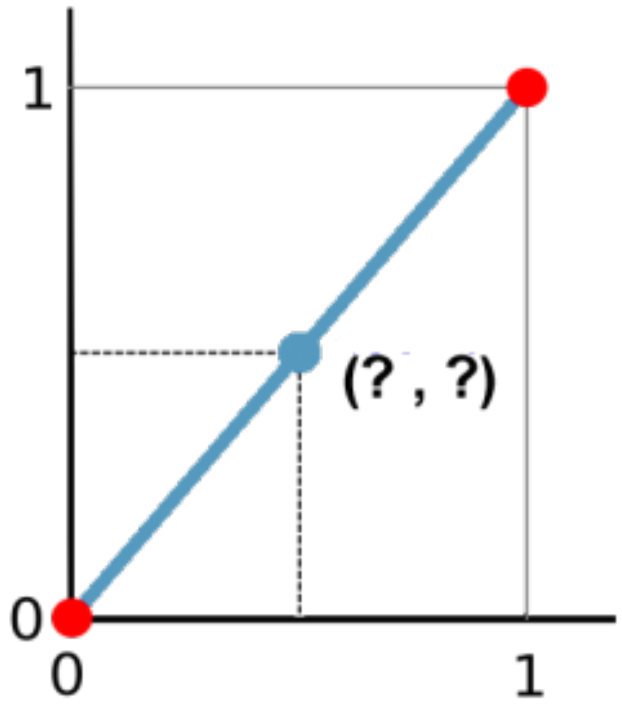
Principles of Spatial Analysis

SHORT LECTURE 02, WEEK 07: INTERPOLATION AND GEOSTATISTICS



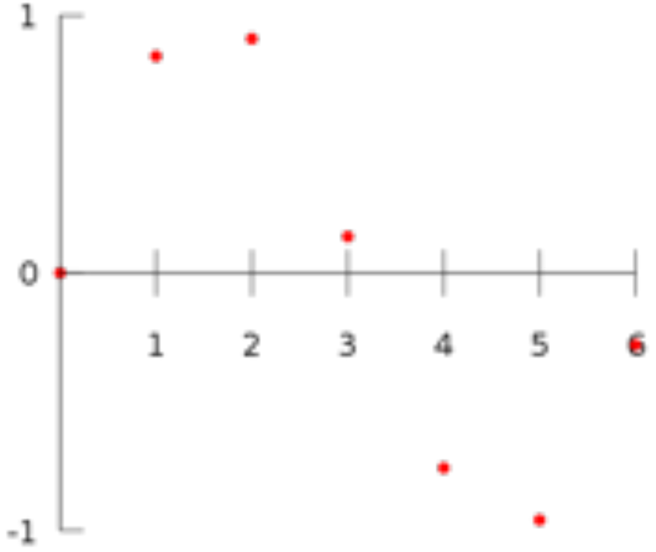
data interpolation

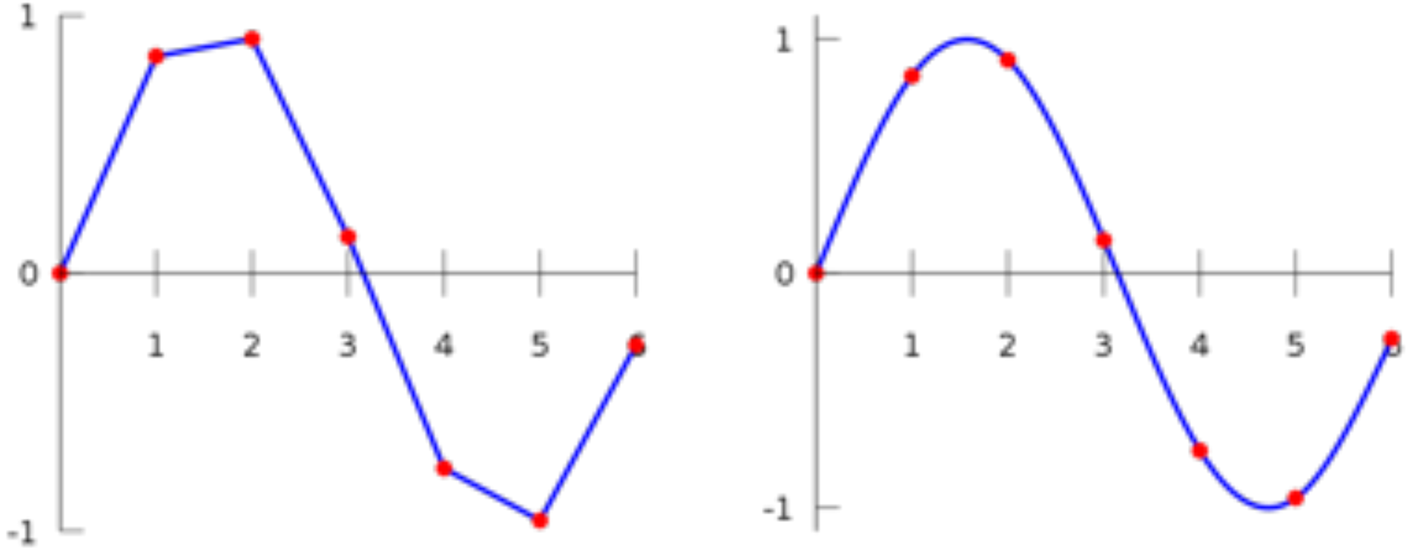
the process of using points with known values to estimate values at other unknown points


1
0
0
1
(? , ?)

data interpolation

data’s distribution will normally determine most suitable method for interpolation


1
0
-1
1
2
3
4
5
6


1
0
-1
1 2 3 4 5 6
1
0
-1
1 2 3 4 5 6

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

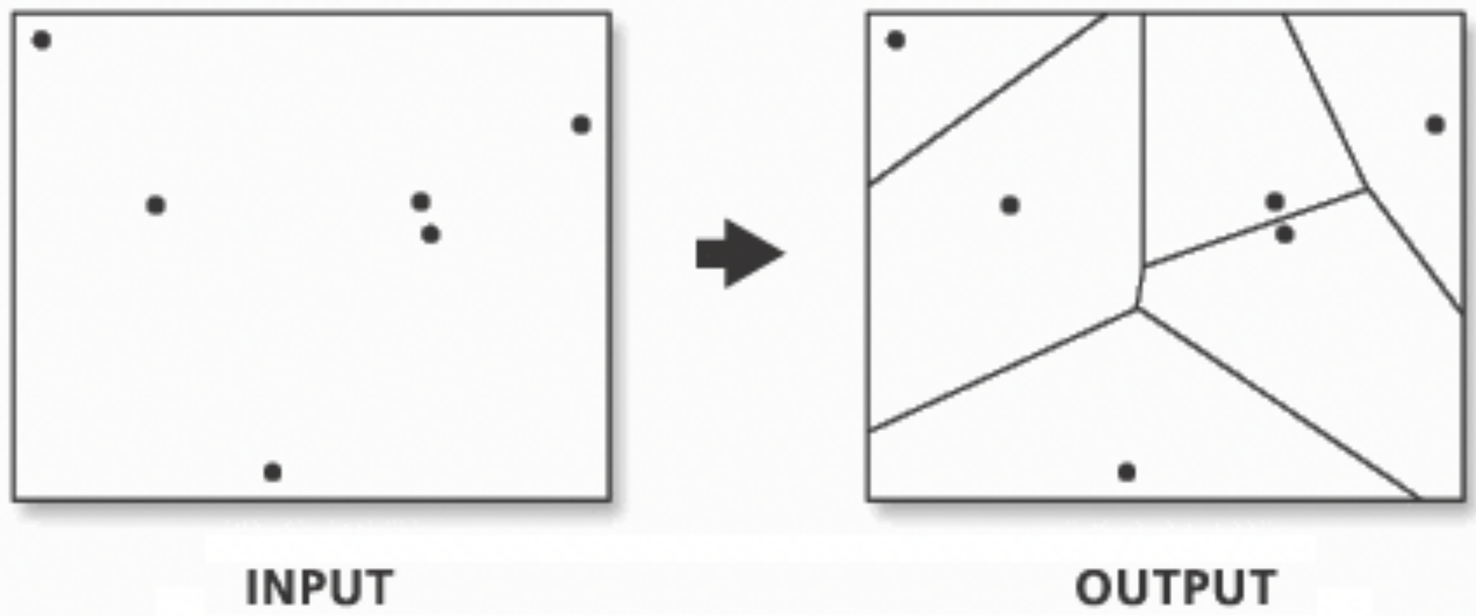
Gimond, M. 2020. *Geodesic geometry*. [online] https://mgimond.github.io/

deterministic methods

assign values to locations based on the surrounding measured values

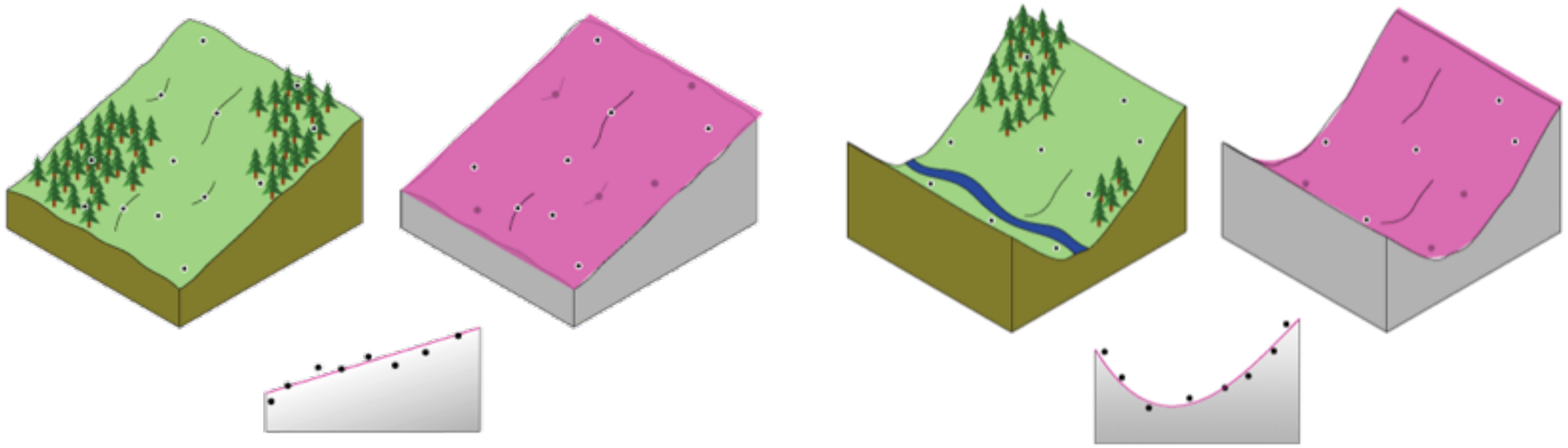
deterministic: natural neighbour

- nearest neighbour interpolation by extending the data, can be spatially expressed through Thiessen polygons


INPUT
OUTPUT

deterministic: trend

- uses a global polynomial interpretation that fits a smooth surface defined by a mathematical function (the polynomial) to the input sample points
- fitting a sheet of paper over a surface



deterministic: trend

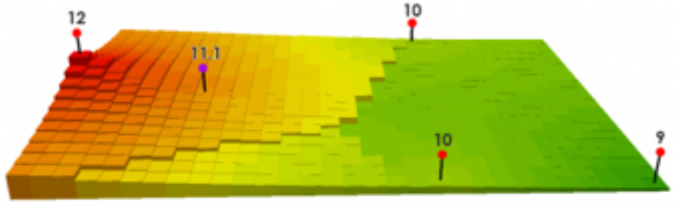
- rarely will the surface be able to pass through all the measured points
- least-squares regression fit: minimises the squared differences among the original values and result surface
- performance can be measured through root mean square error (RMSE)

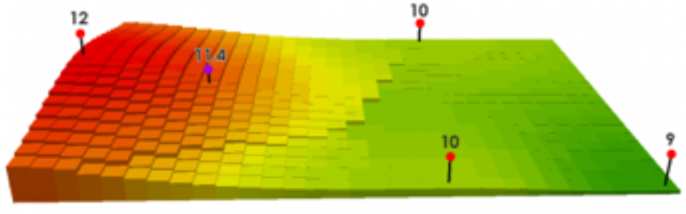
deterministic: IDW

- 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

deterministic: IDW

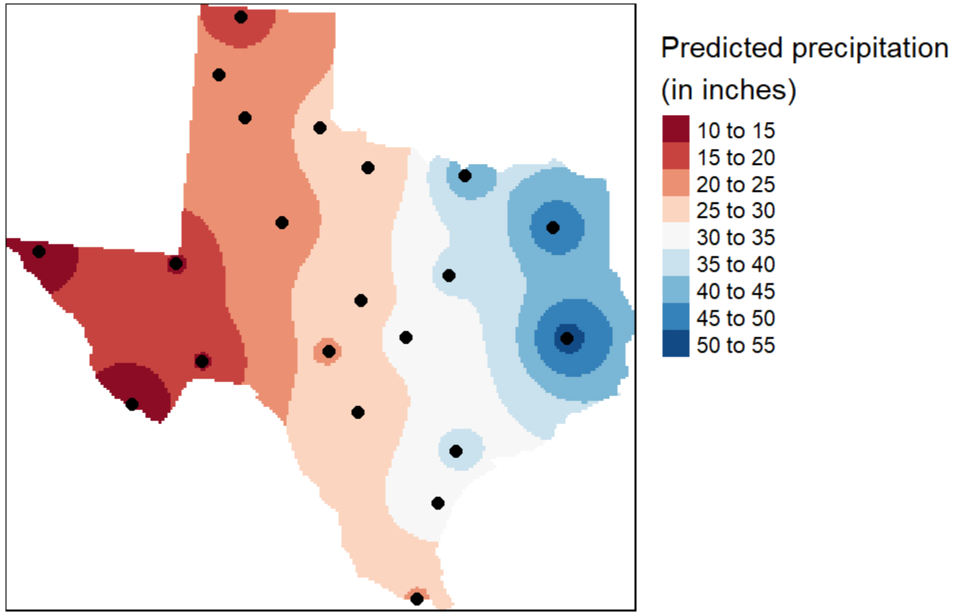
- only use IDW when the set of points are dense enough to capture the extent of the local surface variation needed for analysis
- interpolated surface will always be less than the local maximum value and greater than the local minimum value
- the power setting determines how much influence to give closer points than those further away by enhancing extremes when calculating averages





IDW with different power settings

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



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

geostatistical methods

based on statistical models that include autocorrelation

geostatistical: Kriging

- considers both distance and degree of variation between known data points when estimating values in unknown areas
- most appropriate when there is spatially correlated distance or directional bias in the data
- Kriging is a complex, multistep process

geostatistical: Kriging

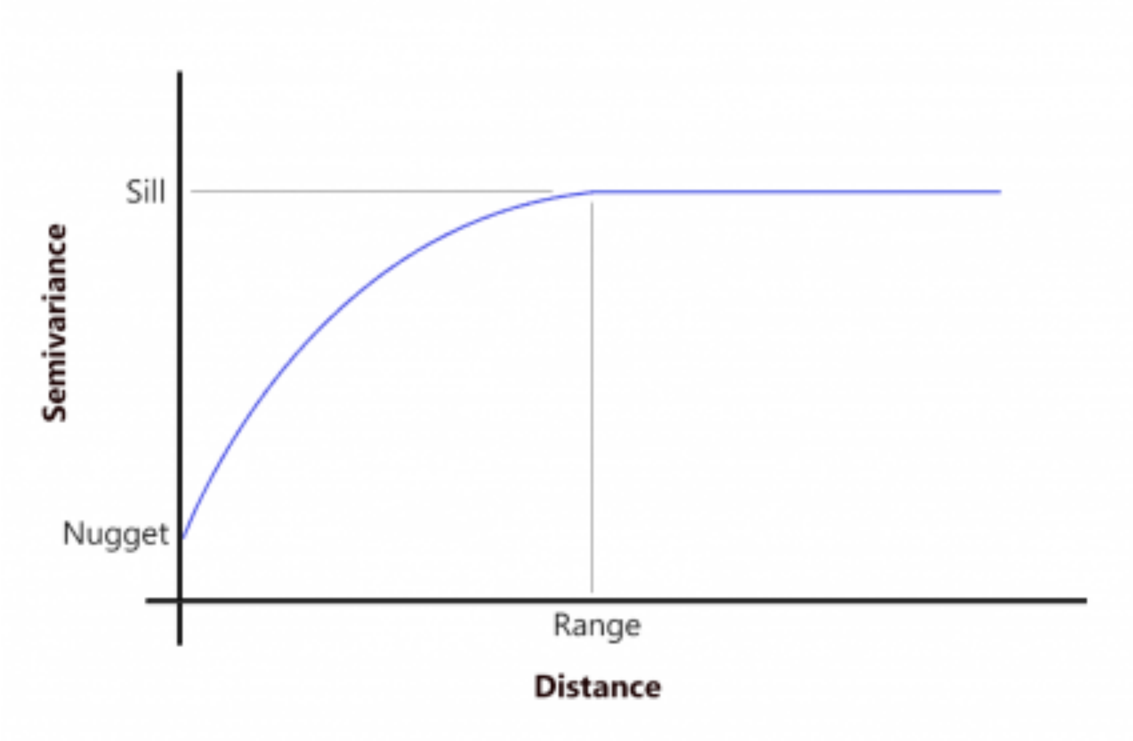
- values of your input data need to be normally distributed
- values are stationary (i.e. local variations do not change in different areas in the map)
- values cannot show a trend or drift (e.g. systematic difference in rainfall)

geostatistical: semi-variogram

- Kriging requires a **semi-variogram**
- closer things are more predictable and have less variability, while distant things are less predictable and less related
- semi-variograms are used to chart how sample values will vary with distance

geostatistical: semi-variogram

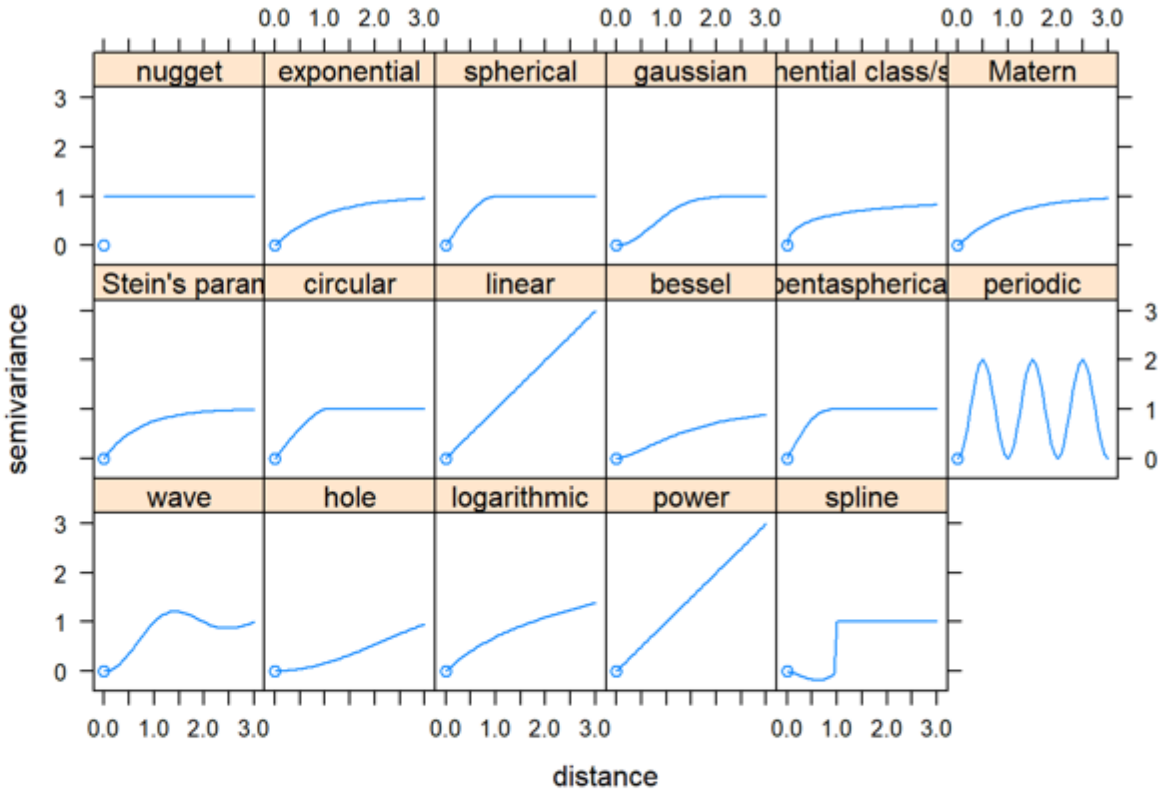
- semi-variogram takes two sample points and calculates the distance between them; subsequently plots the distance versus the semi-variance (i.e. difference squared, then halved)
- often uses lags (grouped distances) with average semi-variance for each lag



Sill: The value at which the model first flattens out (no more spatial autocorrelation)
Range: The distance at which the model first flattens out
Nugget: The value at which the semi-variance intercepts the y-value

geostatistical: Kriging

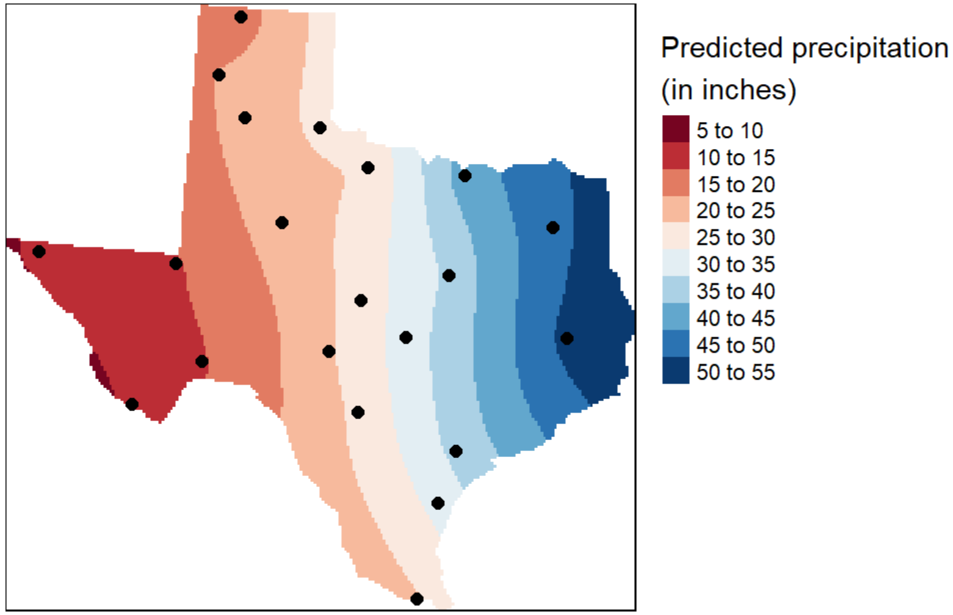
- we need to choose a mathematical model function to interpolate our data
- use the semi-variogram to identify the best fitting model, tweaking the **sill**, **range** and **nugget** parameters



nugget
exponential
spherical
gaussian
nential class/s
Matern
Stein's paran
circular
linear
bessel
pentaspherica
periodic
wave
hole
logarithmic
power
spline
semivariance
distance
0.0 1.0 2.0 3.0
0 1 2 3

geostatistical: Kriging

- chosen mathematical model is used by the Kriging interpolator to predict responses at each location using a weighted average with nearest neighbours
- Kriging appears to give a ‘smoother’ surface than IDW (bull’s eye effect), but more complex to run



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

let’s put it into practice