Spatial Data Analysis in R

Point Patterns and Analysis

Eco 697DR – University of Massachusetts, Amherst – Spring 2022
Michael France Nelson

Assessing Significance

Assessing Significance: Unmarked Point Patterns

- What questions can we ask with univariate point pattern analyses?
- What is our null hypothesis?
- What are the properties of a point pattern under the null?

Assessing Significance: Unmarked Point Patterns

- What questions can we ask with univariate point pattern analyses?
  - Aggregation, dispersion
- What is our null hypothesis?
  - No pattern
- What are the properties of a point pattern under the null?
  - Uniform distribution in space.
  - Quadrat counts follow a Poisson distribution
- How to assess significance?

Significance for aggregation functions

How can we tell if departures from CSR are significant?

Simulations of CSR, or other appropriate point generating process, can generate an *envelope* to use for significance testing.

Stationary/Homogeneous Processes	Sources of Inhomogeneity
• The position of each point does not depend on the positions of other points • points are independent • Points are equally likely to occur anywhere in space • the intensity is constant throughout space • This sounds like Complete Spatial Randomness! (because it is) • Compare to inhomogeneous patterns.	• Gradients • Covariates • Non-constant intensity • Non-independent point positions • Parent/offspring processes • Whether a point pattern is homogeneous or inhomogeneous can depend on your choice of scale and/or sampling unit

Point generating processes

CSR: Poisson process

Inhomogeneous Poisson patterns: non-constant intensity

- Cox process: intensity λ varies in space

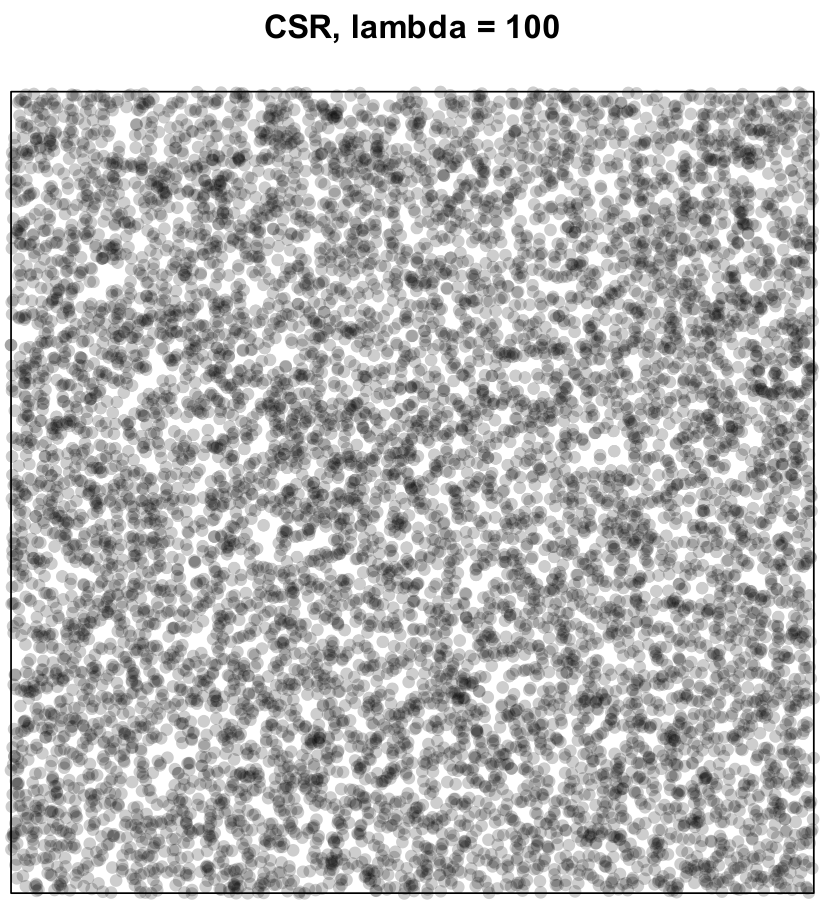
Offspring point patterns: clustering

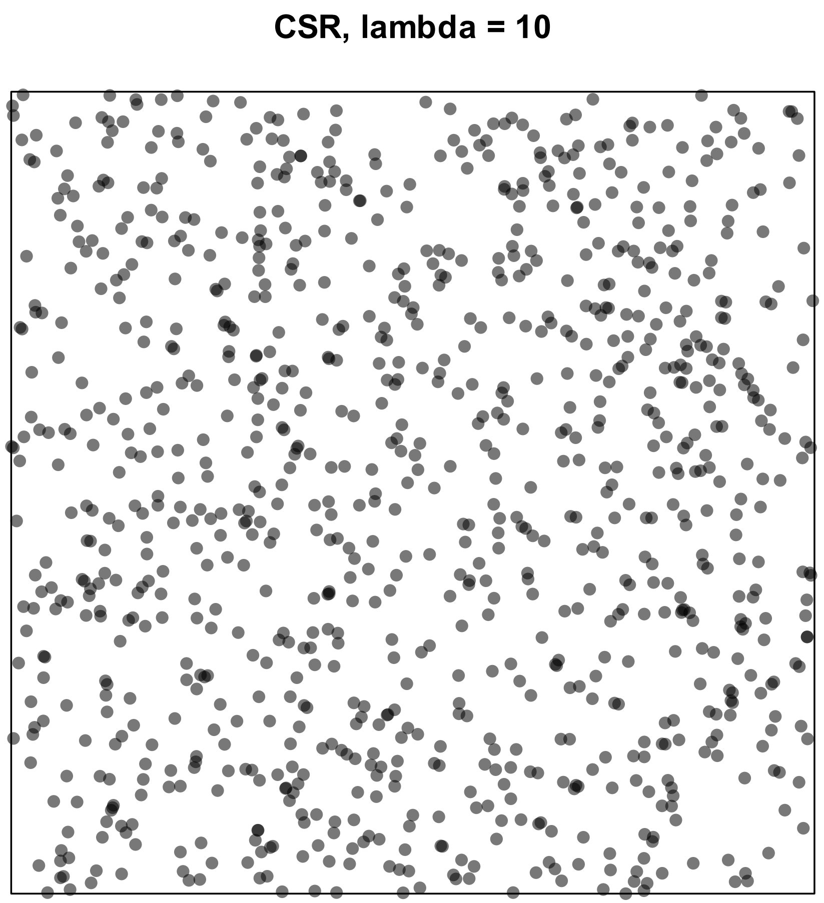
- Parent points are CSR.
- Matérn: offspring are CSR within a radius of parent points.
- Thomas: offspring are Normally-distributed about parent

Processes with repulsion: overdispersion

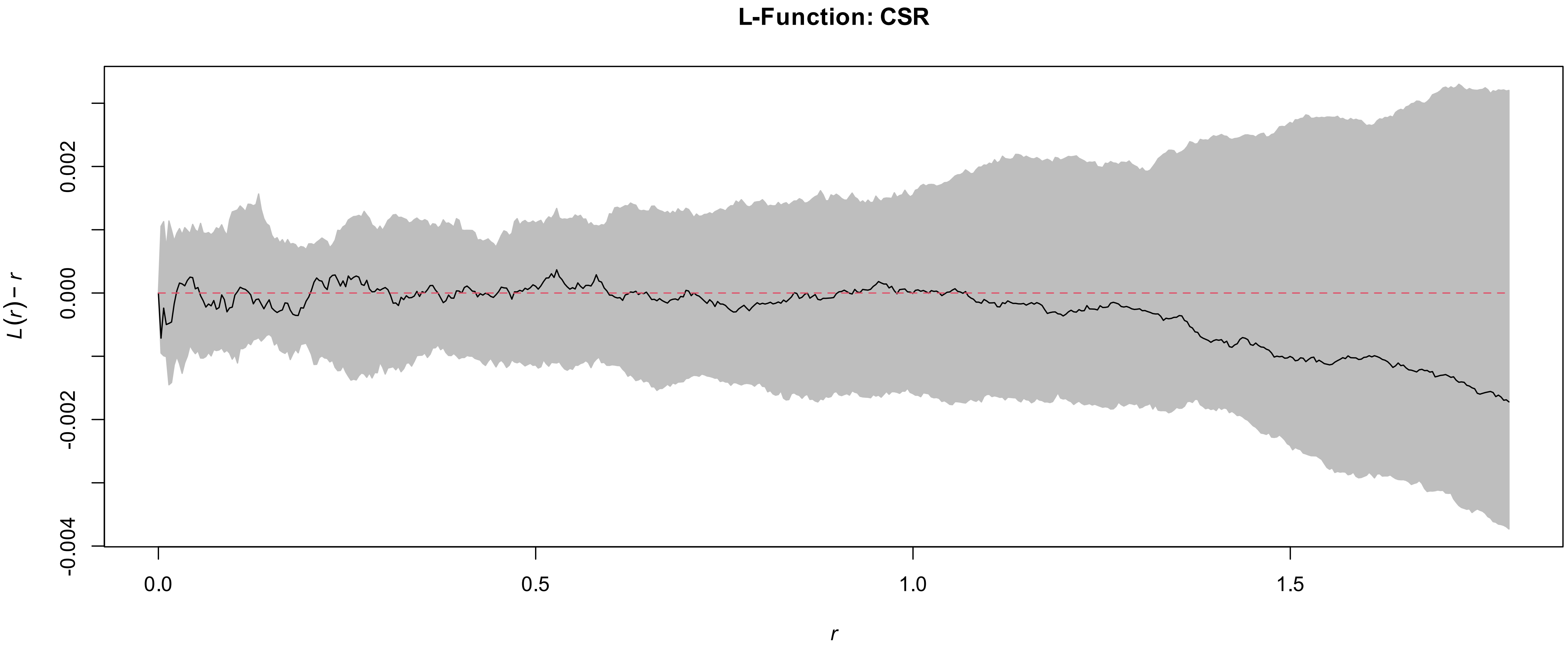
- Inhibition zone
- Hard-core and soft-core processes: minimum core distance

Point Process Examples CSR Pattern





L-function: CSR



Significance: Simulation

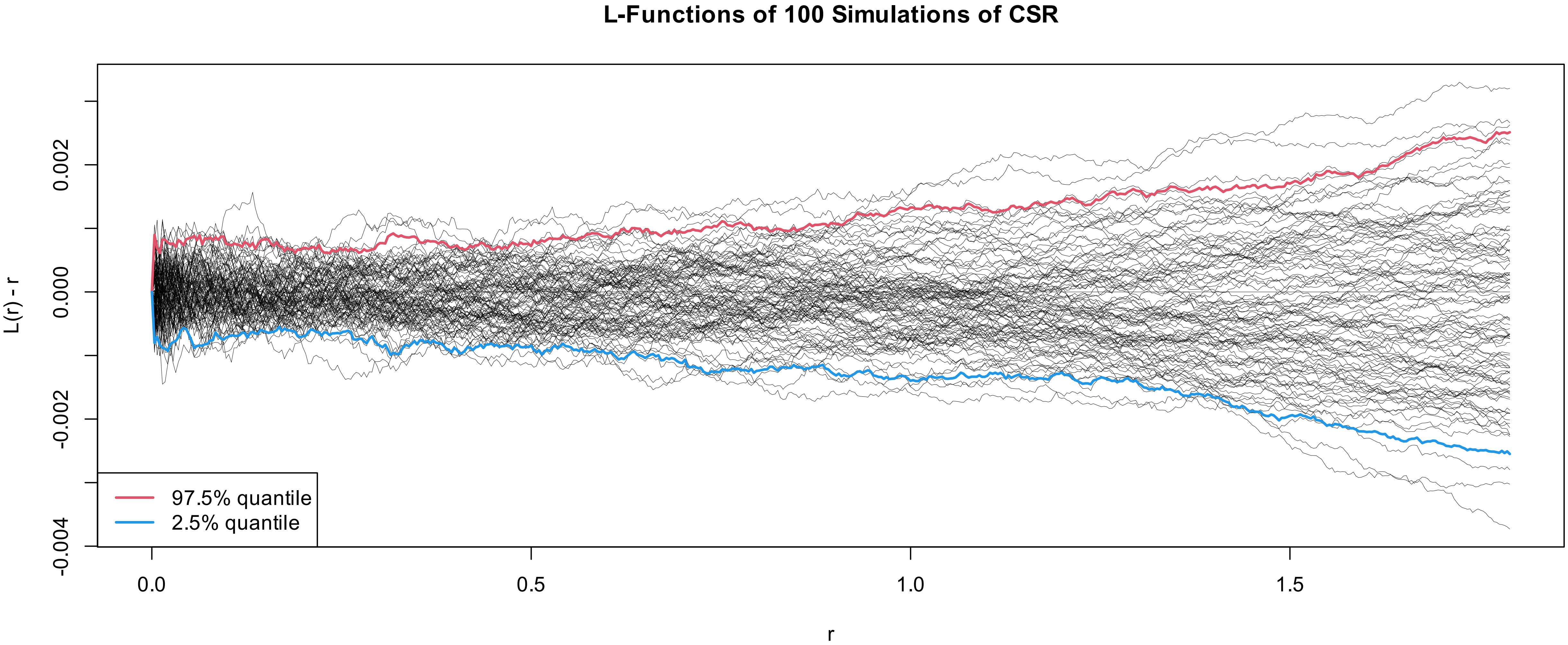
Procedure to create a 95% simulation envelope:

- Simulate a CSR pattern
- Calculate L(r) or g(r) for the simulation, record results
- Repeat steps 1 and 2 n times
- Calculate the upper and lower 2.5% quantiles

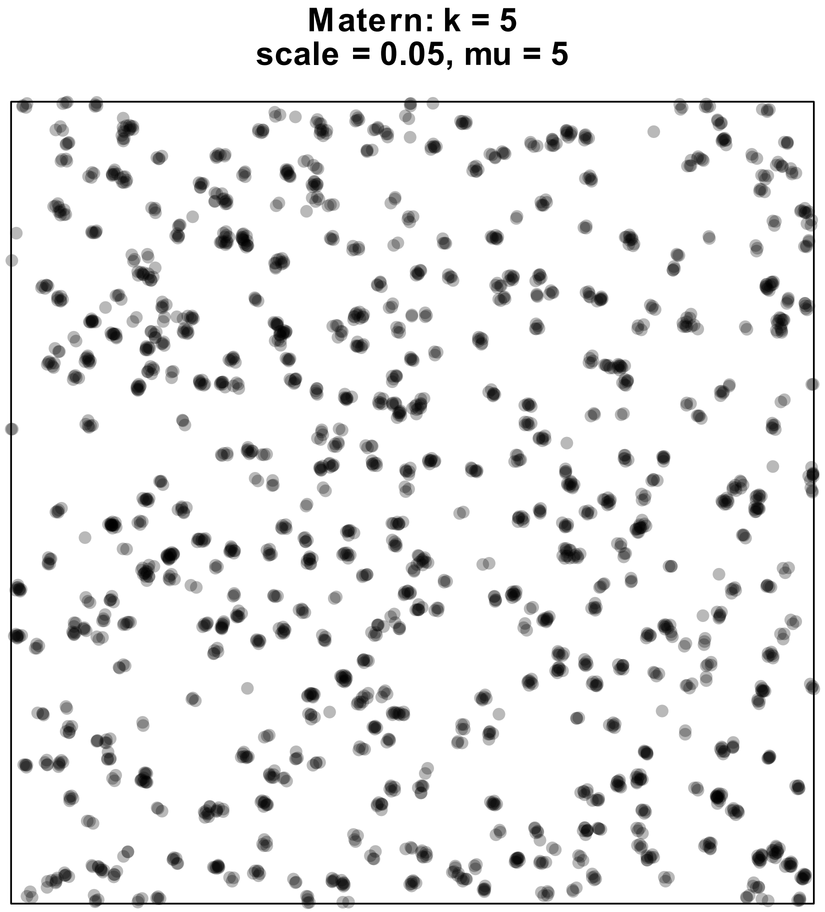
Interpreting the envelope

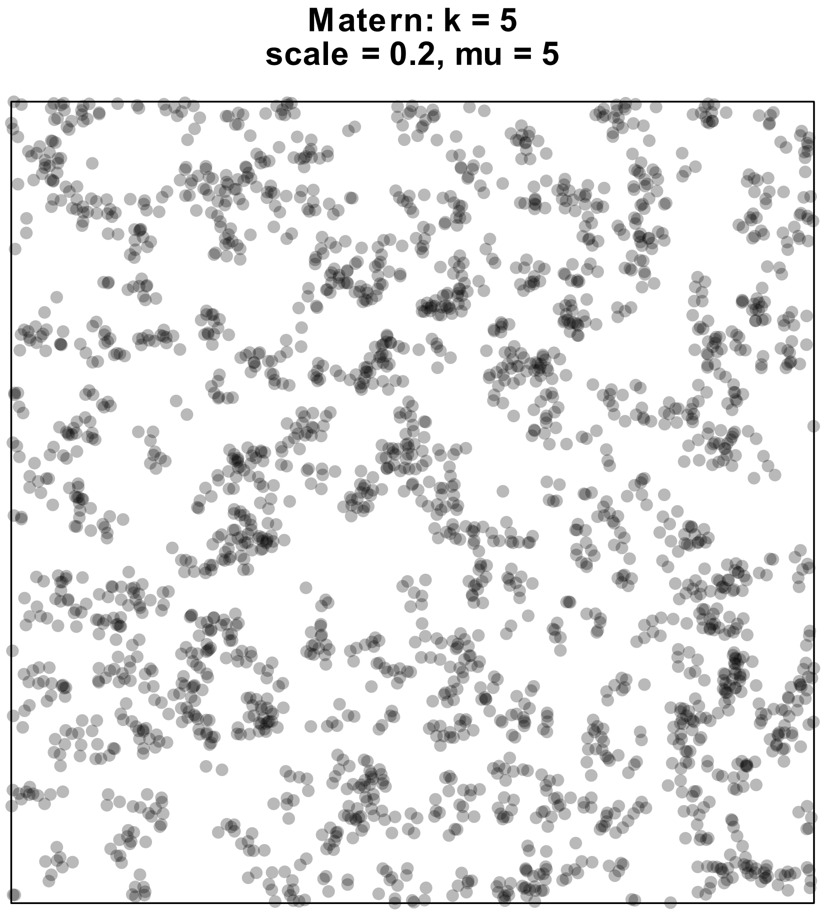
- We know that 95% of simulated CSR values fall within the envelope.
- If the observed L or g (or G, or K) fall within the envelope, they are consistent with CSR.
- If the observed function values are outside the envelope, we can reject the null of CSR

Significance via Simulation

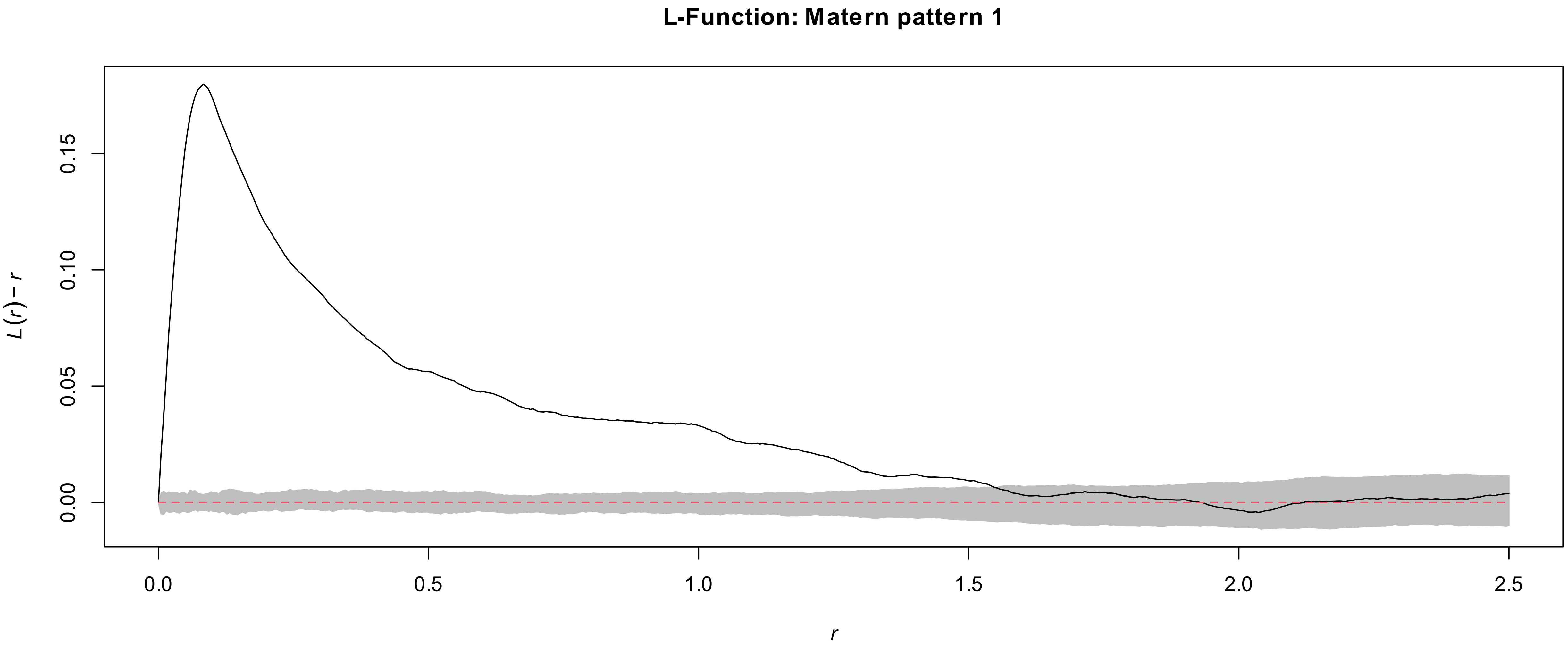


Matern Pattern

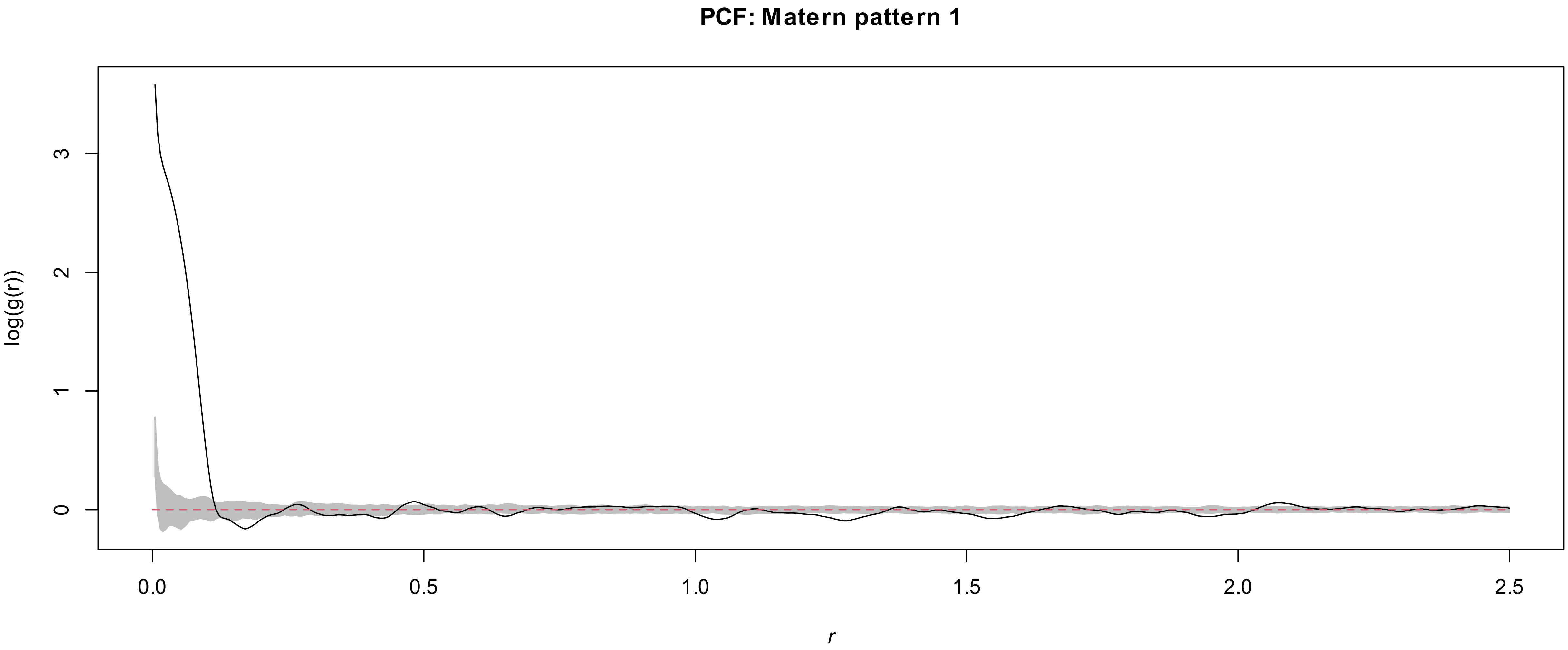




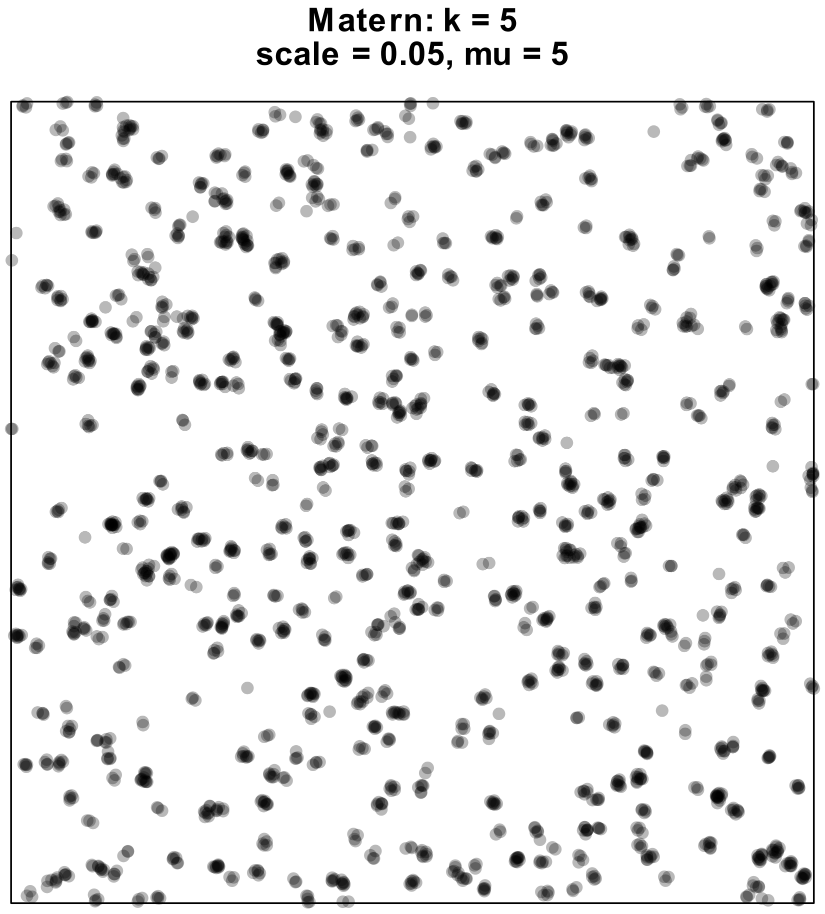
L-function: Matern

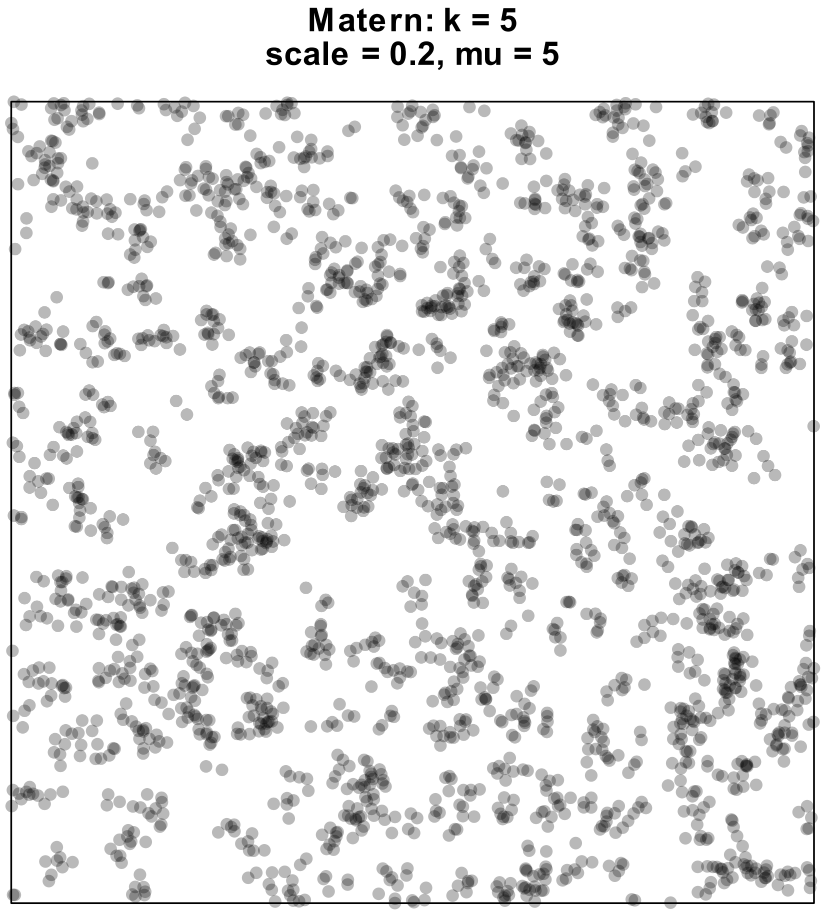


PCF-function: Matern

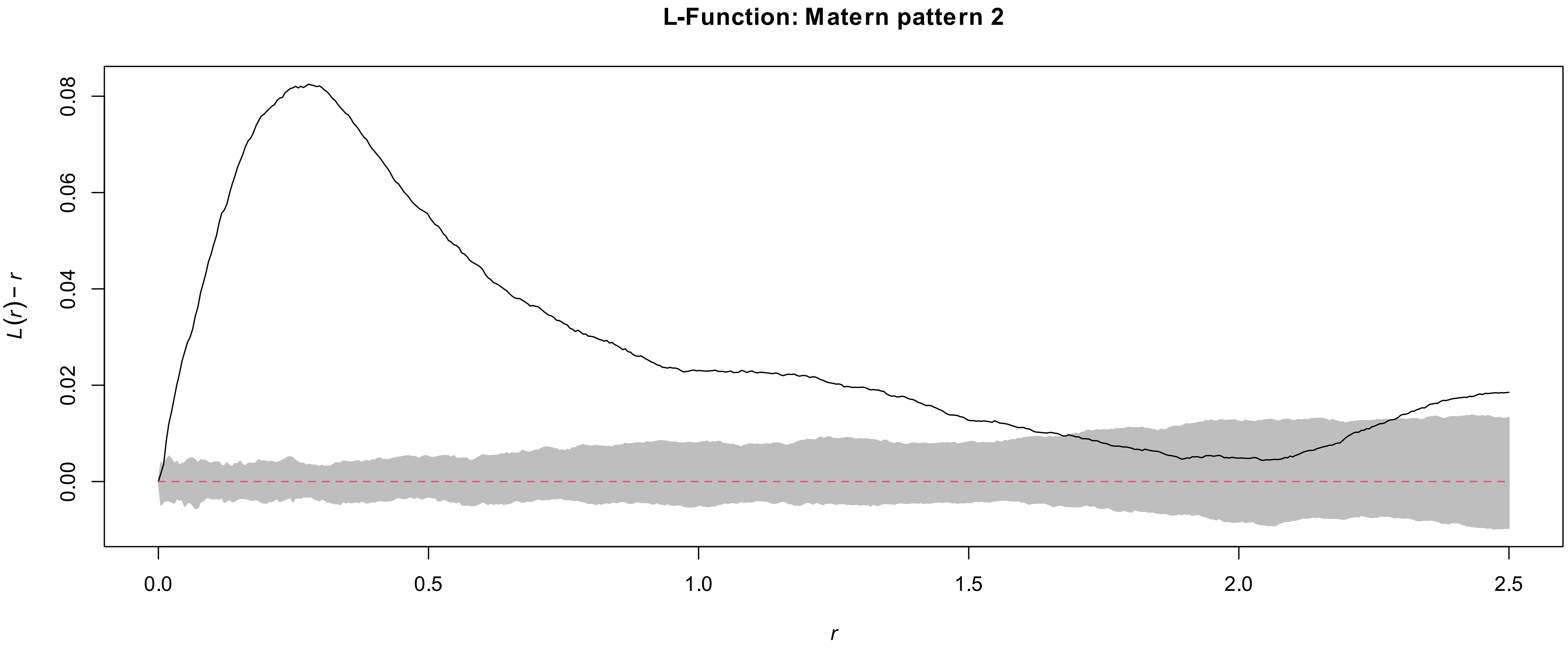


Matern Pattern

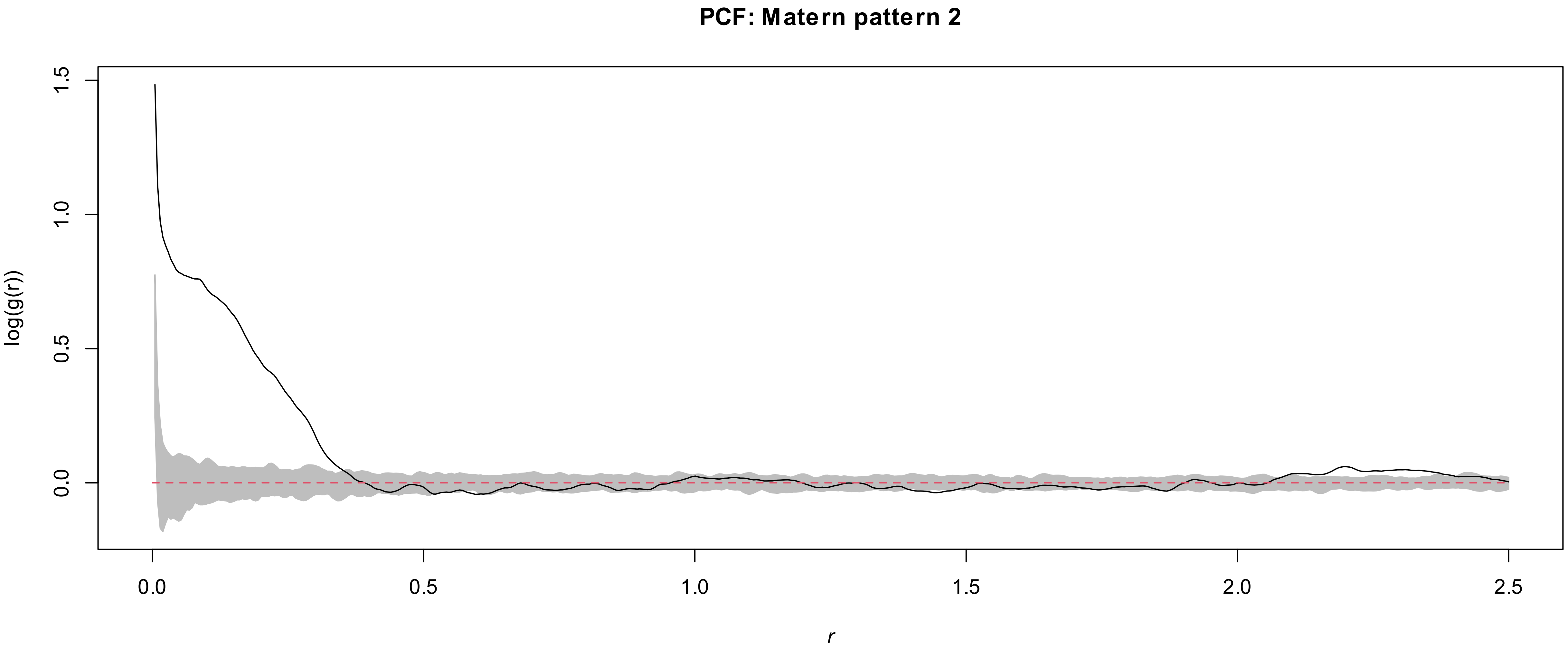




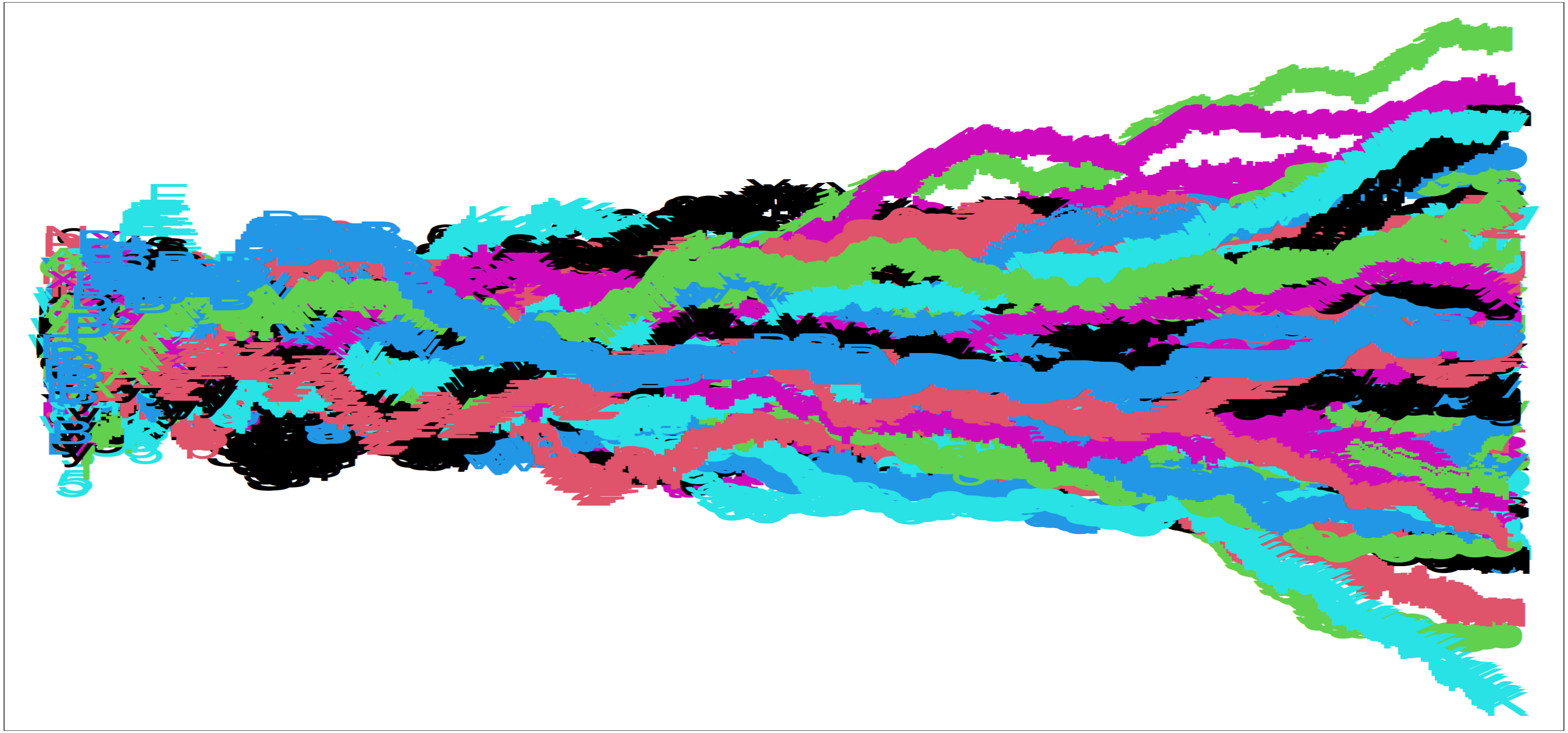
L-function: Matern



G-function: Matern



A Cool Mistake



Interpolation: Inverse Distance

Inverse Distance Weighting Interpolation

A slight detour into chapter 5 material which will leadl to a discussion of marked points.

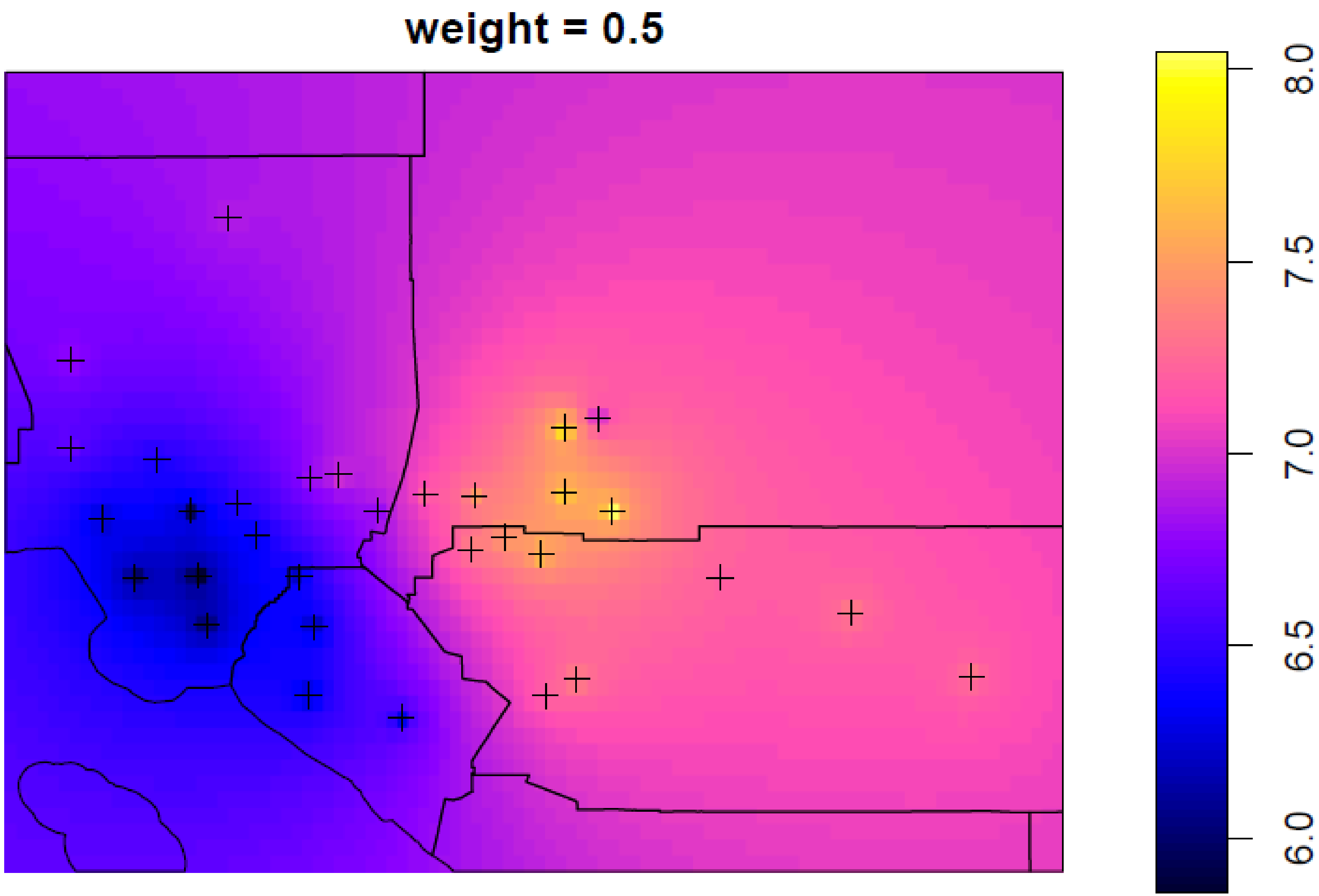
Interpolation: filling in the gaps for areas in which we have no data.

We'll discuss several techniques in this course

- Inverse Distance Weighting (IDW)
  - A non-statistical method that uses a weight matrix (derived from a distance matrix)
  - Allows for weighting the distance of influence of nearby data points.
- Kriging
  - A statistical technique that also estimates uncertainty

Inverse Distance Weighting

- IDW is a simple, non-statistical interpolation method.
- It calculates unknown points as a weighted average of nearby points.
- Weight parameter determines how far the influence of each point extends.



IDW

Template Raster

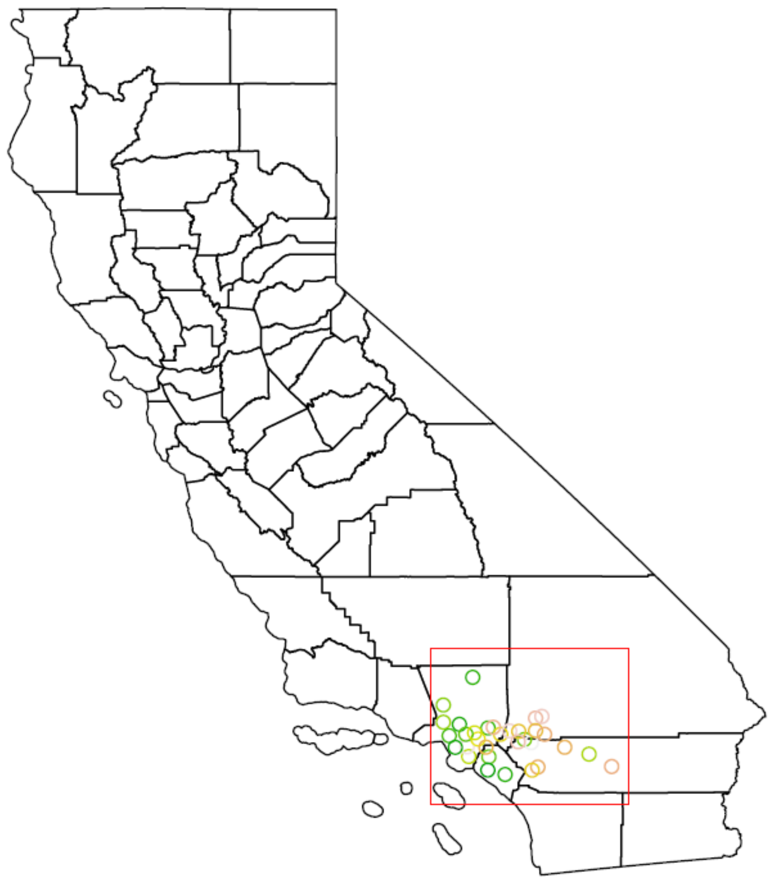
```
# Template Raster
nrows = 50; ncols = 100
ozoneTemplateRaster =
  raster(
    nrow = nrows,
    ncol = ncols,
    ext = extent(border))
ozoneTemplateRaster[, ] = 0
ozoneGrid = as(
  ozoneTemplateRaster,
  'SpatialGridDataFrame')
```

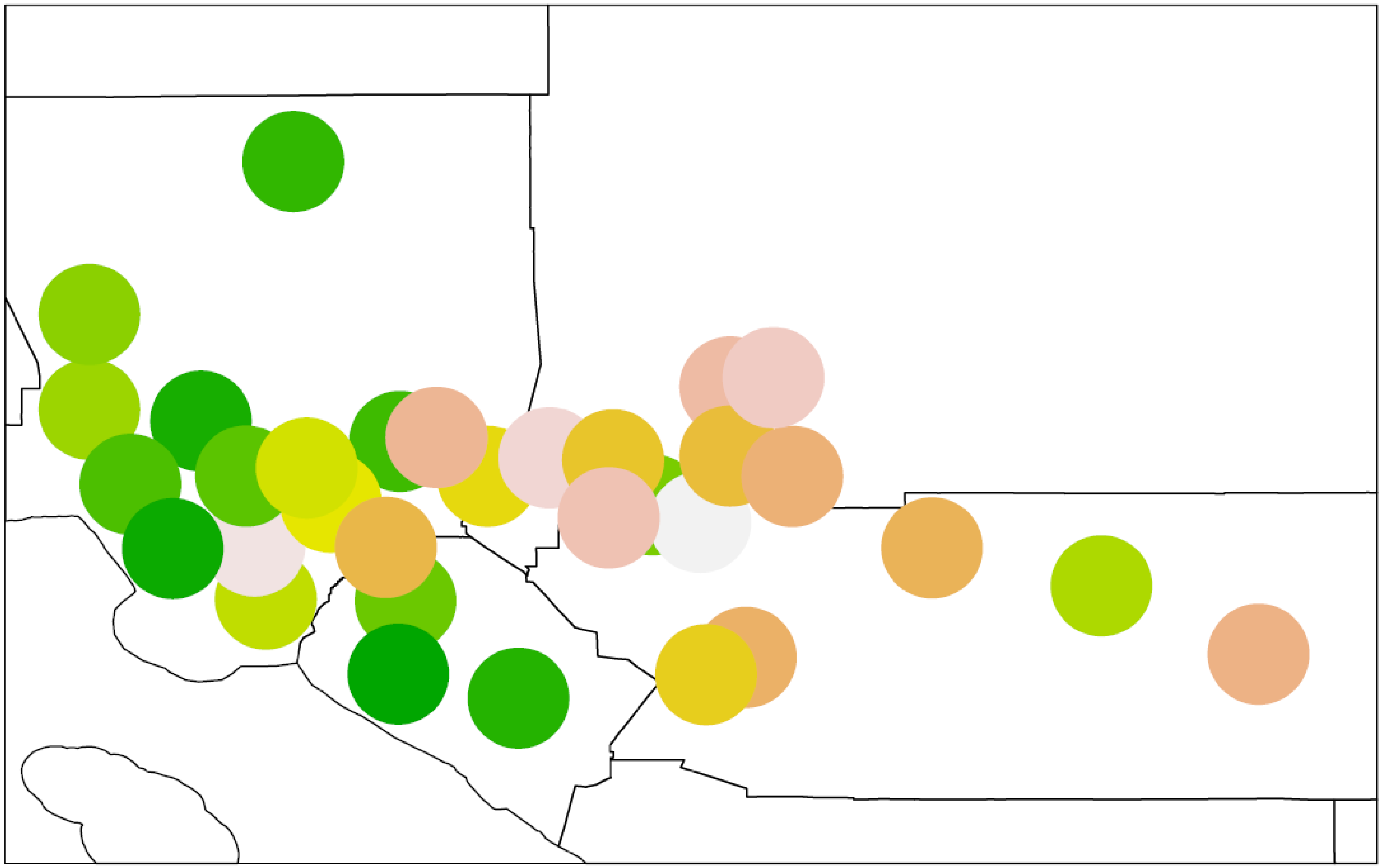
Nearby points are more alike than points further away

IDW requires:

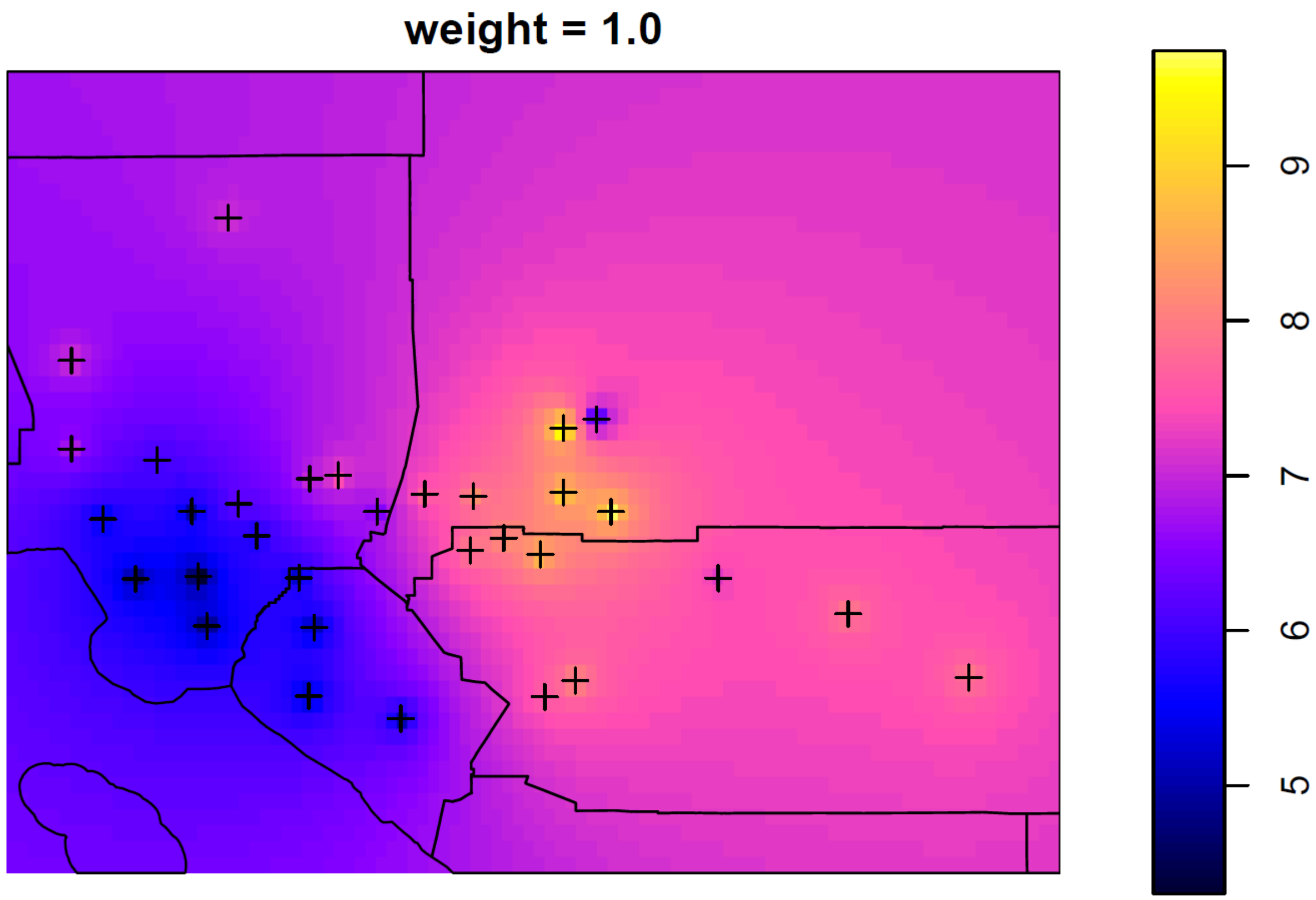
- A set of points (with coordinates) of known values
- A set of points at which to interpolate values
- A weight parameter: this determines the rate of decay of influence.
- A distance matrix: take the reciprocal of the values to get the **inverse distance**
- Apply the distance power parameter: It’s an exponent:
  - b = 1 corresponds to inverse distance
  - b = 2 corresponds to the square of inverse distance

Inverse Distance Weighting Example: LA Ozone

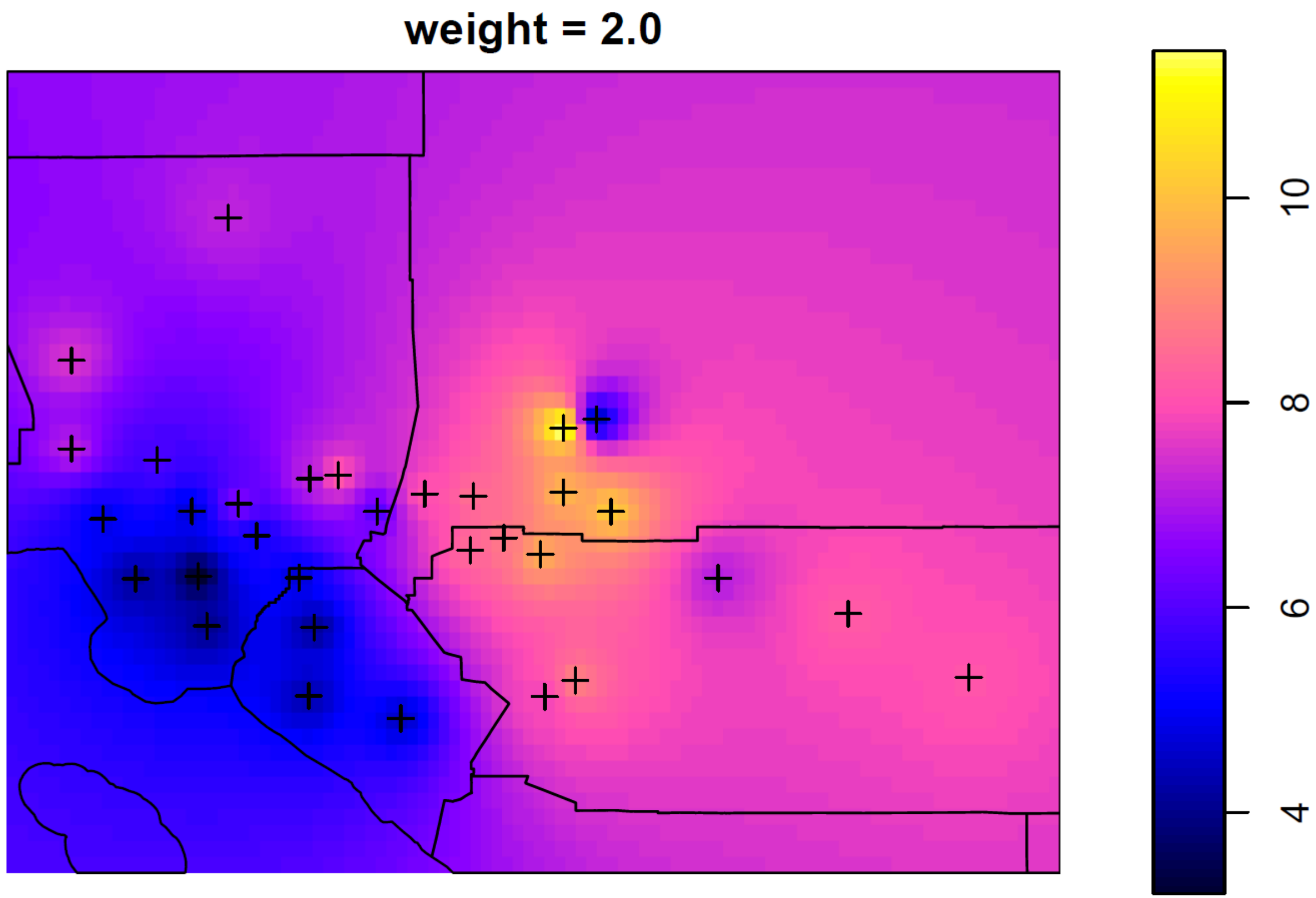




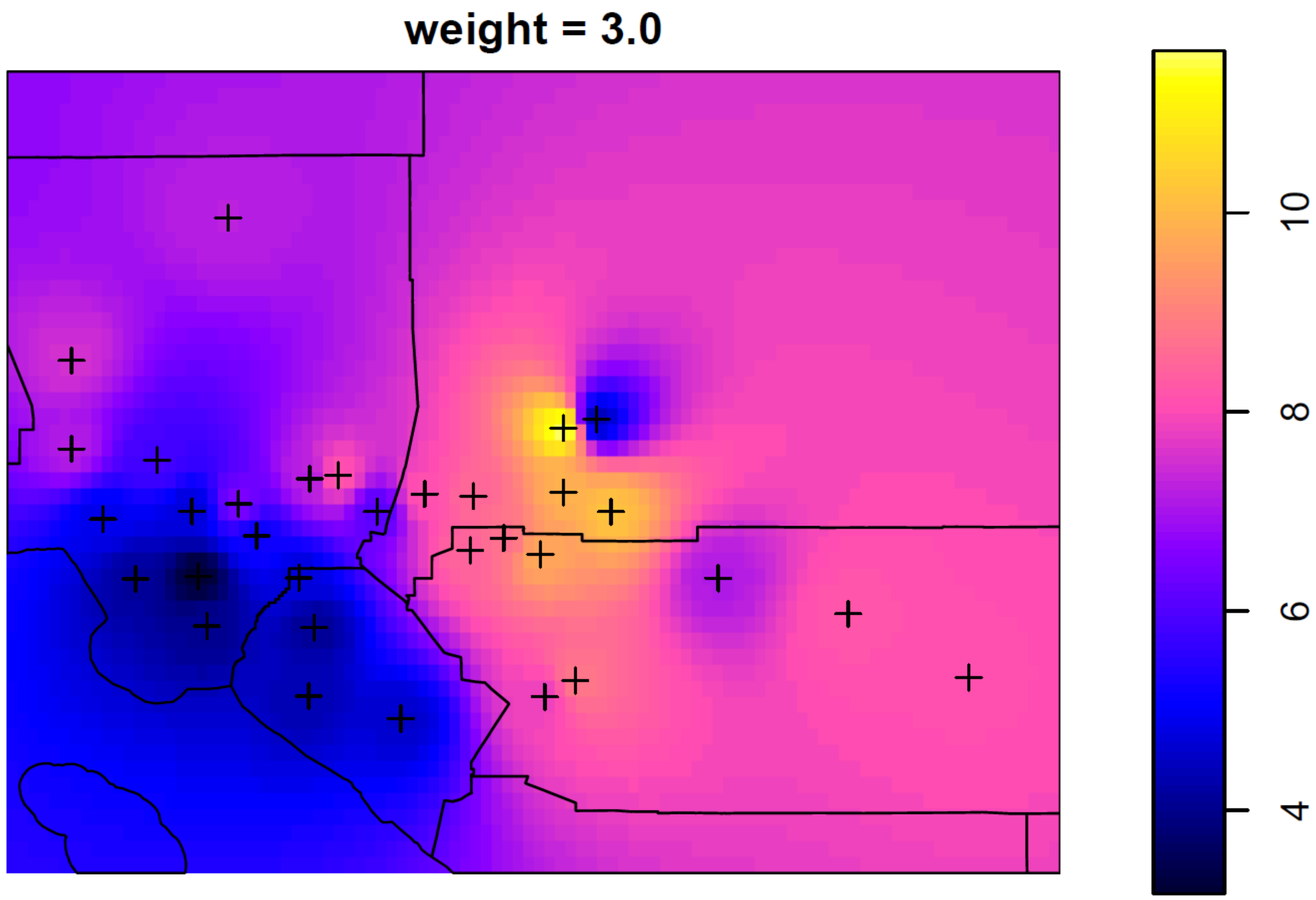
IDW: Example



IDW: Example



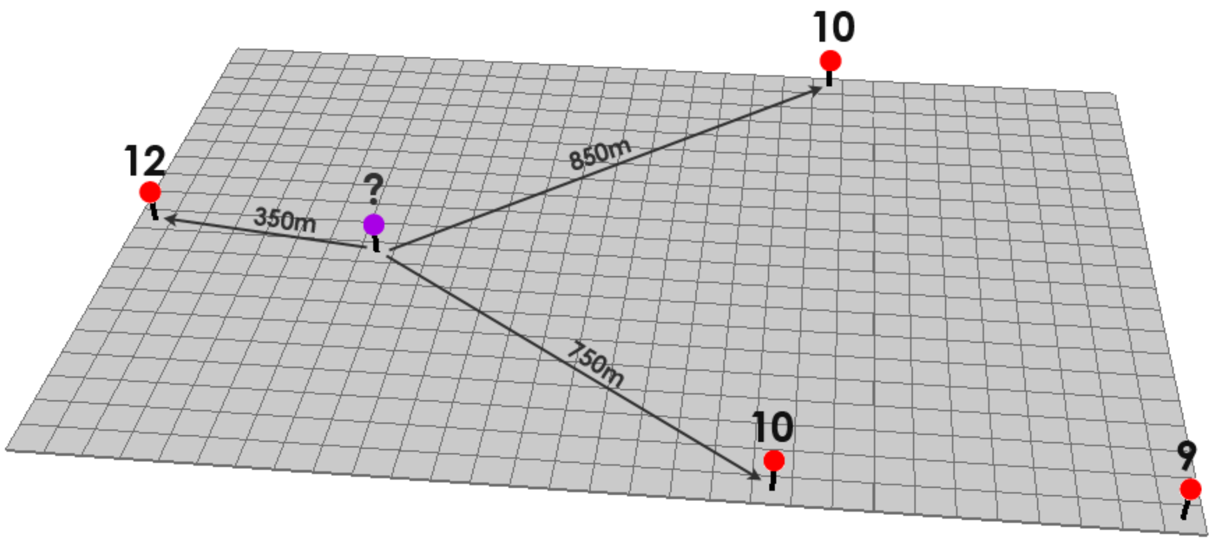
IDW: Example



Example from GIS Geography

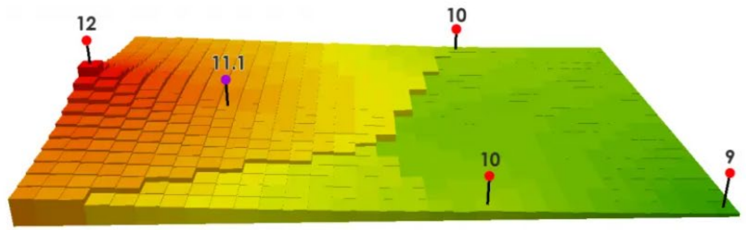
https://gisgeography.com/inverse-distance-weighting-idw-interpolation/

Value of unknown point determined by the three nearest known values



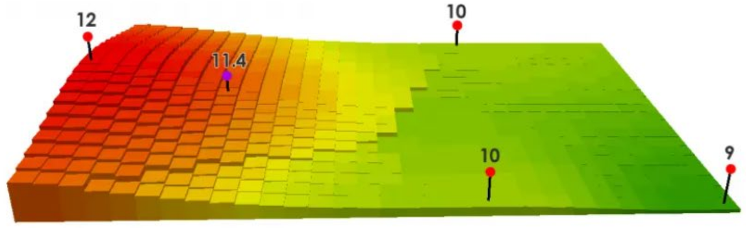
For a **power of 1**, that cell value is equal to:

((12/350) + (10/750) + (10/850)) / ((1/350) + (1/750) + (1/850)) = 11.1



For a **power of 2**, that cell value is equal to:

$= ((12/350^2) + (10/750^2) + (10/850^2)) / ((1/350^2) + (1/750^2) + (1/850^2)) = 11.4$



IDW Things To Note

- Interpolated points that are far form data points approach the global average
- High weighting powers increase the local influence: faster decay
- Low weighting powers reduce the local influence, slower decay
- High weighting powers mean greater smoothing
- Low weighting means values quickly decay to the global average