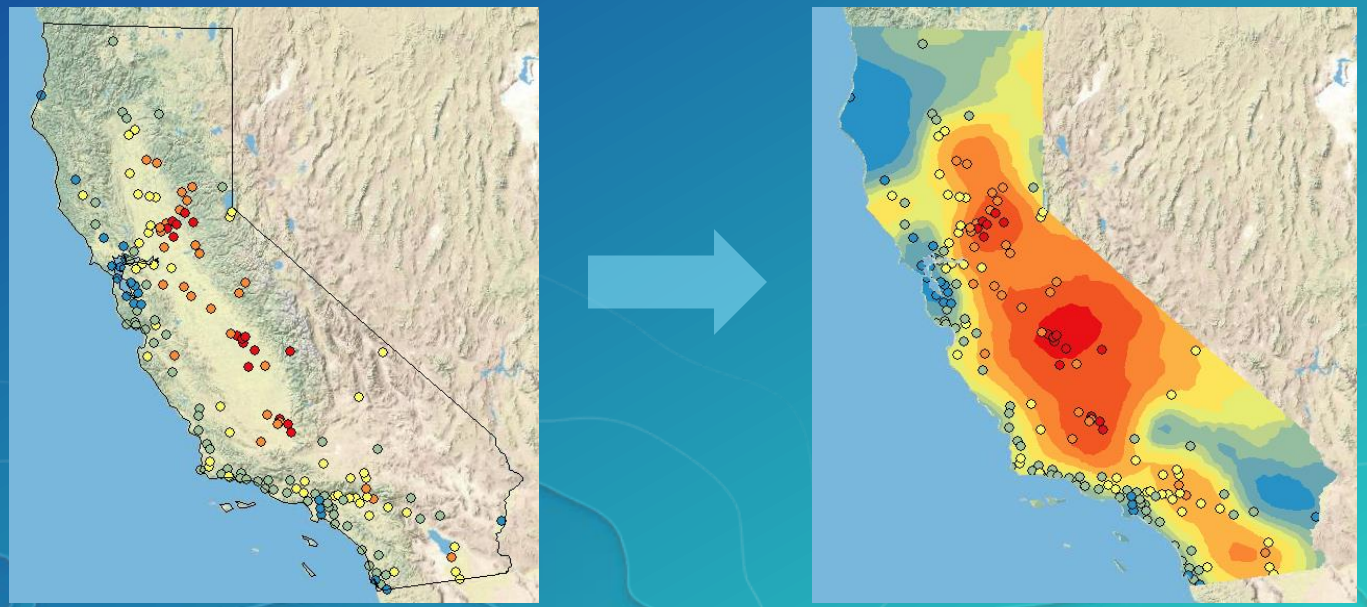
# Geostatistics: Learning kriging through examples

Linda Beale, Konstantin Krivoruchko

# What is interpolation?

- Predict values at unknown locations using values at measured locations
- Assumes spatial autocorrelation
- Many interpolation methods, both statistical and non-statistical



# Why are geostatistical methods important?

- **Non-statistical interpolation methods**
  - Not based in statistical theory
  - Not able to estimate prediction error, unclear assumptions
    - Examples: IDW, Spline, Natural Neighbor, Trend

- **Geostatistical methods**
  - Predictions based on statistical principles and theory
  - Clear assumptions that can be checked
  - Provide measures of uncertainty for predictions

# Kriging

Uses the relationships between your data locations and their values, assuming:

- Data is normally distributed
- Data exhibits stationary (no local variation)
- Data has spatial autocorrelation
- Data is not clustered
  - simple kriging has declustering options
- Data has no local trends
  - local trends can be removed during interpolation (and these trends are accounted for in the prediction calculations)

# How does it work?

Assumes that spatial variation can be decomposed into 3 main components:

1. Deterministic variation or trend/drift

   Trend analysed by trend surface analysis techniques

2. Spatially correlated, random variation

   Spatially correlated variation analysed by computing the semi-variance

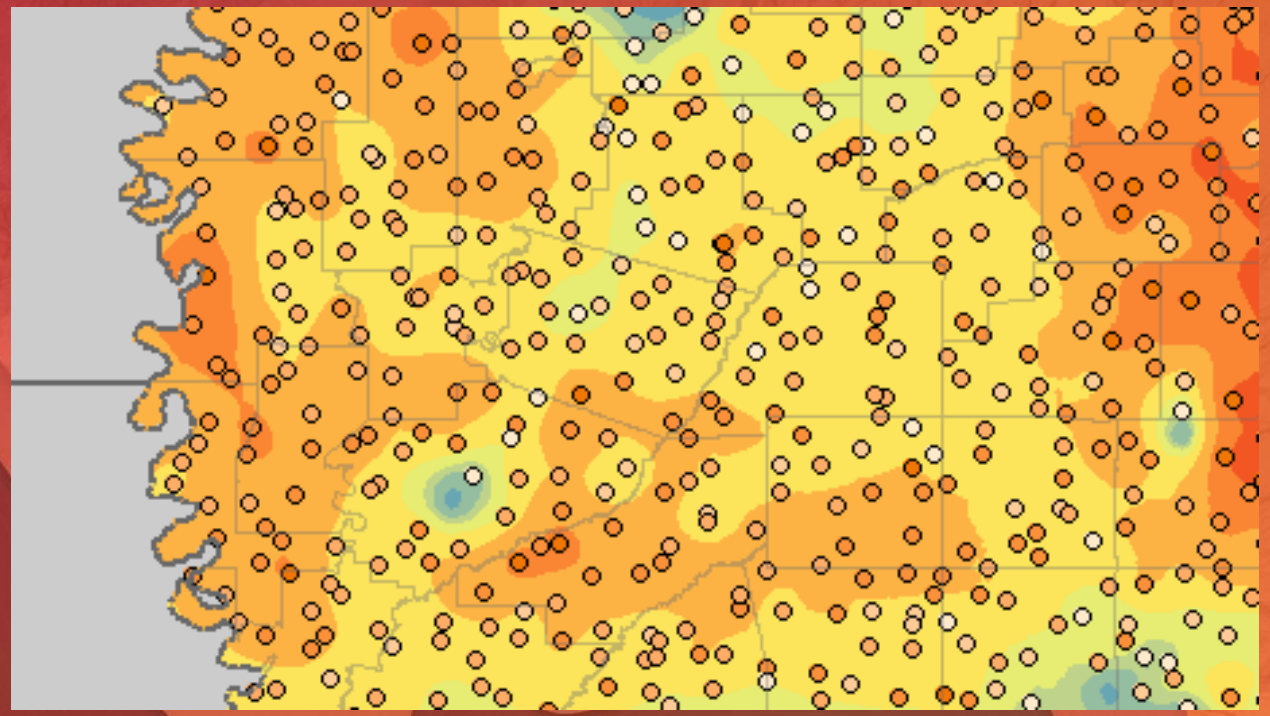
3. Spatially uncorrelated variation (noise)

   Provides measures of the certainty or accuracy of the predictions

# Exploring the data

## When is kriging optimal?

- **Assumptions for kriging**
  - **Normally distributed data** ⇒ **Transformations**
  - **Stationary** ⇒ **Autocorrelation**
  - **No trends** ⇒ **De-trending options**
  - **Not too clustered** ⇒ **De-clustering techniques**



# Explore Data

Understanding your data

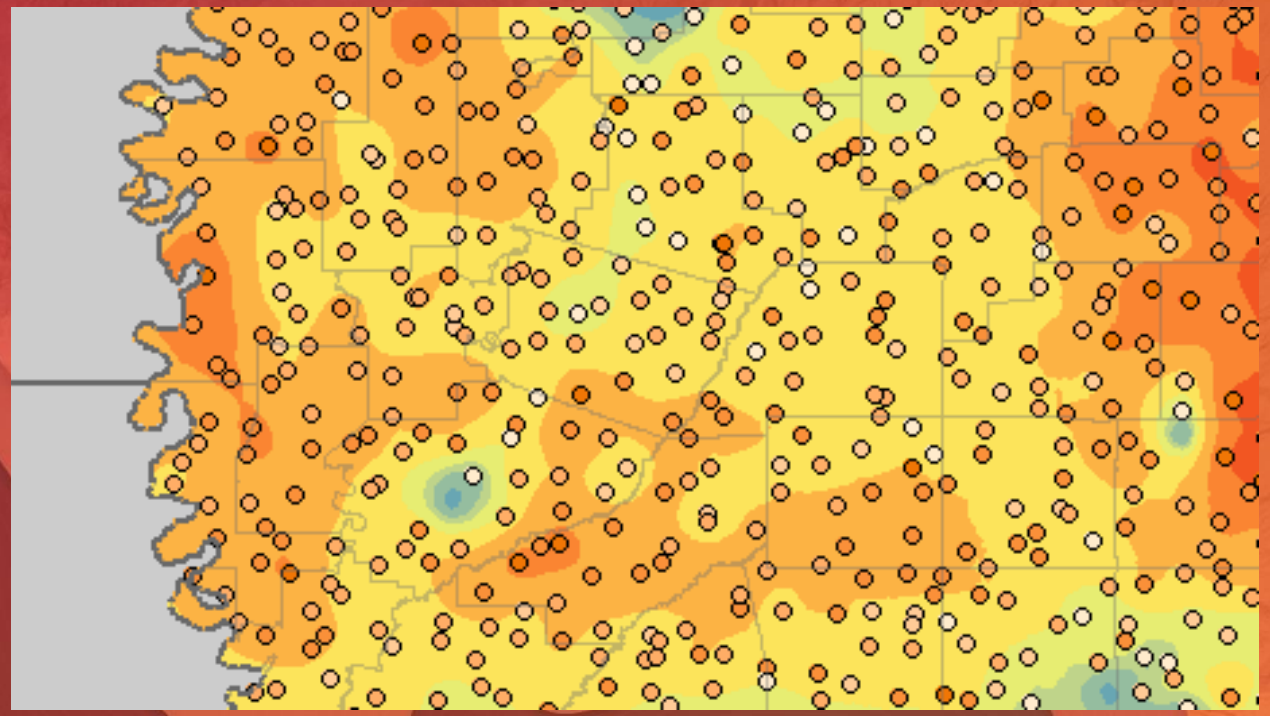
# Types of kriging models

## Assumptions differ by type

- **Ordinary Kriging**
  - Assumes the constant mean is unknown and the data have no trend
- **Simple Kriging**
  - Assumes a constant but known mean value - more powerful than ordinary kriging
- **Universal Kriging**
  - Assumes that there is an overriding trend in the data
- **Indicator Kriging**
  - Uses thresholds to create binary data and then uses ordinary kriging for this indicator data
- **Probability Kriging**
  - Strives to do the same thing as indicator kriging, but it uses cokriging to try to improve results
- **Disjunctive Kriging**
  - Tries to do more than ordinary kriging but many assumptions must be met

# Types of kriging

- Empirical Bayesian Kriging
  - Accounts for the error in estimating the underlying semivariogram through repeated simulations
  - EBK Regression Prediction
- Areal Interpolation
  - Extends kriging theory to data averaged or aggregated over polygons



# Kriging in Action

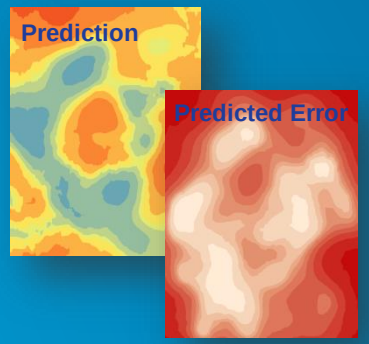
EBK, EBK Regression & Areal

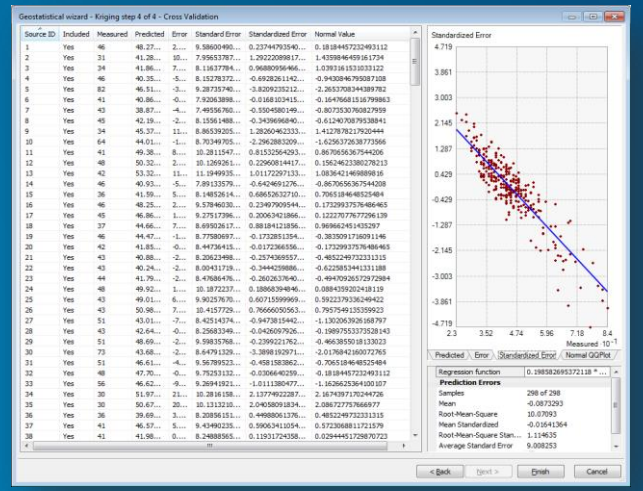
# Selecting the best model

- Predictions should be unbiased
  - Mean prediction error should be near zero (depends on the scale of the data, so,
  - Standardized mean should be nearest to zero
- Predictions should be close to known values
  - Small root mean predictions errors
- Correctly assessing the variability
  - Average standard error nearest the root-mean-square prediction error
  - Standardized root-mean-squaere prediction error nearest to one

# Validate results

- **Cross validation**
- **Create surfaces using a data subset and use the remainder data to validate results**
- **Use predicted error together with predicted surfaces**



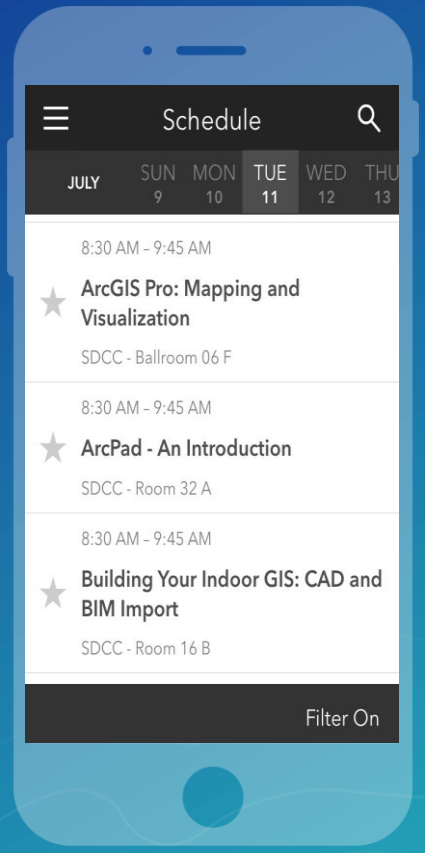



Cross validation does not prove that the model is correct, merely that it is not grossly incorrect (Cressie, 1990)

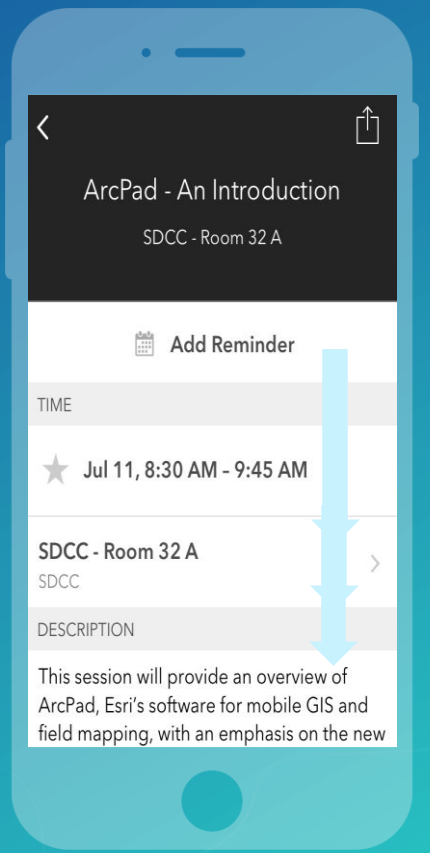
# Please Take Our Survey on the **Esri Events App!**

## Download the Esri Events app and find your event

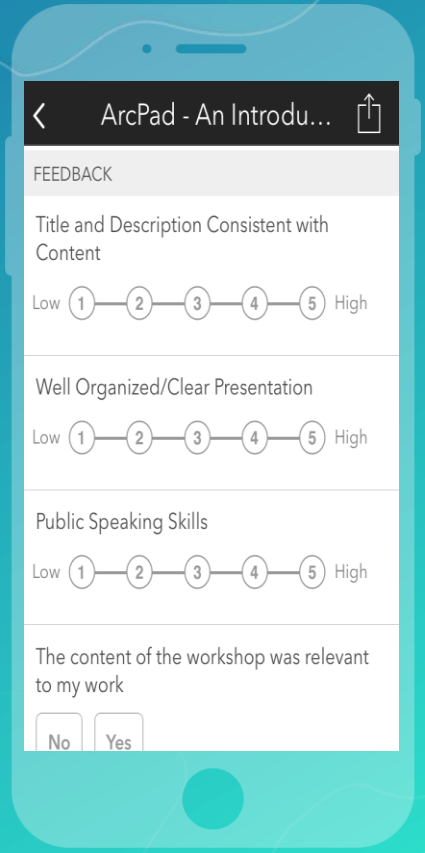

## Select the session you attended




## Scroll down to find the survey




## Complete Answers and Select "Submit"




esri
THE
SCIENCE
OF
WHERE