Esri International User Conference | San Diego, CA

Technical Workshops | July 2011

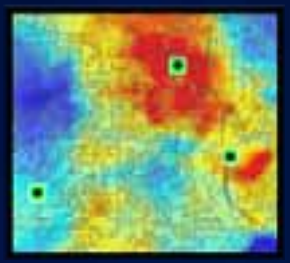
Spatial Statistics: Best Practices

Lauren Rosenshein, MS

Lauren M. Scott, PhD

Workshop Objectives

- Demonstrate an analytical workflow, start to finish
  - 911 Emergency Call Data Analysis



- Dig deeper:
  - Understand tool parameter options
  - Select an appropriate spatial scale for your analysis
  - Explore data relationships
  - Find a properly specified regression model

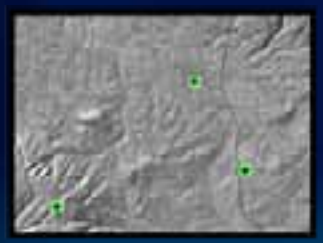
- Point you to additional resources for learning more

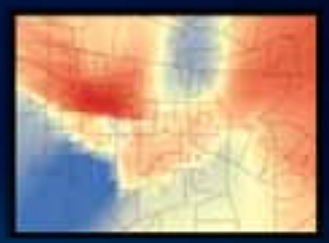
www.esriurl.com/spatialstats

Kindly complete a course evaluation: www.esri.com/sessionevals

Context for the Analysis

- Responding to 911 calls is expensive
- The population is expected to double
- This community has questions!
  - Are existing fire and police resources well sited?
  - What are the factors that contribute to high 911 call rates?
  - What can be done to reduce 911 call volumes?
  - Given population increases, what kinds of call volumes can we expect in the future?

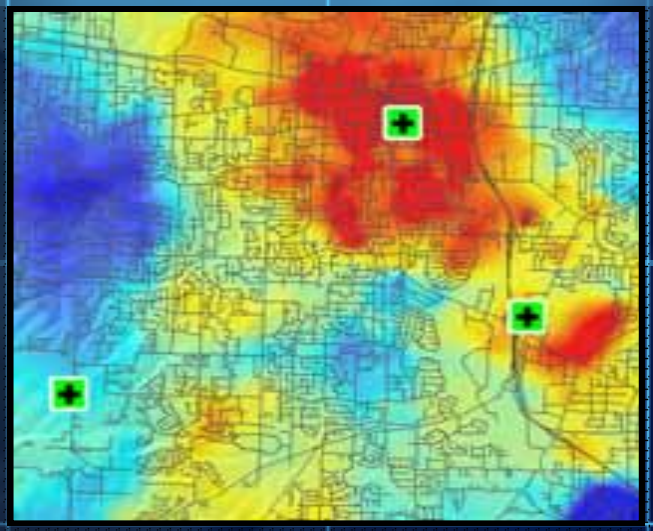


Hot Spot Analysis

- Works by assessing the *values* for each feature within the context of neighboring feature values
- Challenges:
  - What is the analysis field?
    - With incident data, you often need to aggregate the data
  - What is the scale of your analysis?
- Tools:
  - To aggregate your data
    - Integrate
    - Collect Events
  - To find an appropriate scale of analysis
    - Incremental Spatial Autocorrelation (sample script at 10.0)
    - Calculate Distance Band from Neighbor Count

Hot Spot Analysis

Demo



Regression Analysis

- All about answering "why?" questions
  - Ordinary Least Squares (OLS) creates an equation relating your dependent variable (calls) to a set of explanatory variables (population, income, …)

Calls = f(Population?, Income?, Education Levels?, Traffic?, Businesses?)

- Challenge
  - Finding a "properly specified model" (a model you can trust …because it meets all of the assumptions of OLS)
- Tool
  - Exploratory Regression

Did you find a good model?

- Check coefficient significance and sign
  - You want explanatory variables that are good predictors
    - Each explanatory variable coefficient should:
      - Have the expected sign (the expected relationship)
      - Have an asterisk (*) indicating the variable is statistically significant
- Check VIF values for Multicollinearity
  - You want a stable model: no explanatory variable redundancy
    - Variance Inflation Factor (VIF) values should be < 7.5
- Run the Spatial Autocorrelation tool on model residuals
  - You want to be able to trust variable relationships
    - Model over/under predictions should exhibit a random pattern

Did you find a good model?

- Check the Jarque-Bera diagnostic for model bias
  - You want your model to be consistent for the full range of values and throughout the study area
    - Model over/under predictions should be normally distributed

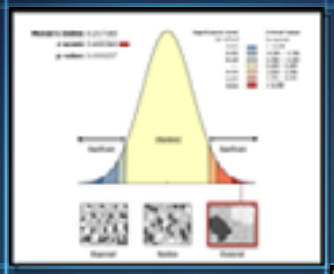
- Check model performance
  - You want a model that effectively explains 911 call volumes
    - Look for models with:
      - Largest Adjusted R^2 value
      - Smallest AICc value

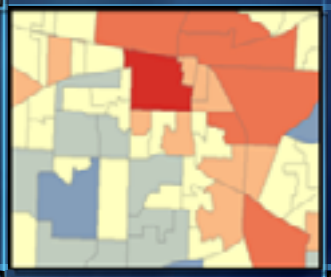
The Wiggle Clause

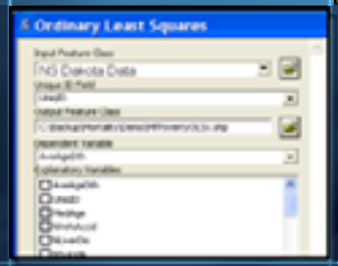
- Exploratory Regression can help you find a properly specified OLS model
  - Tries every combination for a set of explanatory variables
- There is a trade-off:
  - You will learn a lot about your data and about relationships among your variables.
  - You increase your risk for Type I statistical error
  - More likely to get a model that over fits your data
- Best practices:
  - Select your variables carefully
  - Consult your common sense often
  - Validate your final model

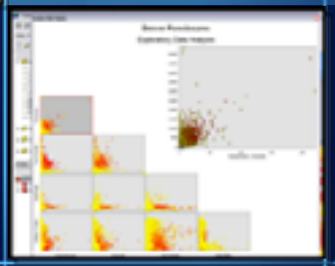
Regression Analysis

Demo



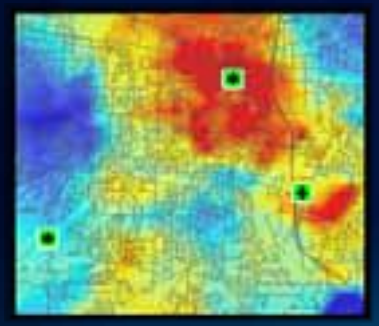






Analysis Results

- **Hot Spot Analysis:** are fire and police stations located well?
- **OLS:** what are the key factors promoting 911 calls?
- **GWR:**
  - where might remediation be most effective?
  - what will call volumes be like in the future given the anticipated growth?
  - are remediation methods effective?



Resources for learning more…

- **Spatial Pattern Analysis: Mapping Trends and Clusters**
  - Tue 8:30 Rm 2; Wed 3:15 Rm 2
- **Modeling Spatial Relationships using Regression Analysis**
  - Tue 10:15 Rm 2; Thu 1:30 Rm 1A/B
- **Spatial Statistics: Best Practices**
  - Tue 3:15 Rm 2; Thu 3:15 Rm 1A/B
- **Using R in ArcGIS**
  - Wed 12:00 Rm 1A/B
- **Road Ahead: Sharing of Analysis (ArcGIS 10.1)**
  - Wed 11:05 Rm 6B

Resources for learning more…

QUESTIONS?

- **www.esriurl.com/spatialstats**
- Short videos
- Articles and blogs
- Online documentation
- Supplementary model and script tools
- Hot Spot, Regression, and ModelBuilder tutorials

LRosenshein@Esri.com
LScott@Esri.com

Kindly complete a course evaluation: www.esri.com/sessionevals