

Modeling Protest in Lafayette Square: An agent-based modeling approach

Thomas Booth

Lead Geospatial Analytics, LMI

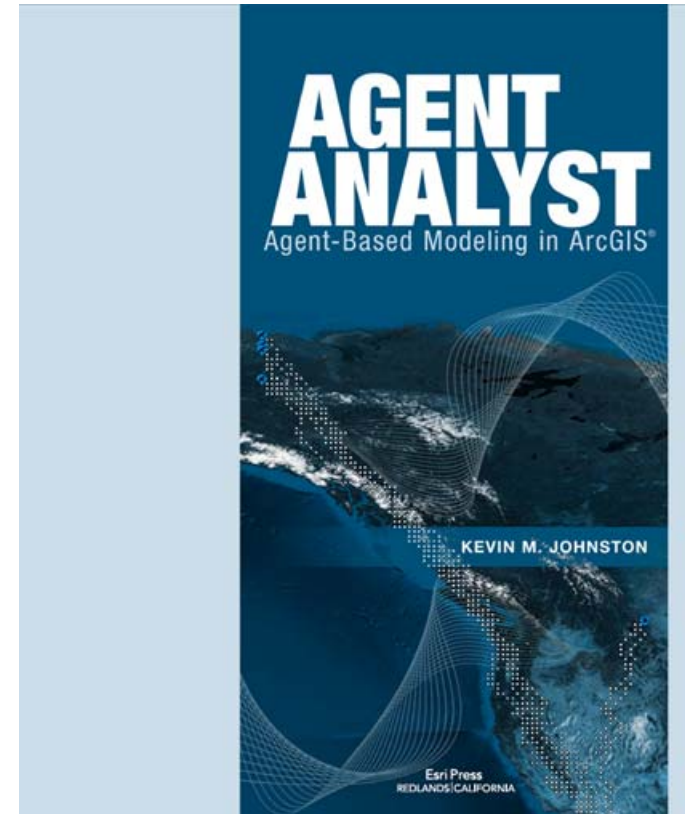
LMI

Agent-Based Modeling Overview

- Agent based modeling is the science of building computational models simulating the interactions of autonomous agents constrained by space and time while measuring the collective effect on a system as a whole.
- Breaking it down
 1. Agents can be any discrete object such as humans, vehicles, animals etc...
 2. Given rules, the agents interact with each other making decisions influenced by their immediate environment.
 3. Capturing agent interactions allows users to explore patterns and behaviors to study complex problems.

Agent-Based Modeling in Geographic Information Systems

- Agent Analyst is an agent based modeling extension built for ESRI's ArcGIS software suite.
- AA fully integrates the open-source REPAST modeling environment by using the programming language NQPython as an interpreter over its Java based engine.
- Importance: GIS experts can now model behavior in a familiar framework, utilizing real geographic data and apply complex spatial analytic processes and analysis to study behavior constrained by space and time.

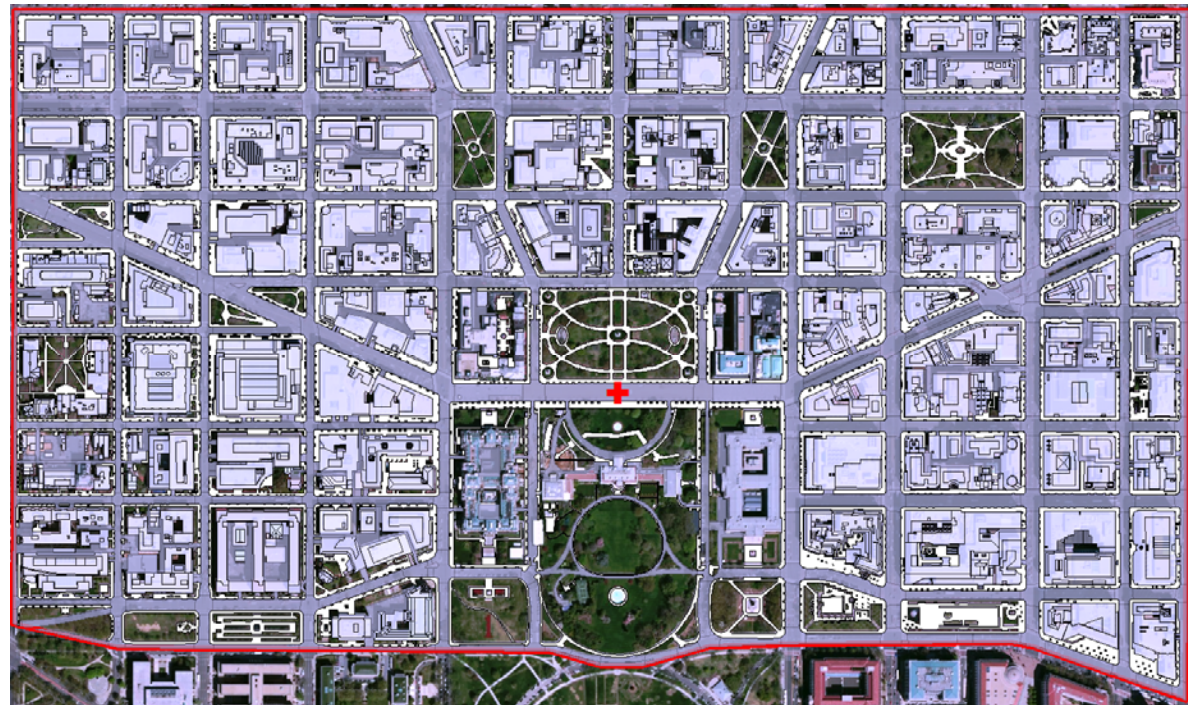


Protest Model in Agent Analyst

- Simplified re-creation of Joshua Epstein's model of Civil Unrest.
- Model a centralized protest on Lafayette Square against a central policing authority.
 - Primary focus on the relationship between protestors and cops.
 - Run experiments testing how changing parameters influence protestor cop behavior and model outcomes.
 - Extensively test the relationship between ABM and GIS.
 - Generate navigable raster datasets mimicking real-world walking paths throughout the project boundary.

DATA

- DCGIS
 - Street Centerline
 - Road Polygons
 - Sidewalk Polygons
 - Building Polygons (3D polygons)
 - Stair Polygons
 - Historic Sites (Parks)
- ESRI World Imagery
- Agents
 - Protestors
 - Cops

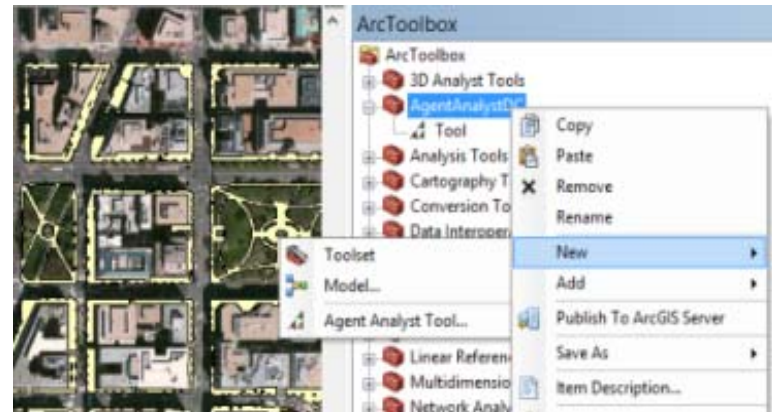


Building ABMs in Agent Analyst

- Agent Analyst uses the open source ABM environment Repast to integrate ABM within ArcGIS.
- Uses programming language called NQPython, or Not Quite Python.
- User friendly GUI, models are accessed through ArcToolbox.
- Agent Analyst is an out of process application.

Data Types:

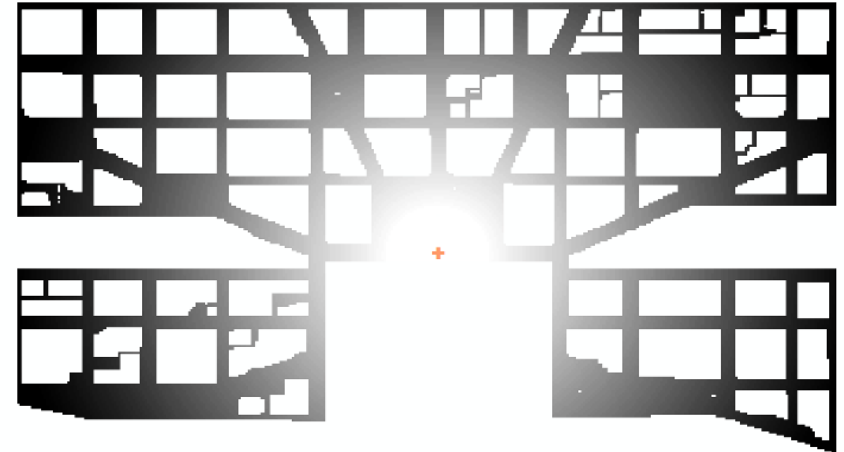
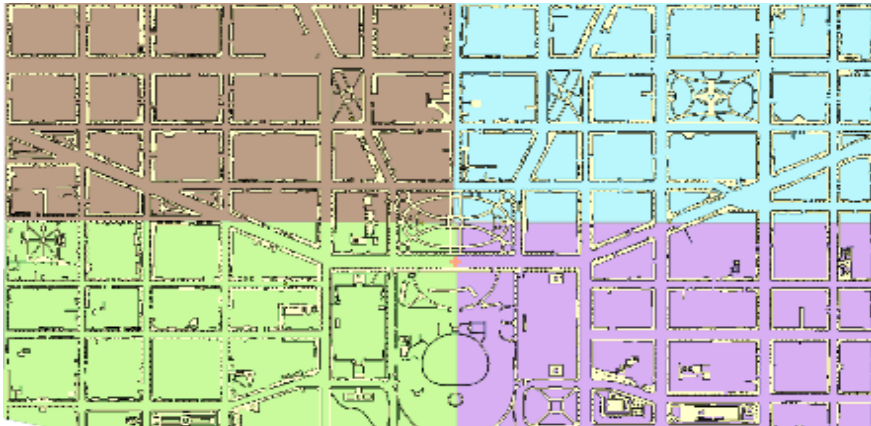
- Point, line and polygon shapefiles.
- Raster datasets*: euclidean distance, cost surfaces etc.



Protest Model Setup & Pre-Processing

Agent to Shapefile Interaction:

- Vector Agents: Protestors, police forces, and the central protest point are tied to three separate shapefiles. Each point in the shapefile represents one agent.
- Raster datasets: Euclidean distance datasets (bottom right) will serve as the primary navigation layer in the protest model. All the vector agents (protestors and cops) are bound to this layer for two reasons. First, the pixel values represent where they are related to the central protest point. Second, areas off the raster dataset indicate absolute barriers.
- Quadrant euclidean distance surfaces were generated to randomly establish starting locations for all agents at the model initiation. (bottom left)



Programming the Protest Model

Model Parameters

- Model Parameters drive the decision making of the agents within the framework of the model...
 - **Grievance Value:** At model initiation, all protestors are assigned a random grievance value from 0 to .20. This parameter defines the threshold at which a protestors turns from peaceful to violent protest.
 - **Arrest Threshold:** This parameter can be defined as a proxy variable for risk assessment. If a cops arrests exceeds the arrest threshold, then protestors will evade.
 - Cop/Protestor Vision: The vision parameter specifies the distance in meters that protestors and cops can “see” their surroundings.

And many more...

Agent Parameters

- The most important variable outside of an agents own grievance value is its protest level. Each protestor is either peaceful, active, arrested or evading at all times.

Capturing Protest Realism: Tasks for agents at each model iteration

- Step 1: Identify all surrounding cells and identify which cell is the closest to the protest center.
- Step 2: Using your vision, check to see if there are other protestors within your vision and get their grievance values.
- Step 3: Recalculate your own grievance value accordingly.
- Step 4: Check to see if there are any active cops who have exceeded the arrest threshold limit. If true, then evade, else keep moving toward protest.
- Step 5: Move most desired cell and then repeat.

Coding & Tracking Simulations

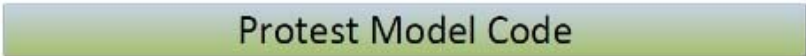
- In order to capture the essence of protestor behaviors, one must capture every aspect of the model to use data mining techniques to elicit common behavioral patterns.
- Some captured fields include... Activity level, grievance level, current location, distance to protest center, quad assigned to, time stamp, protestor density, protestor active density, average grievance, cop count, cop arrests etc...

Output Data Usage:

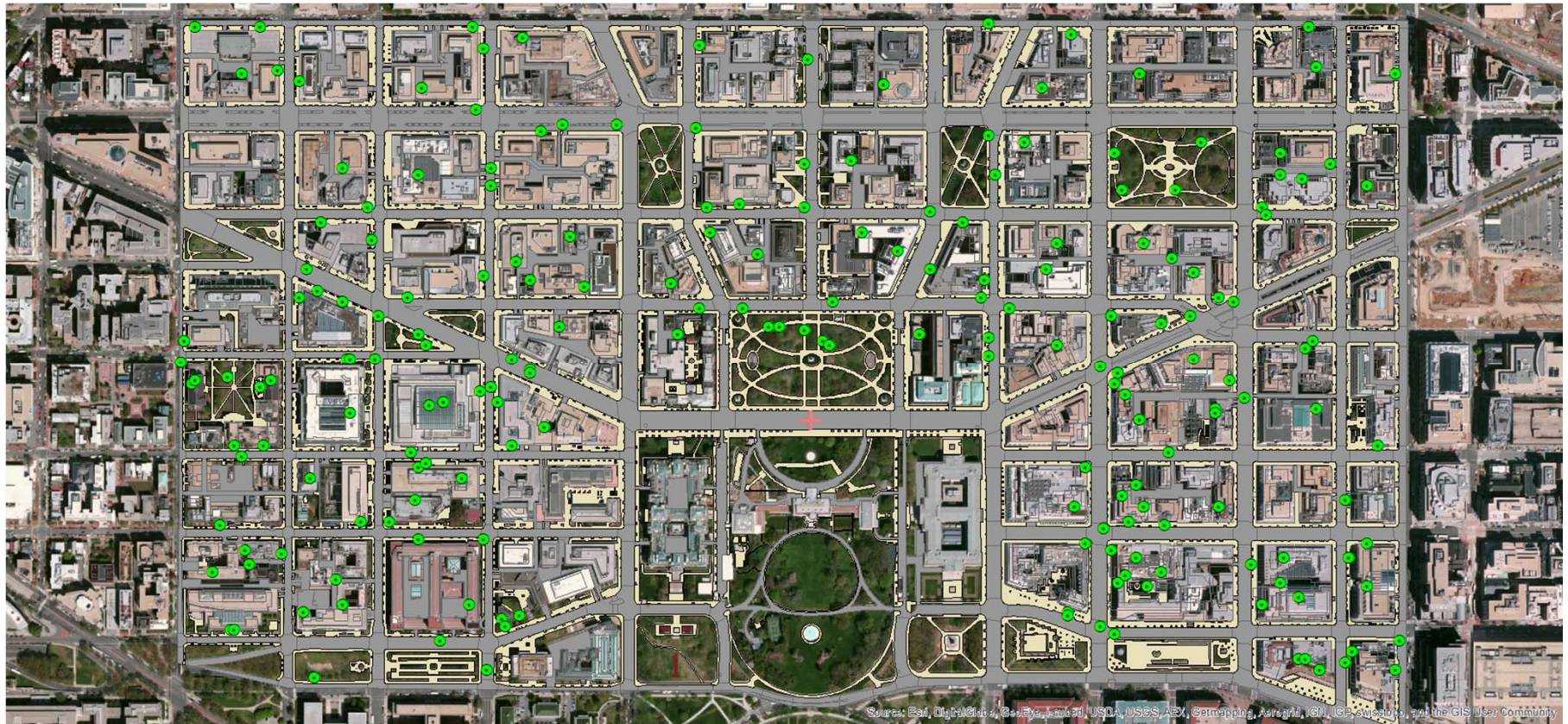
- Using ArcGIS's animation tool we can recreate the model simulations for further viewing analysis as well as time series analysis for all simulations at the same model steps.
- Perform different spatial statistical and analytical techniques (i.e. kernel densities, interpolation algorithms)
- Further statistical analysis in R. Use machine learning libraries to help us identify key variables in either deterring protests or inhibiting protests. Also run post-processed datasets through algorithms to predict which protestors will make it to active protest at the protest center. (Boruta, randomForest)

Tracking Data Stats

- Model simulations took 12 hours to complete.
- Each tracking shapefile had over 75k observations.



Simulation Recreated using Animation Tool



- Cops
 - Roaming
 - Actively Pursuing
- Protestors
 - Peaceful Protest
 - Active Protest
 - Arrested
 - Evading Arrest

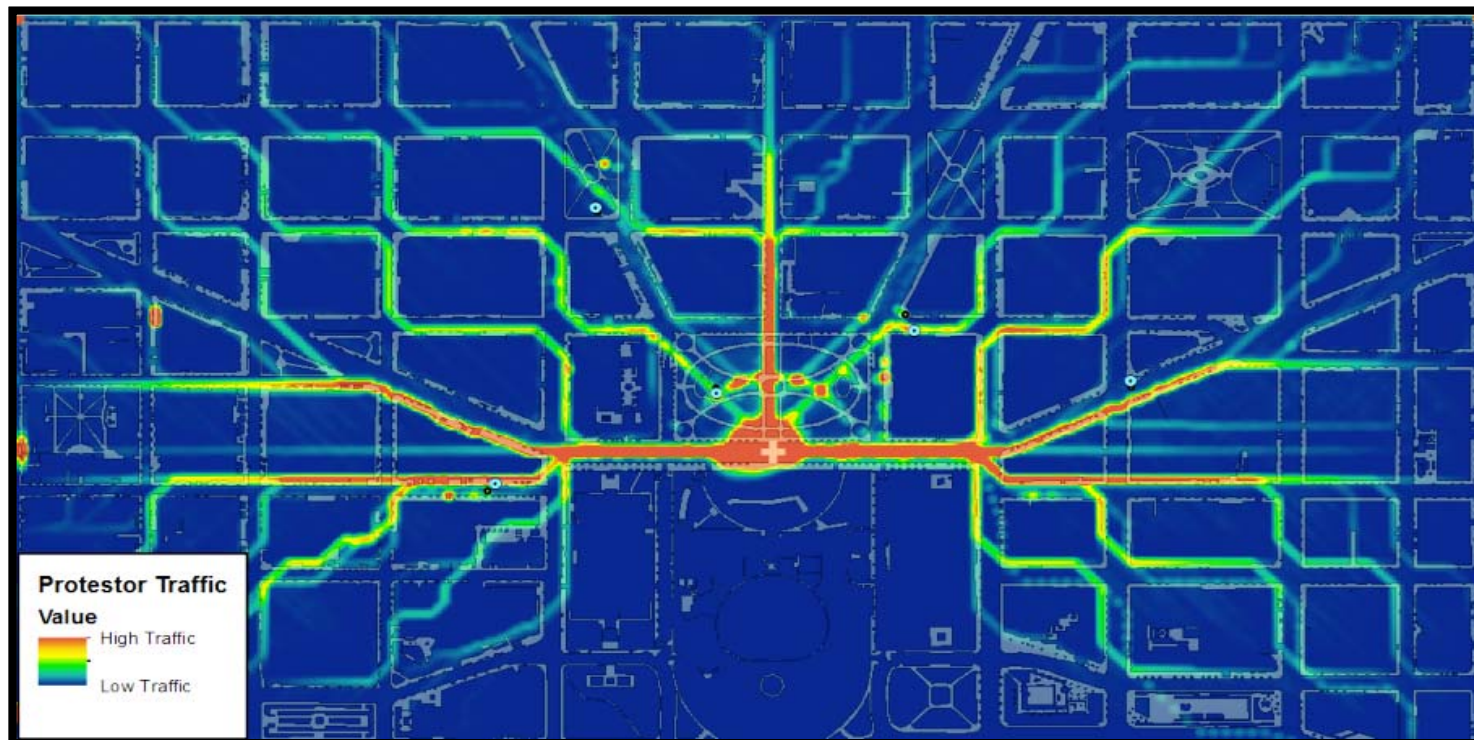
Post-Processing and Results (Kernel Density)

Kernel Density

- After compiling all simulations (1.5 million observations), a kernel density layer identifying the most heavily traveled routes by the protestors.

Purpose

- This dataset could be utilized by a central authority to place road blocks or to plan against an active protest.



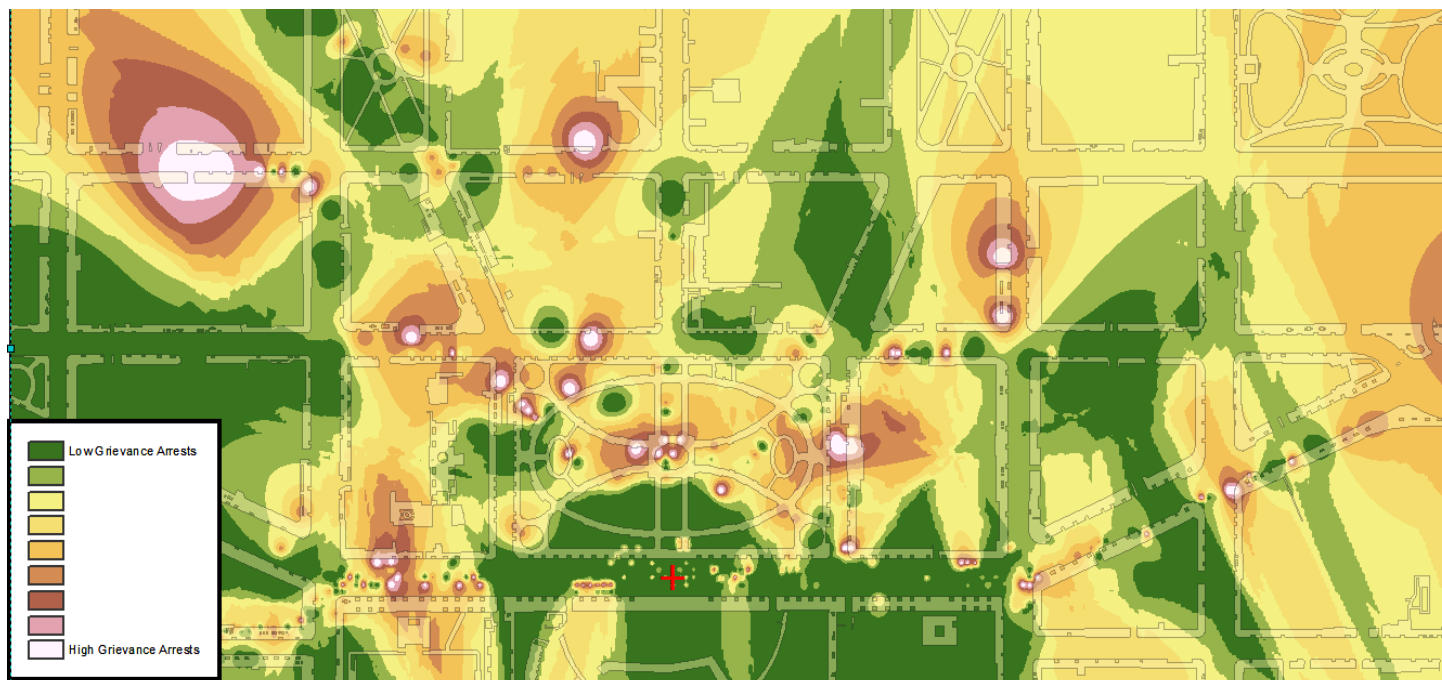
Post-Processing and Results (IDW)

Inverse Distance Weighted

- The IDW interpolation algorithm was utilized to identify the highest grievance value arrest areas in the project landscape. The figure below indicates grievance values at the time of arrest.

Purpose

- Again this dataset could be utilized by a central policing authority to key in on areas of the highest violence during a protest.



Arrest Animation Model



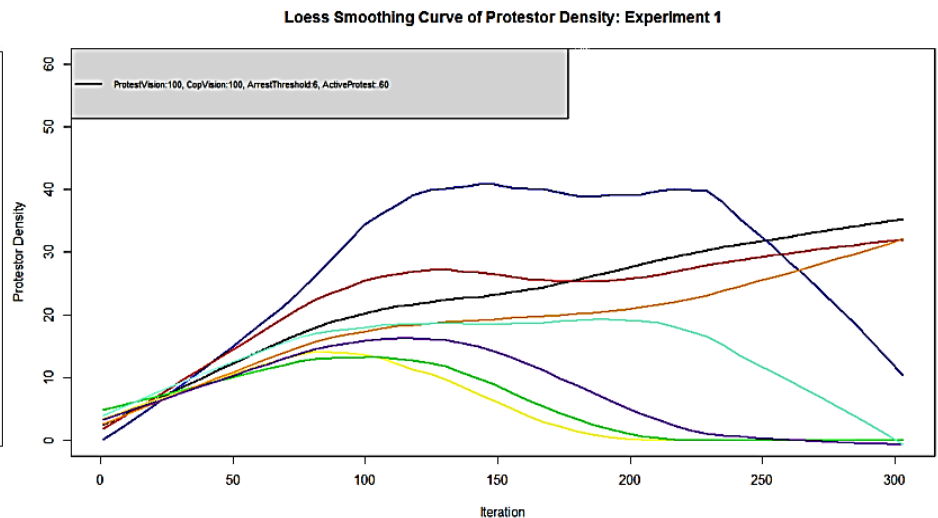
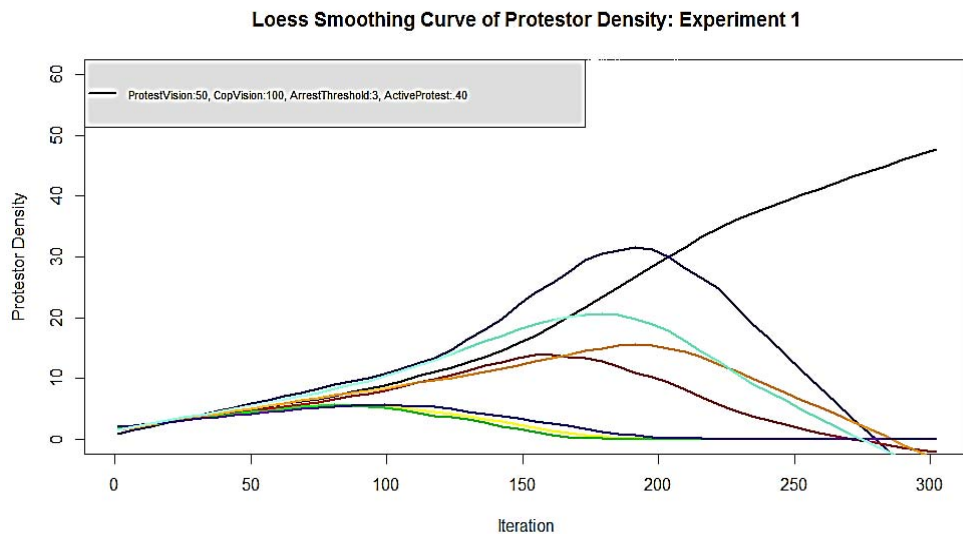
Loess Curves Exploratory Analysis

Loess Smoothing Curves

- In order to statistically gauge which experiment reached full protest, Loess Smoothing Curves were applied to each simulation exercise.
- Loess curves locally weight points closest to the permutation value, resulting in a curved fitted line to the data.

Result

- Experiment 2 reached full protest more often than its counterpart. Experiment 2 had a slightly higher arrest threshold and grievance threshold. (takes longer to get to full protest)



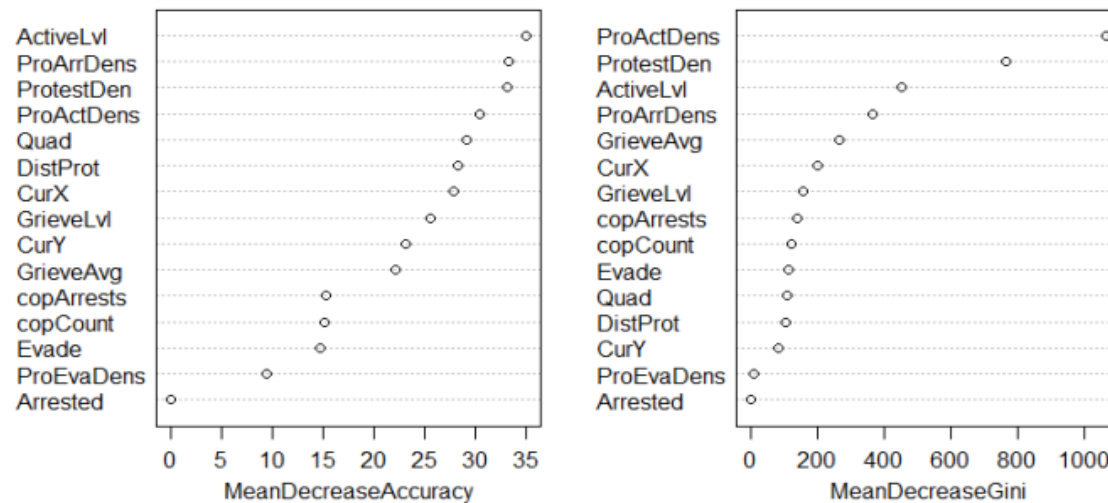
Random Forest Variable Importance

- Random Forest is an open source machine learning package in R. Its an advanced data classification algorithm that identifies key variables and can be used to model predictions on a dependent variable.

Result

- The two graphs below show variable importance when predicting which agents reach the protest center.
 - Mean Decrease Accuracy: Quantifies how much inclusion of each variable reduces classification error.
 - Mean Decrease Gini: Quantifies which variables play a larger role in splitting data throughout the course of the algorithm.

Random Forest Variable Importance Plot

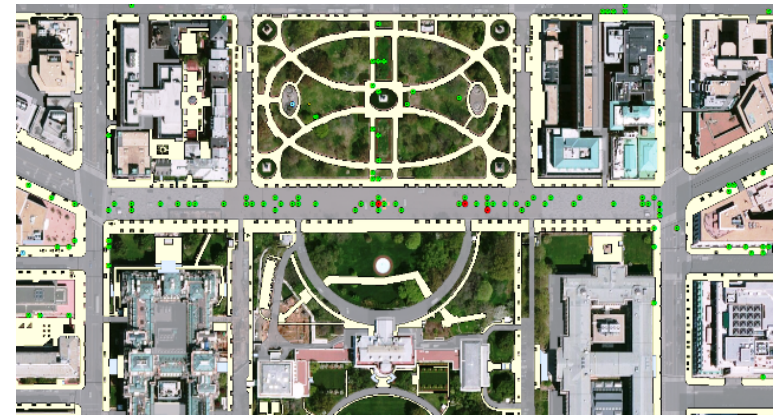


Strengths of ABM with Geospatial Analysis



Pre-Model Benefits

- Data creation (continuous surfaces, point, polygon, polyline)
- Pre-processing of datasets (editing to fit data requirements)
 - Raster and vector editing for high resolution and accurate datasets
- Streamline workflows (build repeatable models)
 - Python scripting
 - Model Builder



Post-Model Benefits

- Post-processing of model outcomes (making sense of the output data)
- Spatial Analysis
 - Spatial interpolation, density operations and techniques
 - Raster calculations, PCA Analysis (Map Algebra & Calculus operations)
- Spatial Statistics
 - Analyzing spatial autocorrelation, clusters and patterns
 - Regressions (OLS and GWR)

Strengths & Weaknesses of Agent Analyst

Strengths

- REPAST interface integrated into ArcGIS, the most widely used and respected geospatial software on the market.
- Captured data is immediately stored in feature classes for post-processing
 - Point, Line or Polygon shapefiles and feature classes can be written to.
- Model simulations can easily be replayed in ArcGIS's animation tool or using a third party animation software.
 - In ArcMap, ArcScene or ArcGlobe (for realistic three dimensional viewing)
- AA works great with GIS datasets.
 - Traditional Geometries: Point, Line and Polygon shapefiles
 - Elevation Surfaces: DEMs, Terrains and other elevation
 - Raster Surfaces: Euclidean Distances, Cost-Surfaces, Slope, Aspects and many others

Weaknesses

- Tracking and capturing model data is computationally expensive
 - Protest Model: 200 agents, a time step averages around 20 seconds.
 - Each simulation creates around 70,000 to 90,000 individual point features.
- Limited documentation and support community
 - The developers finally released a book with sample models, data and tutorials showcasing some basic and advanced functionality. (Kevin M. Johnston, 2013)
- Fairly steep learning curve for beginner and intermediate programmers
- AA is an out of process application
 - Agents can't generate new geoprocessed datasets "in-run"
- Experimentation flexibility

Agent Analyst Evaluation

- Agent Analyst has proven a powerful geospatial agent-based modeling tool.
- Provides GIS professionals a robust package to expand upon standard analysis techniques of locate, quantify and visualizing spatial relationships.
- Expands GIS horizons by modeling non-traditional problems such as behaviors constrained by space and time.
- ABM's fusion within a GIS environment enhances the science altogether with powerful analysis tools that ArcGIS provides in raster dataset generation and data modification.
- **For more information and to download Agent Analyst please visit...**

<http://resources.arcgis.com/en/help/agent-analyst/>

