Assessment and Evaluation of Individually Calibrated Journey-To-Crime Geographic Profiling Models

Michael Leitner
Department of Geography and Anthropology
Louisiana State University

Tania Pal
Department of Geography and Anthropology
Louisiana State University

Research Question 1

- Comparing the accuracy of Journey-To-Crime (JTC) Geographic Profiles (GP) created from
  - individually calibrated distance decay functions with
  - using the default values in CrimeStat (3.0)

Research Question 2

- Comparing the accuracy of individually calibrated JTC GP with alternative modern GP models (Rigel and Dragnet) and simple spatial distribution measures (spatial mean, spatial median, center of minimum distance).

- This research question is answered by comparing the results from this study with results from previous research.

Rationale

- If there were no difference, whether JTC GP are created from default or individually calibrated distance decay functions, then
  - Default parameter values should be used when creating JTC GP
  - Distance decay functions do not need to be individually calibrated
- This would save time and resources (personal, money)
- This comparative analysis has never been done before.

Definition - Geographic Profile

… is a decision support tool used by law enforcement to make estimates about the likely location of a serial offender's haven.

Comparison - Geographic Profiling Models

Paulsen, Derek J. 2006 “Connecting the Dots: Assessing the Relative Accuracy of Geographic Profiling Software”. *Policing: An International Journal of Police Strategies and Management*. Vol 29, Issue 2, pp. 306-334.

Compares various accuracy measures across different GPs

- JTC GP (using *CrimeStat III* default parameters)
- Rigel
- Dragnet
- simple spatial distribution measures (spatial mean, spatial median, center of minimum distance)

Results - GP Comparison (Paulsen, 2006)

• Probability strategies (JTC GP, Rigel, Dragnet) are not substantially more accurate than spatial distribution measures (spatial mean, spatial median, center of minimum distance)

• Consistent with previous findings (Levine 2002, Snook *et al*. 2005)

Research - GP Comparison

Extend the research by Paulsen (2006) by using individually calibrated distance decay functions instead of the default values in CrimeStat (3.0).

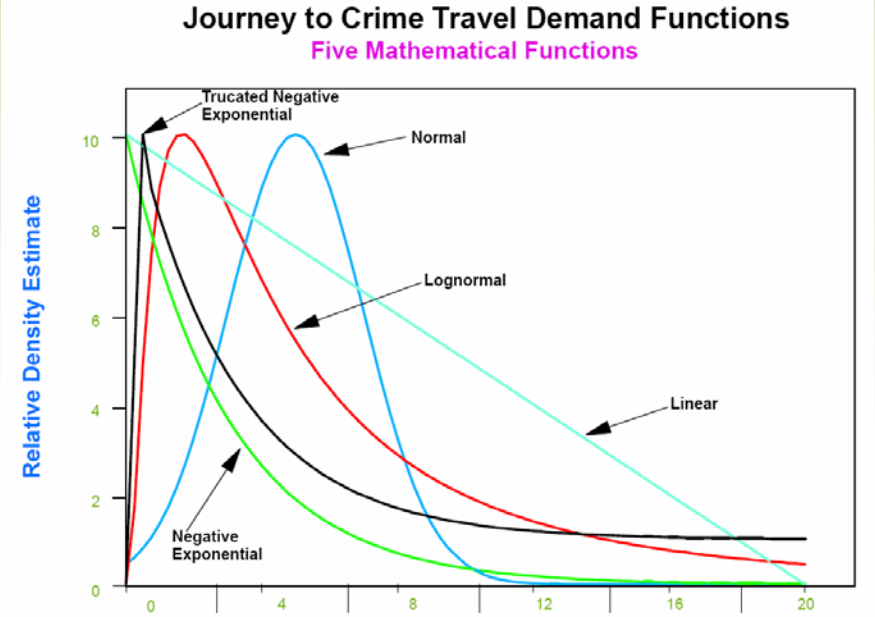
- Same data set
- Same size of Geographic Profile (x-, y-coordinates of lower left and upper right corner of GP)

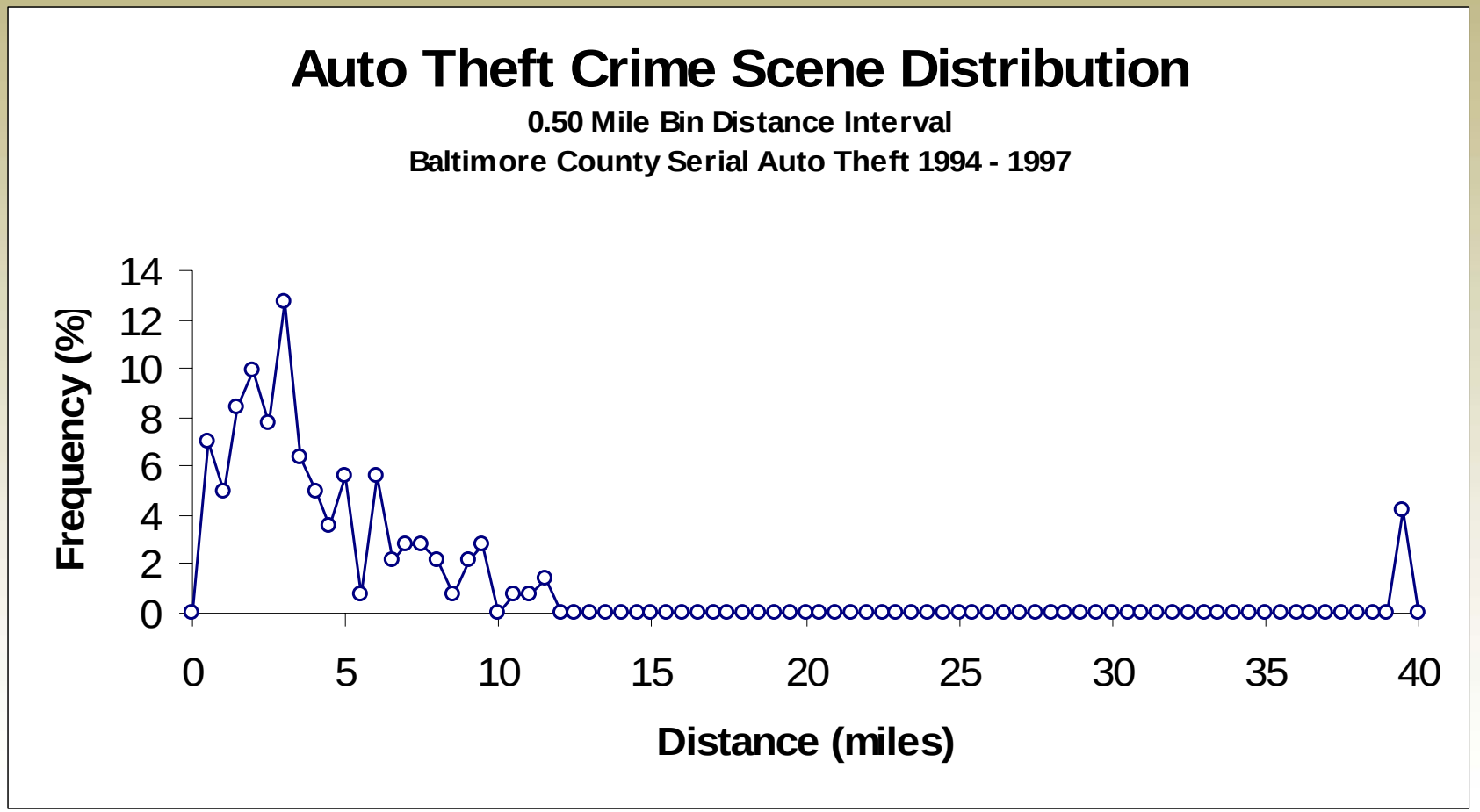
Data & Study Area

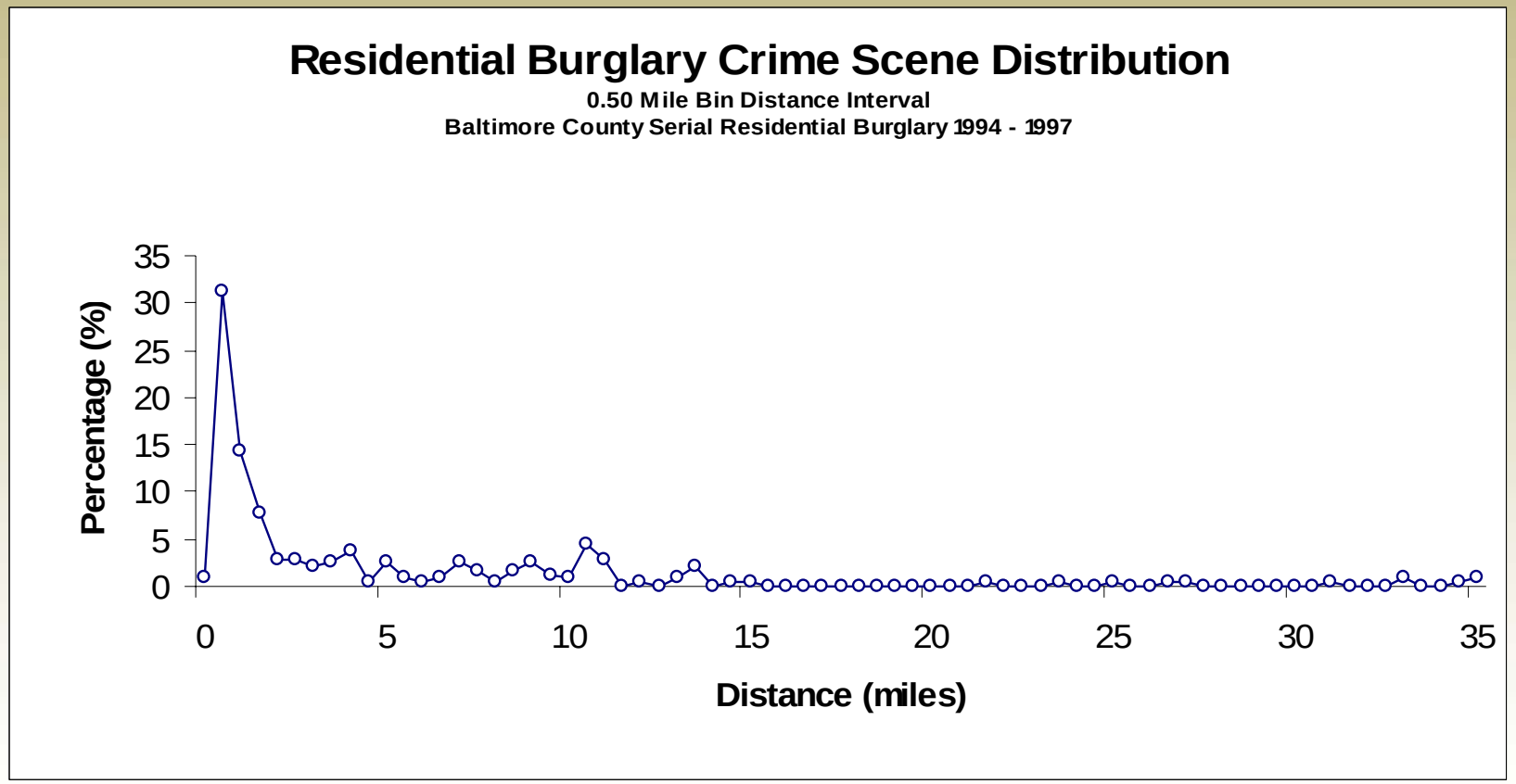
- 247 Serial Crimes
- Nine different crime types
  - Larceny (51 serial crimes)
  - Arson (4 serial crimes)
  - Auto theft (31 serial crimes)
  - Robbery (commercial-76, street-17, mixed-15)
  - Rape (1 serial crime)
  - Burglary (residential-51, commercial-1)
- Both crime locations and actual “haven” known
- Three or more offenses in each crime series
- 1994 – 1997
- Baltimore County, Maryland

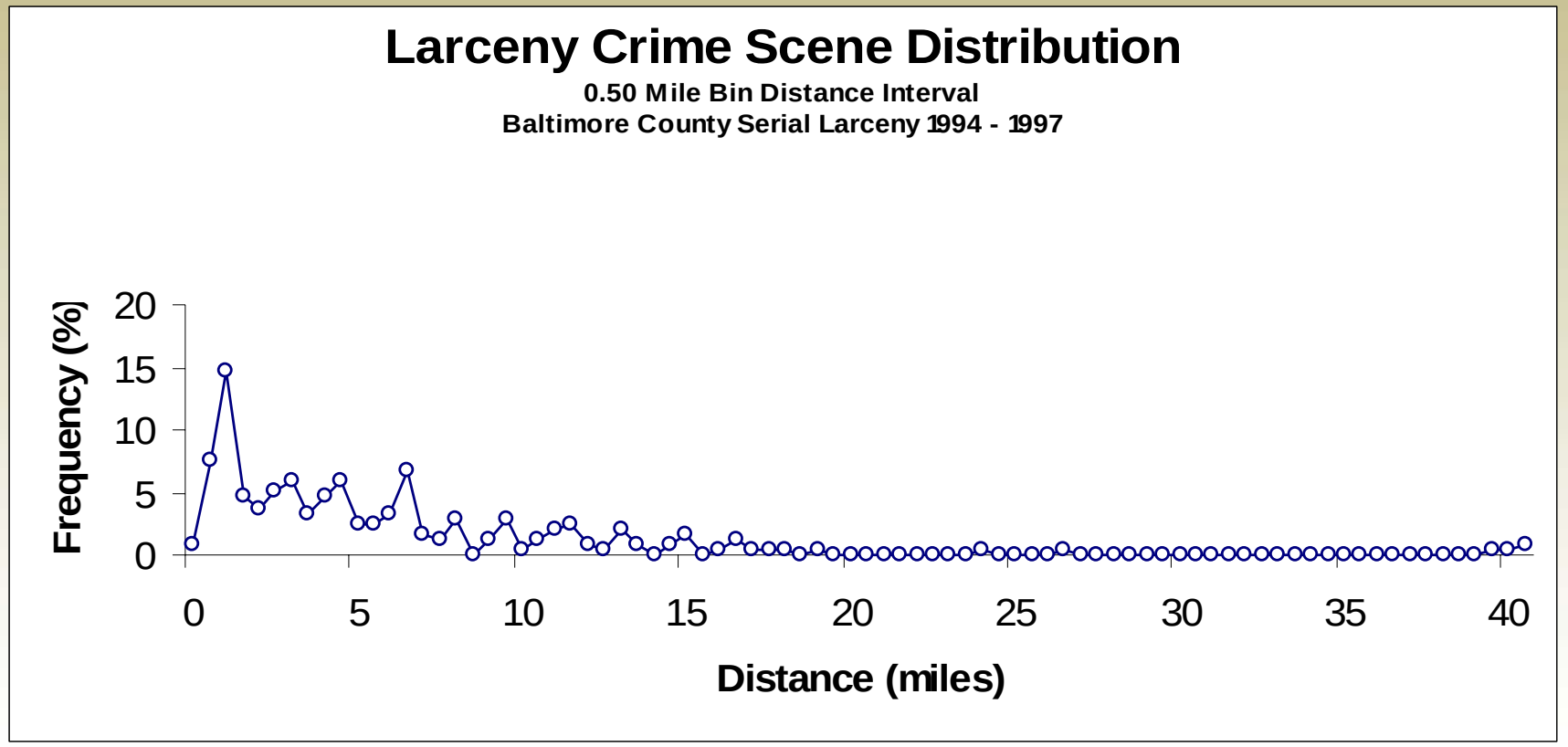
JTC GP Method

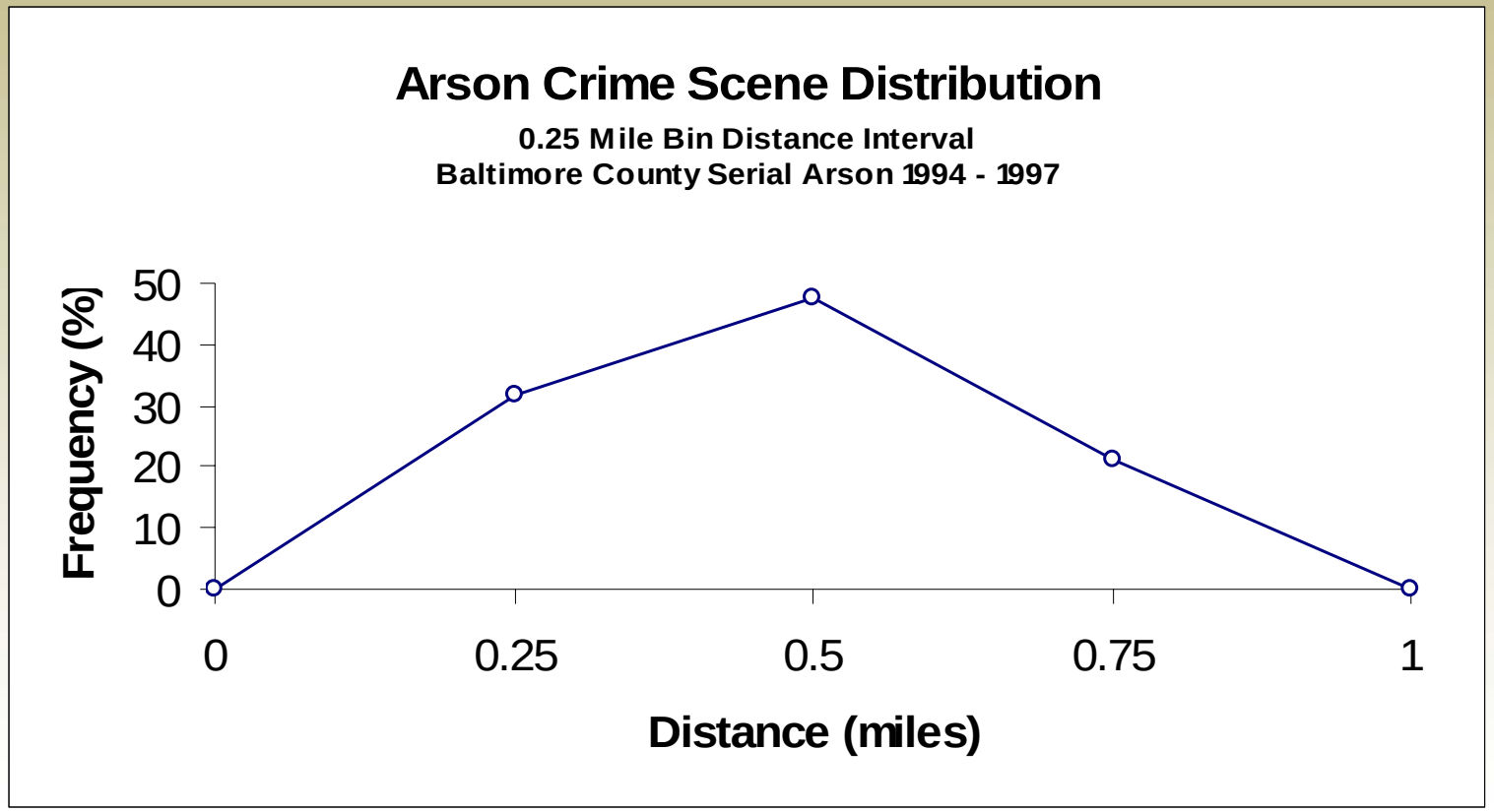
- **Calibration group:** Many (serial) offenders for which travel patterns to and from the crime location are known
- These travel patterns are modeled with various **distance decay functions** (modeling=estimation of parameters, calibration)
- **Test group:** One serial offender with known crime locations
- **JTC GP:** It integrates the crime location from the test group with the calibrated distance decay function.

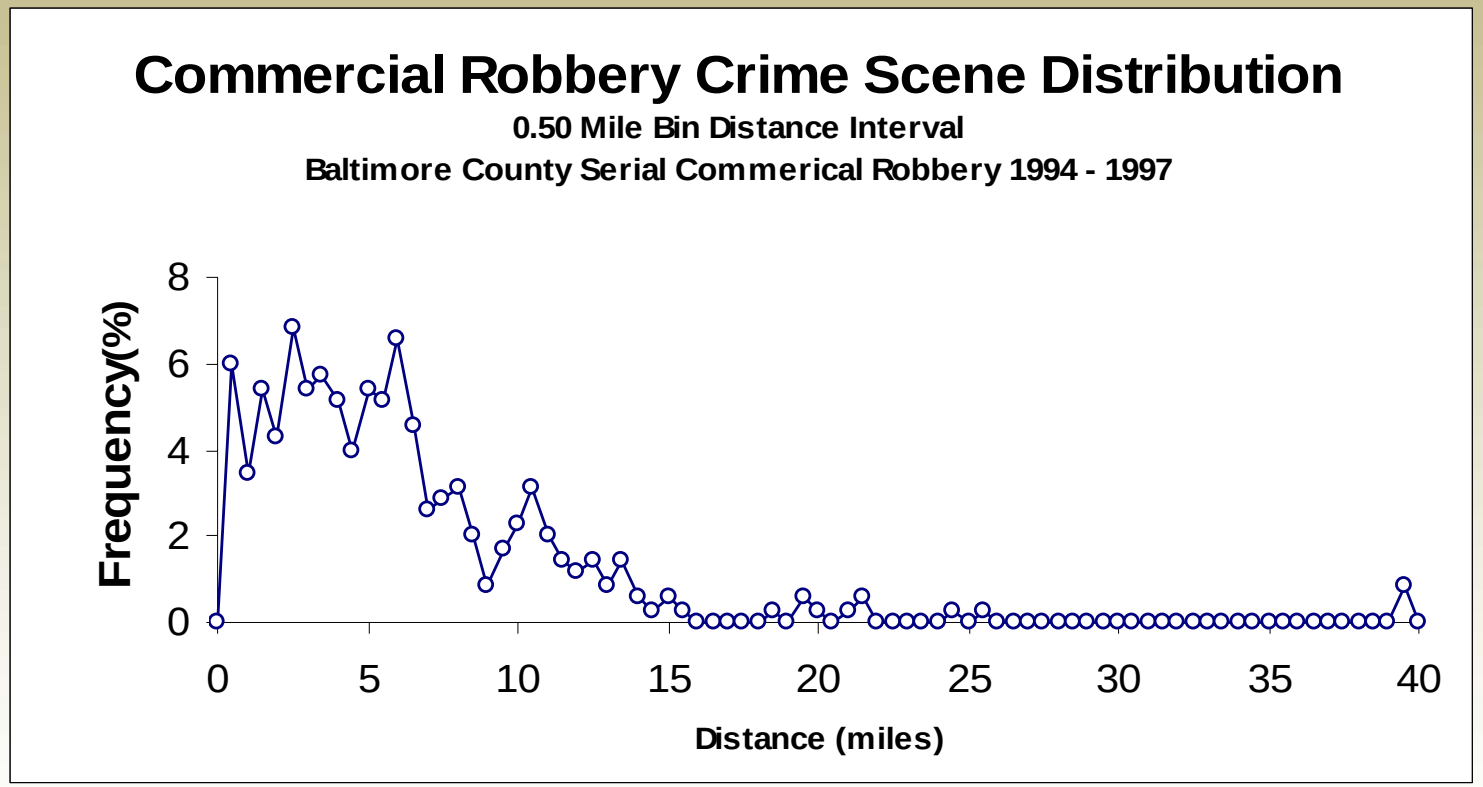


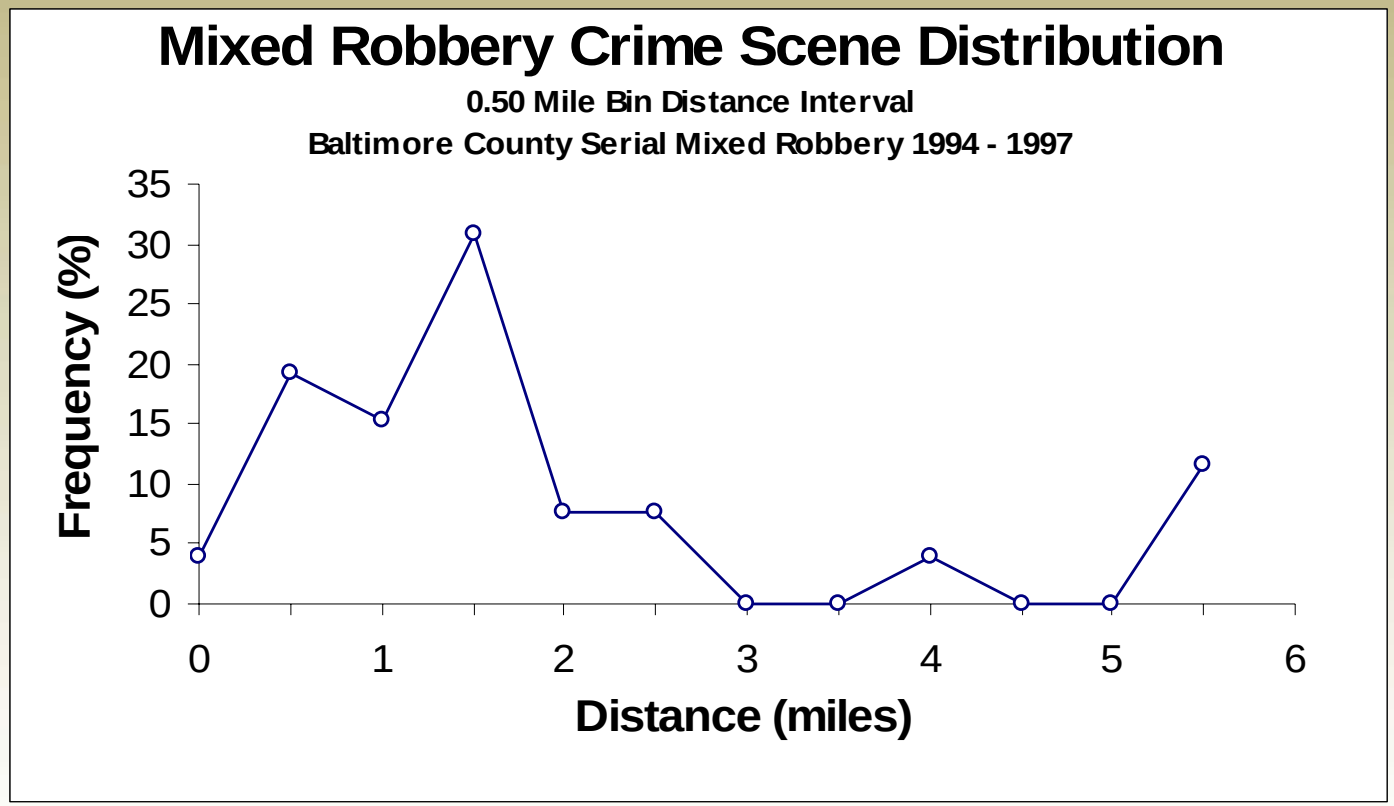

Auto Theft Crime Scene Distribution
0.50 Mile Bin Distance Interval
Baltimore County Serial Auto Theft 1994 - 1997
Frequency (%)
14
12
10
8
6
4
2
0
0
5
10
15
20
25
30
35
40
Distance (miles)

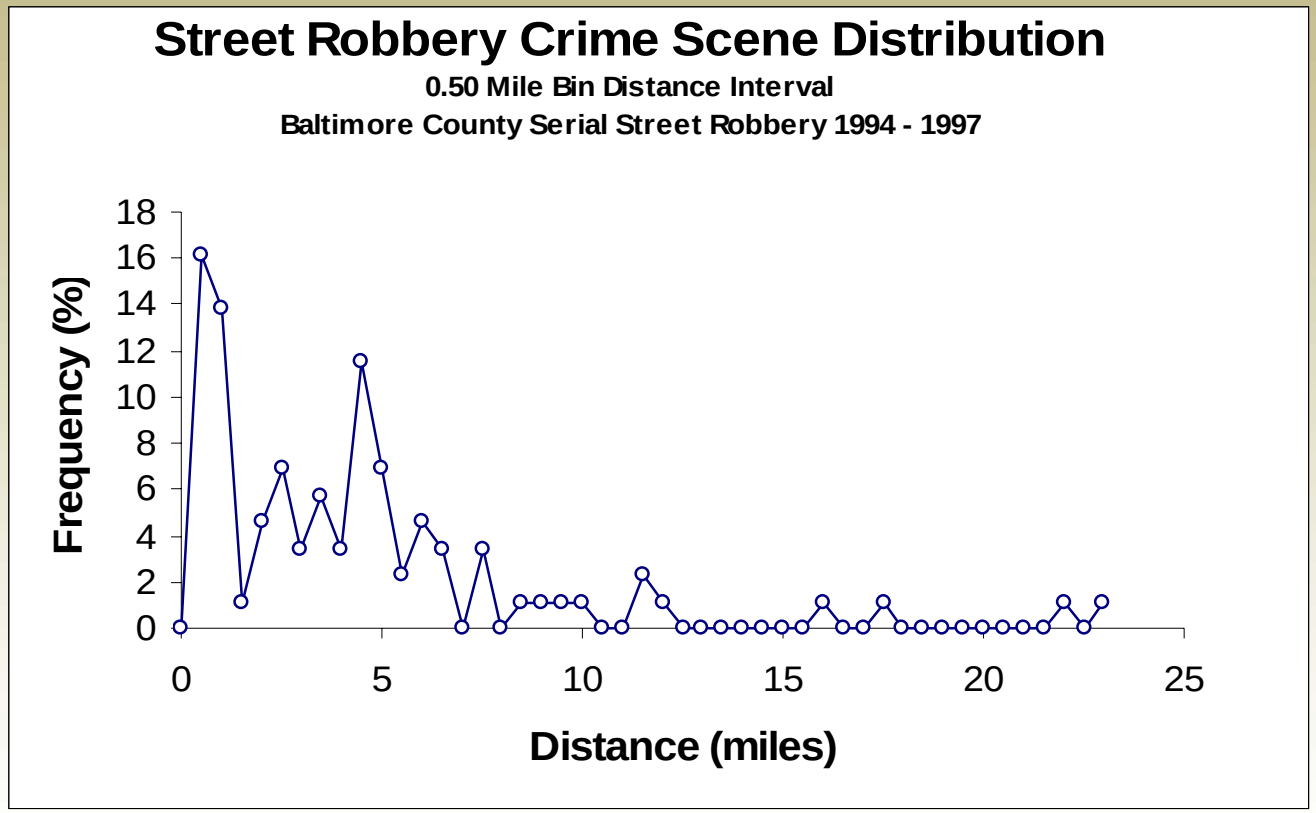

Residential Burglary Crime Scene Distribution
0.50 Mile Bin Distance Interval
Baltimore County Serial Residential Burglary 1994 - 1997
Percentage (%)
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
35
Distance (miles)


Larceny Crime Scene Distribution
0.50 Mile Bin Distance Interval
Baltimore County Serial Larceny 1994 - 1997
Frequency (%)
20
15
10
5
0
0
5
10
15
20
25
30
35
40
Distance (miles)


Arson Crime Scene Distribution
0.25 Mile Bin Distance Interval
Baltimore County Serial Arson 1994 - 1997
Frequency (%)
50
40
30
20
10
0
0
0.25
0.5
0.75
1
Distance (miles)


Commercial Robbery Crime Scene Distribution
0.50 Mile Bin Distance Interval
Baltimore County Serial Commerical Robbery 1994 - 1997
Frequency(%)
8
6
4
2
0
0
5
10
15
20
25
30
35
40
Distance (miles)


Mixed Robbery Crime Scene Distribution
0.50 Mile Bin Distance Interval
Baltimore County Serial Mixed Robbery 1994 - 1997
Frequency (%)
35
30
25
20
15
10
5
0
0
1
2
3
4
5
6
Distance (miles)


Street Robbery Crime Scene Distribution
0.50 Mile Bin Distance Interval
Baltimore County Serial Street Robbery 1994 - 1997
Frequency (%)
18
16
14
12
10
8
6
4
2
0
0
5
10
15
20
25
Distance (miles)

Procedure - Using default parameters for calibrating distance decay functions

- The procedure is explained using the larceny dataset with 51 serial crimes

- **Step 1:** Select the first serial crime from the larceny dataset (**test group**).
- **Step 2:** Create five different JTC GP using the default parameter values for each of the five different distance decay functions implemented in CrimeStat (3.0).
- **Step 3:** Repeat Steps 1-2 for each serial crime in the larceny dataset.

This procedure results in 51 JTC GP for each of the five default calibrated distance decay functions.

Procedure - Using parameter values from individually calibrated distance decay functions

- The procedure is explained using the larceny dataset with 51 serial crimes
- **Step 1:** Remove the first serial crime (**test group**) from the larceny dataset.
- **Step 2:** Calibrate five different distance decay functions from the remaining 50 serial crimes (**calibration group**).
- **Step 3:** Create a JTC GP for each of the five different distance decay functions for the first larceny serial crime removed from the dataset in Step 1.
- **Step 4:** Repeat Steps 1-3 for each serial crime in the larceny dataset.

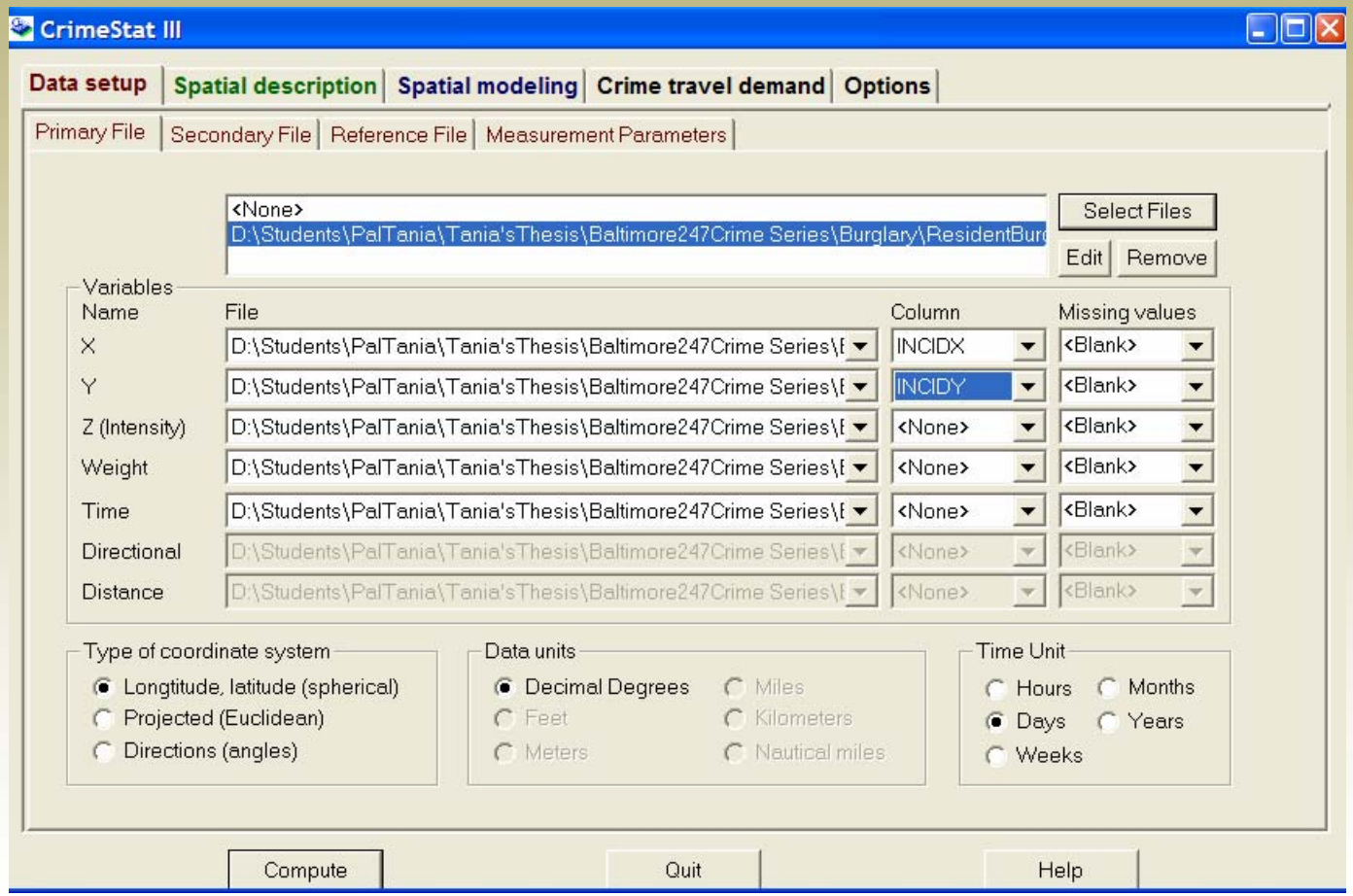
This procedure results in 51 JTC GP for each of the five individually calibrated distance decay functions.

Comparison between GP calculated from default and individually calibrated distance decay functions

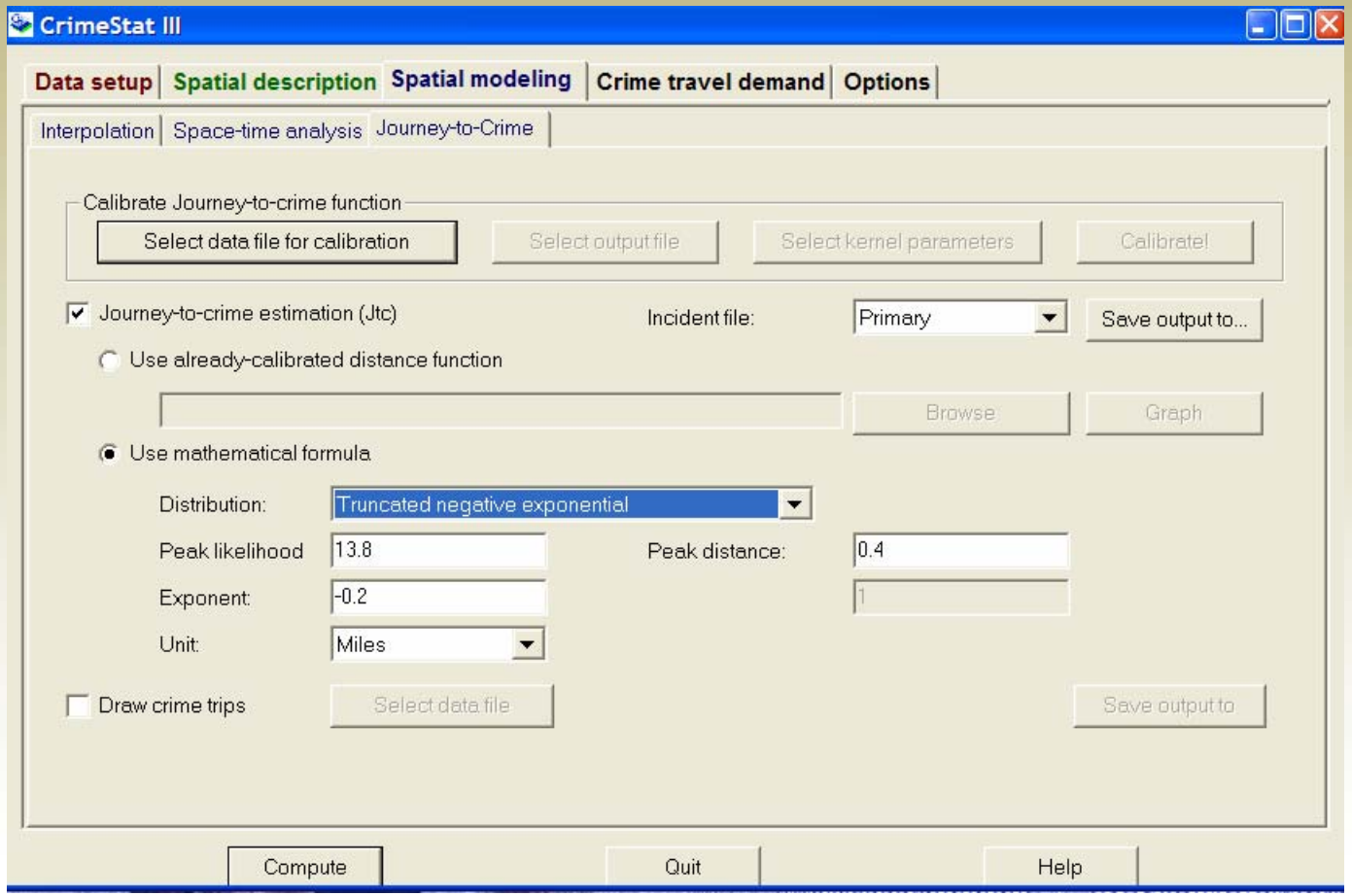
- The comparison is conducted for
  - All crime types
  - Five different distance decay functions (linear, negative exponential, truncated negative exponential, normal, lognormal), and
  - Three different comparison measurements (error distance, search area size, and hit score percentage).

- For each crime type, distance decay function, and comparison measurement a paired-samples t-Test is used to compare the individually calibrated with the default calibrated JTC GP.

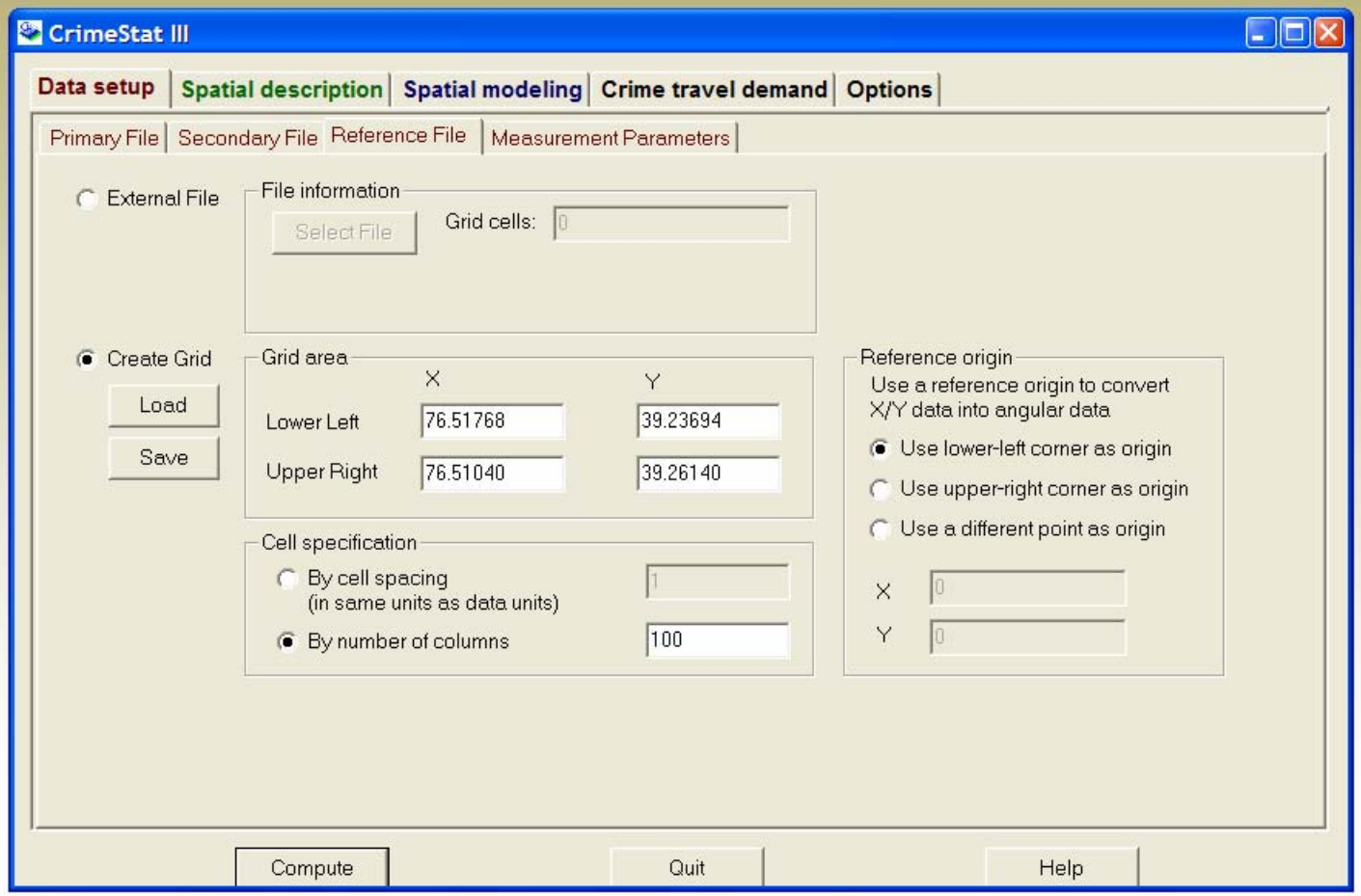
JTC GP in CrimeStat III



JTC GP in CrimeStat III



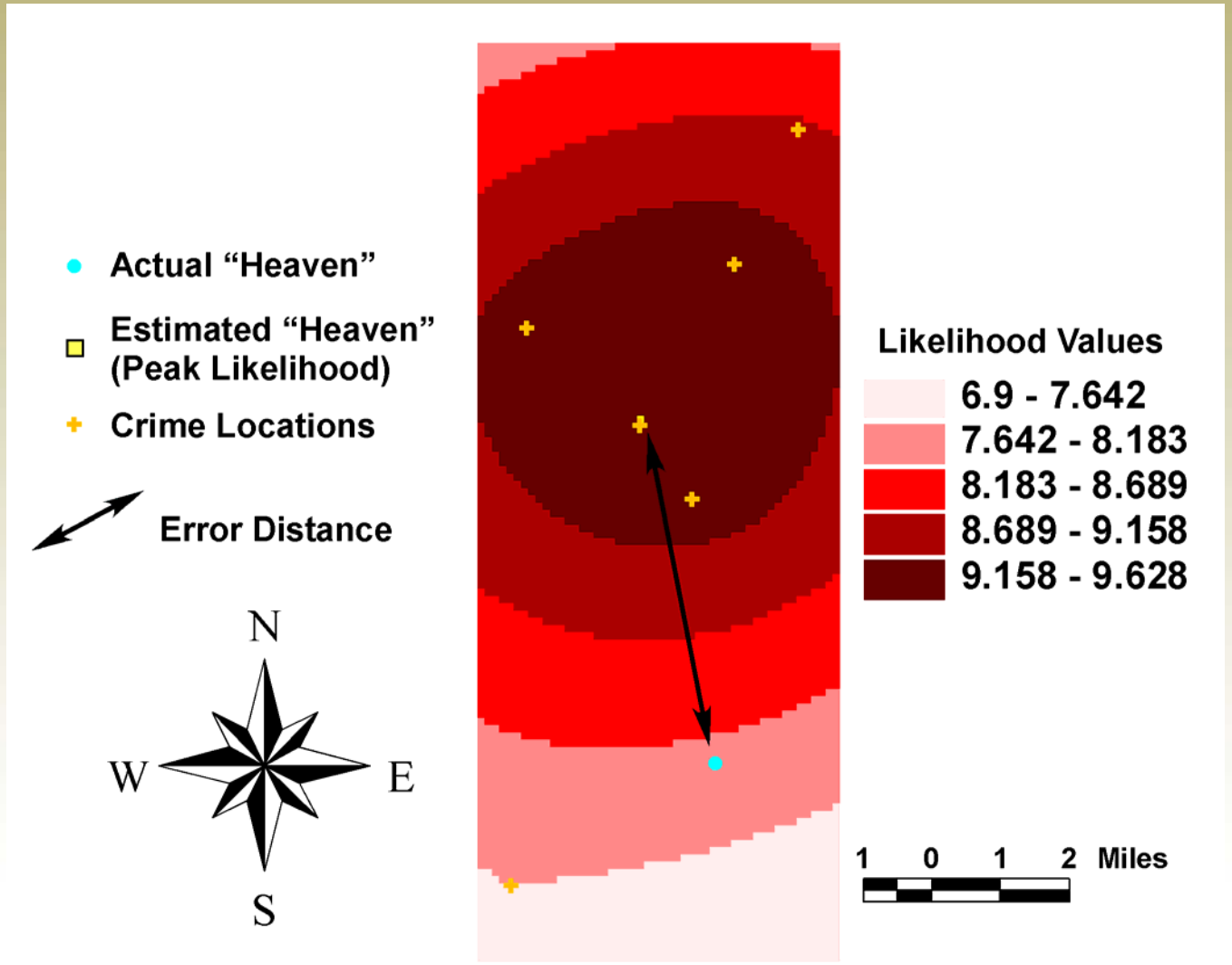
JTC GP in CrimeStat III



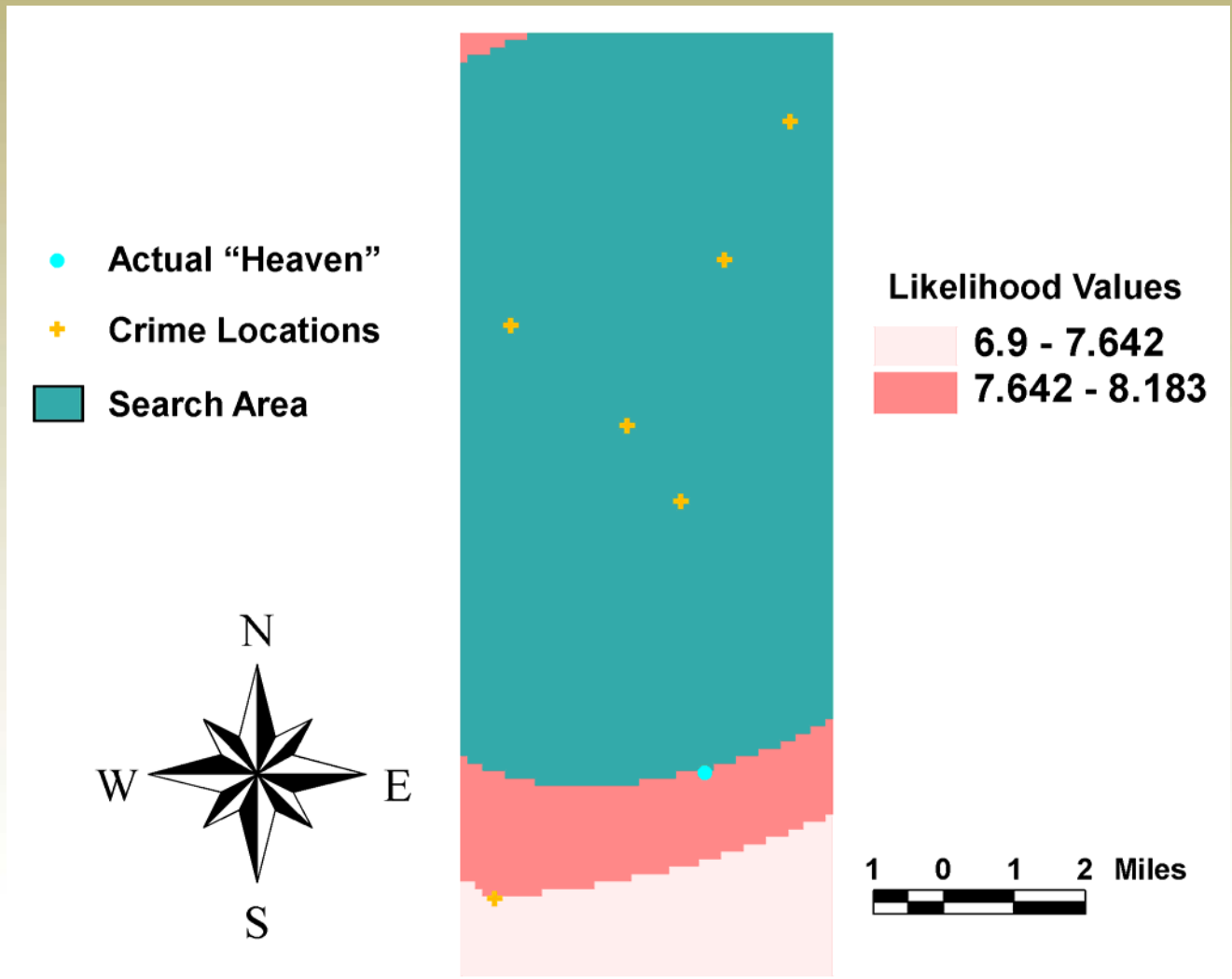
Measures to compare between GP Results

- Error distance
  - Straight-line distance between the actual and the predicted “haven”
- Search area size
  - Area of all cells with a a probability score equal to or higher than the probability score assigned to the actual haven (hit score)
- Hit score percentage, search cost
  - Ratio between search area and study area
  - The lower the hit score percentage, the more accurate the GP

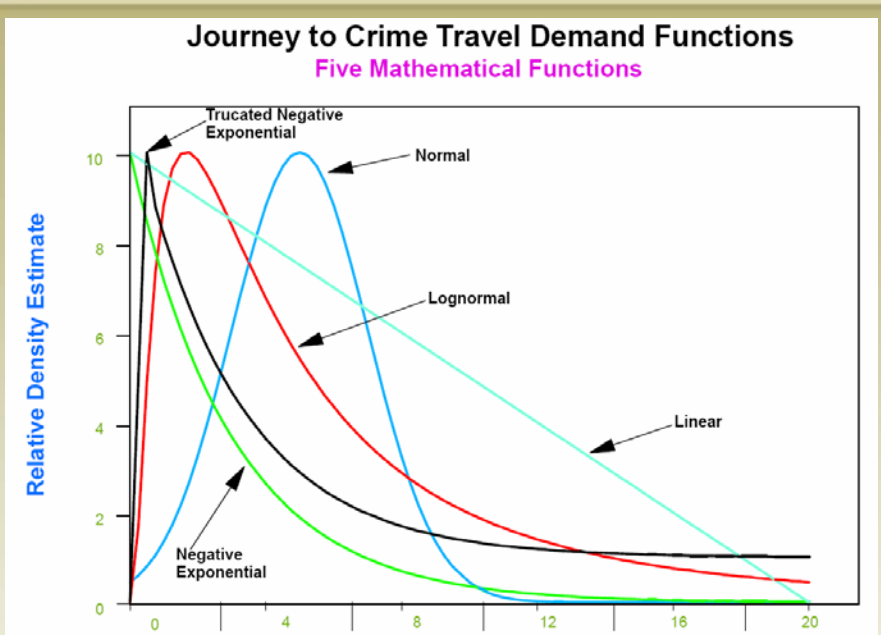
Comparison Measure - **Error Distance**



Comparison Measure – **Search Area**



Auto Theft - Linear Function

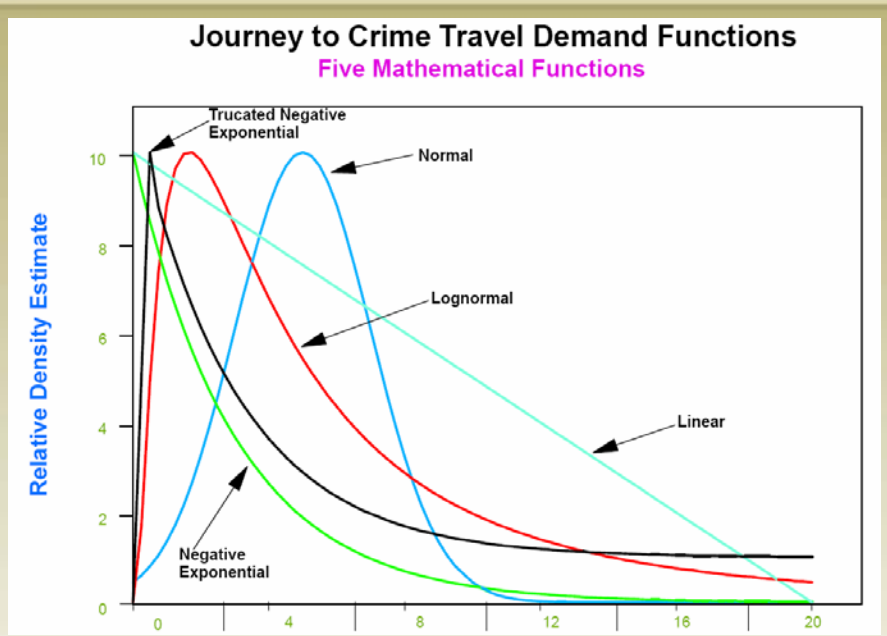


$$f(d_{ij}) = A + B * d_{ij}$$

- Default values: A = 1.9, B = -0.06

- Range of calibrated values: A = 2.000 to 2.089
  B = -0.0069 to -0.0074
  R^2 = 0.302 to 0.364

Auto Theft - Negative Exponential Function

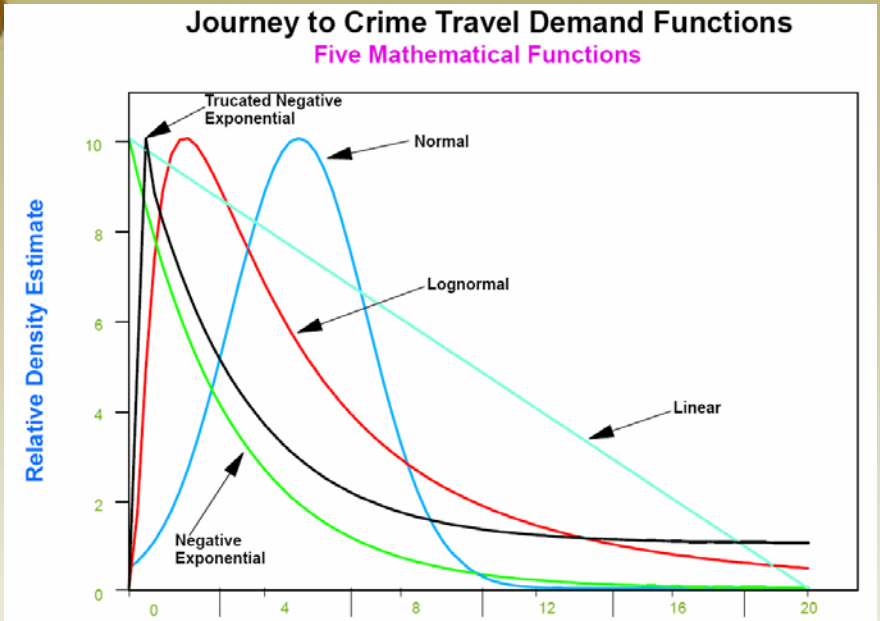


$$f(d_{ij}) = A*e^{-B*d_{ij}}$$

- Default values: A = 1.89, B = -0.06

- Range of calibrated values:
  A = 0.0069 to 0.6480
  B = -0.380 to -0.407
  R^2 = 0.399 to 0.441

Auto Theft - Normal Function



$$Z_{ij} = \frac{(d_{ij} - \text{MeanD})}{S_d}$$

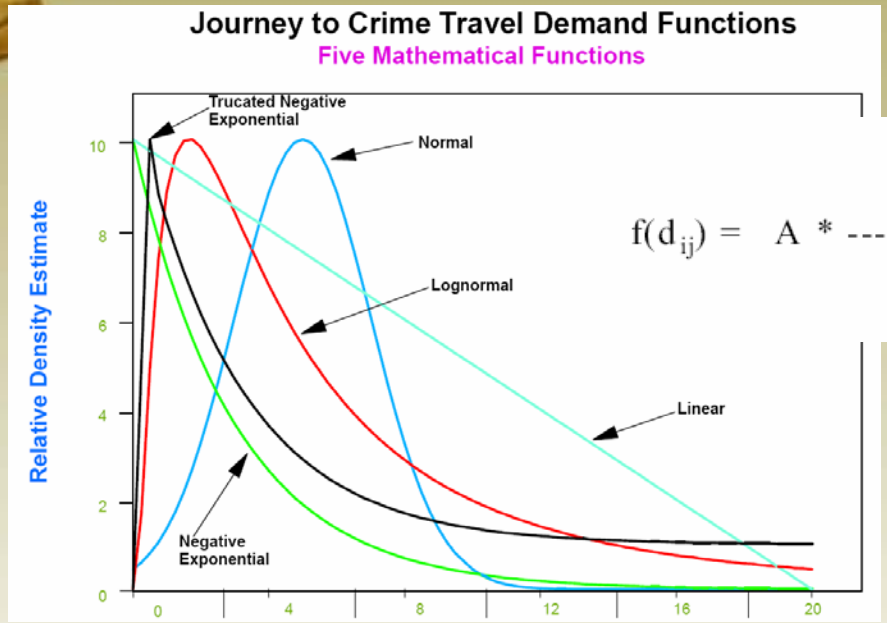
$$f(d_{ij}) = A * \frac{1}{S_d * \text{SQRT}(2\pi)} * e^{-0.5 * Z_{ij}^2}$$

- Default values: MeanD = 1.9, S_d = 4.6, A = 29.5

- Range of calibrated values:

  MeanD = 19.7497
  S_d = 11.4382
  A = 48.5810 to 51.4010
  R^2 = 0.432 to 0.500

Auto Theft - Lognormal Function



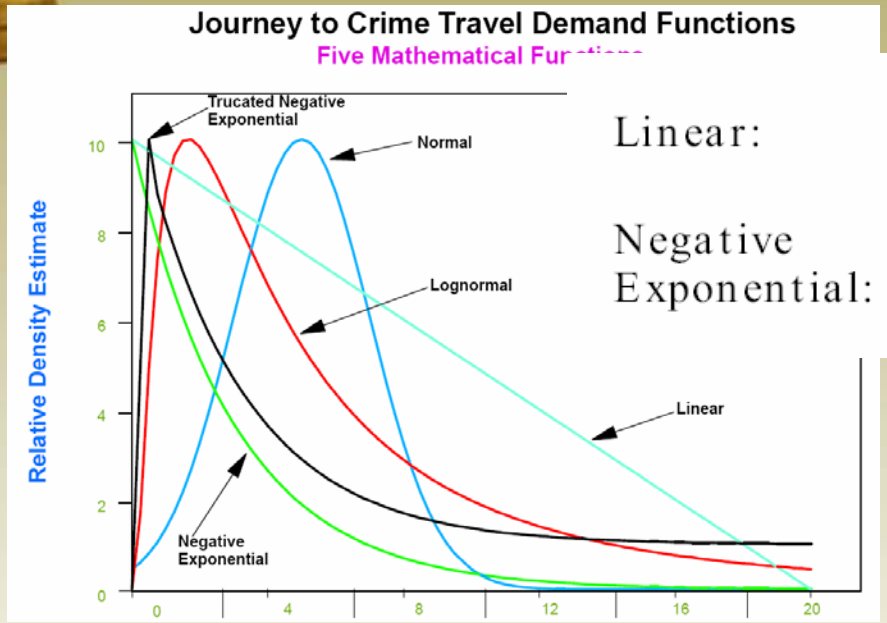
$$f(d_{ij}) = A * \frac{1}{d^2_{ij} * S_d * SQRT(2\pi)} * e^{-[\ln(d^2_{ij}) - MeanD]^2 / 2 * s_d^2}$$

- Default values: MeanD = 4.2, Sd = 4.6, A = 8.6

- Range of calibrated values:

  MeanD = 19.7497
  Sd = 11.4382
  A = -0.3510 to 0.6220
  R^2 = 0.000

Auto Theft - Truncated Negative Exponential Function



Linear: $f(d_{ij}) = 0 + B*d_{ij} = B*d_{ij}$ for $d_{ij} \geq 0$, $d_ij \leq d_p$

Negative Exponential: $f(d_{ij}) = A*e^{-C*d_{ij}}$ for $X_i > d_p$

- Default values: $d_p = 0.4$
  peak likelihood = 13.8,
  C = -0.2
- Range of calibrated values: $d_p = 1.375$ to 2.875
  peak likelihood = 5.8394 to 7.6923
  C = -0.301 to -0.383
  $R^2 = 0.286$ to 0.395

Auto Theft – Comparing **Error Distance** (m) using a Paired-Samples T Test

Distance Decay Function	**Calibrated**	**Default**	**T-Test Statistic**	**Significance (2-tailed)**
Linear	4766	4766		
Negative Exponential	5354	5338	0.544	0.591
Normal	**9545**	**4807**	**4.916**	**0.000**
Lognormal	7181	6508	1.353	0.186
Truncated Neg. Exp.	5173	4994	0.375	0.710

Auto Theft – Comparing **Hit Score Percentage** using a Paired-Samples T Test

Distance Decay Function	**Calibrated**	**Default**	**T-Test Statistic**	**Significance (2-tailed)**
Linear	34.34	34.34		
Negative Exponential	43.29	36.31	1.772	0.098
Normal	**65.85**	**43.93**	**2.845**	**0.013**
Lognormal	51.98	39.58	1.448	0.17
Truncated Neg. Exp.	43.23	46.69	-0.296	0.771

Auto Theft – Comparing **Search Area** (mi^2) using a Paired-Samples T Test

Distance Decay Function	**Calibrated**	**Default**	**T-Test Statistic**	**Significance (2-tailed)**
Linear	34.34	34.34		
Negative Exponential	40.55	30.08	2.056	0.059
Normal	**62.06**	**30.87**	**3.189**	**0.007**
Lognormal	**62.04**	**30.86**	**3.189**	**0.007**
Truncated Neg. Exp.	32.16	36.98	-1.121	0.281

Summary of Results

- For serial offenses involving Auto Thefts, comparison measurements to not show significant different results, whether or not distance decay functions are calibrated.
- Unexpectedly, the normal distance decay function using the default parameters produces more accurate GP for all three comparison measurements.
- Preliminary results indicate that spending time and resources to calibrate distance decay functions individually may not be necessary.

Future Research

- Redo this analysis with marauder type serial offenders only (i.e., remove the commuter type serial offenders).
- Redo this comparison analysis with a Bayesian JTC routine.
- As always, redo this analysis for other study areas and different set of serial offense data.