Chapter 14:[1]
Bayesian Journey-to-Crime Modeling

Ned Levine
Ned Levine & Associates
Houston, TX

Richard Block
Loyola University
Chicago, IL

Table of Contents

Table of Contents (continued)

Chapter 14:

Bayesian Journey-to-crime Modeling

The Bayesian Journey-to-crime module (Bayesian Jtc) includes a set of tools for estimating the likely residence location of a serial offender. It is an extension of the Journey-to-crime routine (Jtc) that uses a travel distance function to make an estimate about the likely residence location of a serial offender. The Bayesian Jtc routine adds information about specific origins of offenders who committed crimes in the same locations to the Jtc to update the estimate. Before proceeding with this chapter, users should be thoroughly familiar with the material on Jtc modeling discussed in Chapter 13.

First, the theory behind the Bayesian Jtc routine will be described. While this material is not essential for running the routine, it does provide the background behind the routine. Users who want to go immediately into the routine should skip to the data section on p. 14.10.

Second, data requirements will be discussed. Third, the routine will be illustrated with data from Baltimore County and from Chicago. Fourth, the use of probability filters as extensions will be illustrated. Fifth, and finally, some guidelines are provided for analysts.

Bayesian Probability

Bayes Theorem is a formulation that relates the conditional and marginal probability distributions of random variables. The *marginal probability* distribution is a probability independent of any other conditions. Hence, P(A) and P(B) is the marginal probability (or just plain probability) of A and B respectively.

The *conditional probability* is the probability of an event given that some other event has occurred. It is written in the form of P(A|B) (i.e., event A given that event B has occurred). In probability theory, it is defined as:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} \tag{14.1}$$

Conditional probabilities can be best be seen in contingency tables. Table 14.1 below shows a possible sequence of counts for two variables (e.g., taking a sample of persons and counting their gender - male = 1 v. female = 0, and their age - older than 30 = 1 v. 30 or younger = 0). The probabilities can be obtained just by counting:

P(A) = 30/50 = 0.6
P(B) = 35/50 = 0.7
P(A and B) = 25/50 = 0.5
P(A or B) = (30+35-25)/50 = 0.8
P(A|B) = 25/35 = 0.71
P(B|A) = 25/30 = 0.83

However, if four of these six calculations are known, Bayes Theorem can be used to solve for the other two. Two logical terms in probability are the 'and' condition and the 'or' condition. Usually, the symbol ∩ is used for 'and' ∪ is used for 'or', but writing it in words might make it easier to understand.

Table 14.1:
Example of Determining Probabilities by Counting

	A has NOT occurred	A has occurred	**TOTAL**
B has NOT occurred	10	5	**15**
B has occurred	10	25	**35**
TOTAL	**20**	**30**	**50**

The following two theorems define these.

1.   The probability that *either* A *or* B will occur is:

$$P(A \ or \ B) = P(A) + P(B) - P(A \ and \ B)$$ (14.2)

2.   The probability that *both* A *and* B will occur is:

$$P(A \ and \ B) = P(A) * P(B|A) = P(B) * P(A|B)$$ (14.3)

Bayes Theorem relates the two equivalents of the *'and'* condition together:

$$P(B) * P(A|B) = P(A) * P(B|A)$$ (14.4)

$$P(A|B) = \frac{P(A)*P(B|A)}{P(B)} \qquad\qquad (14.5)$$

The theorem is sometimes called the 'inverse probability' in that it can invert two conditional probabilities:

$$P(B|A) = \frac{P(B)*P(A|B)}{P(A)} \qquad\qquad (14.6)$$

By plugging in the values from the example in Table 14.1, the reader can verify that Bayes Theorem produces the correct results, for example:

$$P(B|A) = \frac{0.7*0.71}{0.6} = 0.83 \qquad\qquad (14.7)$$

Bayesian Inference

In the statistical interpretation of Bayes Theorem, the probabilities are estimates of a random variable. Let θ be a parameter of interest and let X be some data. Thus, Bayes Theorem can be expressed as:

$$P(\theta|X) = \frac{P(X|\theta)*P(\theta)}{P(X)} \qquad\qquad (14.8)$$

Interpreting this equation, $P(\theta|X)$ is the probability of θ given the data, X, and is called the *posterior probability* (or posterior distribution). $P(\theta)$ is the probability that θ has a certain distribution and is often called the *prior probability*. $P(X|\theta)$ is the probability that the data would be obtained given that θ is true and is often called the *likelihood function* (i.e., it is the likelihood that the data will be obtained given the distribution of θ). Finally, $P(X)$ is the marginal probability of the data, the probability of obtaining the data under all possible scenarios; essentially, it is the data.

The equation can be rephrased in words:

Posterior probability that θ is true given the data, X =	Likelihood of obtaining the data given θ is true *	Prior probability of θ

$$\text{data, X} \quad = \quad \frac{\text{--}}{\text{Marginal probability of X}} \qquad (14.9)$$

In other words, this formulation allows an estimate of the probability of a particular parameter, θ, to be updated given new information. Since θ is the prior probability of an event,

given some new data, *X*, Bayes Theorem can be used to update the estimate of θ. The prior probability of θ can come from prior studies, an assumption of no difference between any of the conditions affecting θ, or an assumed mathematical distribution. The likelihood function can also come from empirical studies or an assumed mathematical function. Irrespective of how these are interpreted, the result is an estimate of the parameter, θ, given the evidence, *X*.

A point that is often made is that the prior probability of obtaining the data (the denominator of the above equation) is not known or cannot easily be evaluated. The data are what was obtained from some data gathering exercise (either experimental or from observations). Thus, it is not easy to estimate it. Consequently, often the numerator only is used for estimate the posterior probability since:

$$P(\Theta|X) \propto P(X|\Theta) * P(\Theta) \tag{14.10}$$

where $\propto$ means 'proportional to'. In some statistical methods (e.g., the Markov Chain Monte Carlo simulation, or MCMC, discussed in Chapters 17, 18 &19), the parameter of interest is estimated by thousands of random simulations using approximations to $P(X|\theta)$ and $P(\theta)$ respectively.

The key point is that estimates of parameters can be systematically *updated* by additional information. The formula requires that a prior probability for the estimate be given with new information being added which is *conditional* on the prior estimate, meaning that it takes into account information from the prior. Bayesian approaches are increasingly being used to provide estimates for complex calculations that previously were intractable (Denison, Holmes, Mallilck, & Smith, 2002; Lee, 2004; Gelman, Carlin, Stern, & Rubin, 2004). Our regression module includes the use of the MCMC algorithm to estimate complex equations.

Application of Bayesian Inference to Journey-to-crime Analysis

Bayes Theorem can be applied to the Journey-to-crime methodology. In the Journey-to-crime (Jtc) method, an estimate is made about where a serial offender is living. The Jtc method produces probability estimates based on an assumed travel distance function (or, in more refined uses of the method, travel time). That is, it is assumed that an offender follows a typical travel distance/time function. This function can be estimated from prior studies (Canter & Gregory, 1994; Canter, 2003) or from creating a sample of known offenders - a calibration sample (see Chapter 13; Levine, 2000) or from assuming that every offender follows a particular mathematical function (Rossmo, 1995; 2000). Essentially, it is a prior probability for a particular location, $P(\theta)$. That is, it is a guess about where the offender lives on the assumption that the offender of interest is following an existing travel distance model.

However, additional information from a sample of known offenders where both the crime location and the residence location are known can be added. This information would be obtained from arrest records, each of which will have a crime location defined (a 'destination') and a residence location (an 'origin'). If these locations are then assigned to a set of zones, a matrix that relates the origin zones to the destination zones can be created (Figure 14.1). This is called an ***origin-destination*** matrix (also known as a *trip distribution* or *O-D* matrix, for short).

In this figure, the numbers indicate crimes committed in each destination zone which originated from each origin zone (i.e., where the offender lived). For example, taking the first row in Figure 14.1, there were 37 crimes that were committed in zone 1 and in which the offender also lived in zone 1; there were 15 crimes committed in zone 2 in which the offender lived in zone 1; however, there were only 7 crimes committed in zone 1 in which the offender lived in zone 2; and so forth.

Note two things about the matrix. First, the number of origin zones can be (and usually is) greater than the number of destination zones because crimes can originate outside the study area. Second, the marginal totals have to be **equal**. That is, the number of crimes committed in all destination zones must equal the number of crimes originating in all origin zones.

This information can be treated as the likelihood estimate for the Journey-to-crime framework. That is, if a certain distribution of incidents committed by a particular serial offender is known, then this matrix can be used to estimate the likely origin zones from which offenders came, independent of any assumption about travel distance. In other words, this matrix is equivalent to the likelihood function in Equation 14.8, which is repeated below:

$$P(\Theta|X) = \frac{P(X|\Theta)*P(\Theta)}{P(X)} \qquad\qquad \text{repeat (14.8)}$$

The estimate of the likely origin location of a serial offender can be improved by updating the Jtc estimate, *P(θ)*, with information from an empirically-derived likelihood estimate, *P(X|θ)*.

Figure 14.2 illustrates the process. Suppose a serial offender committed crimes in three zones. These are shown in terms of grid cell zones. In reality, most zones are not grid cells but are irregular. However, illustrating with grid cells makes the process more understandable. Using an O-D matrix based on those cells, only the destination zones corresponding to those cells are selected (Figure 14.3). This process is repeated for all serial offenders in the calibration file. Destination zones that are repeated by different serial offenders are counted multiple times, once for each occurrence. This results in marginal totals that correspond to frequencies for those serial offenders who committed crimes in the selected zones. The marginal totals are then

Figure 14.1:
Crime Origin-Destination Matrix

Crime destination zone

	1	2	3	4	5	.	.	.	.	N	Σ
1	**37**	15	21	4	3					12	*346*
2	7	**53**	14	0	4					15	*1050*
3	12	9	**81**	7	6					33	*711*
4	4	10	6	**12**	1					0	*84*
5	8	7	28	2	**24**					14	*178*
.											
.											
.											
M	12	5	43	3	10					**92**	*1466*
Σ	*153*	*276*	*1245*	*99*	*110*					*812*	**43,240**

Crime origin zone

converted to probabilities. In other words, the distribution of crimes is *conditioned* on the locations that correspond to where the serial offender of interest committed his or her crimes. It is a conditional probability.

But, what about the denominator of the Bayesian formula, P(X)? Essentially, it is the spatial distribution of all crimes irrespective of which particular model or scenario we are exploring. In practice, it is very difficult, if not impossible, to estimate the probability of obtaining the data under all circumstances. Therefore, only the numerator in equation 14.8 is estimated and the final probabilities are re-scaled so that they sum to 1.0 over the study area:

$$P(\Theta|X) \propto k * P(X|\Theta) * P(\Theta) \qquad (14.11)$$

where *k* is a scaling constant.

We are going to change the symbols at this point so the *Jtc* represents the distance-based Journey-to-crime estimate, *O* represents an estimate based on an origin-destination matrix, and *O|Jtc* represents the particular origins associated with crimes committed in the same zones as that identified in the *Jtc* estimate. Therefore, there are three different probability estimates of where an offender lives:

1.   A probability estimate of the residence location of a single offender based on the location of the incidents that this person committed and an assumed travel distance function, *P(Jtc)*;

2.   A probability estimate of the residence location of a single offender based on a general distribution of all offenders, irrespective of any particular destinations for incidents, *P(O)*. Essentially, this is the distribution of origins irrespective of the destinations; and

3.   A probability estimate of the residence location of a single offender based on the distribution of offenders given the distribution of incidents committed by other offenders who committed crimes in the same location, *P(O|Jtc)*.

Therefore, Bayes Theorem can be used to create an estimate that combines information both from a travel distance function and an origin-destination matrix in which the posterior probability of the Journey-to-crime location taking into account the origin-destination matrix is proportional to the product of the prior probability of the Journey-to-crime function, *P(Jtc)*, and the conditional probability of the origins for other offenders who committed crimes in the same locations, *P(O)*. This will be called the ***product*** probability.

Figure 14.2:

Bayesian Journey to Crime Routine

Selecting Zones Where Offender Committed Crimes

Zones where offender committed crimes

Legend
- Incidents committed by offender
- Grid overlay
- Selected grid cells
- City of Baltimore
- Baltimore County

0 0.5 1 2 3 Miles
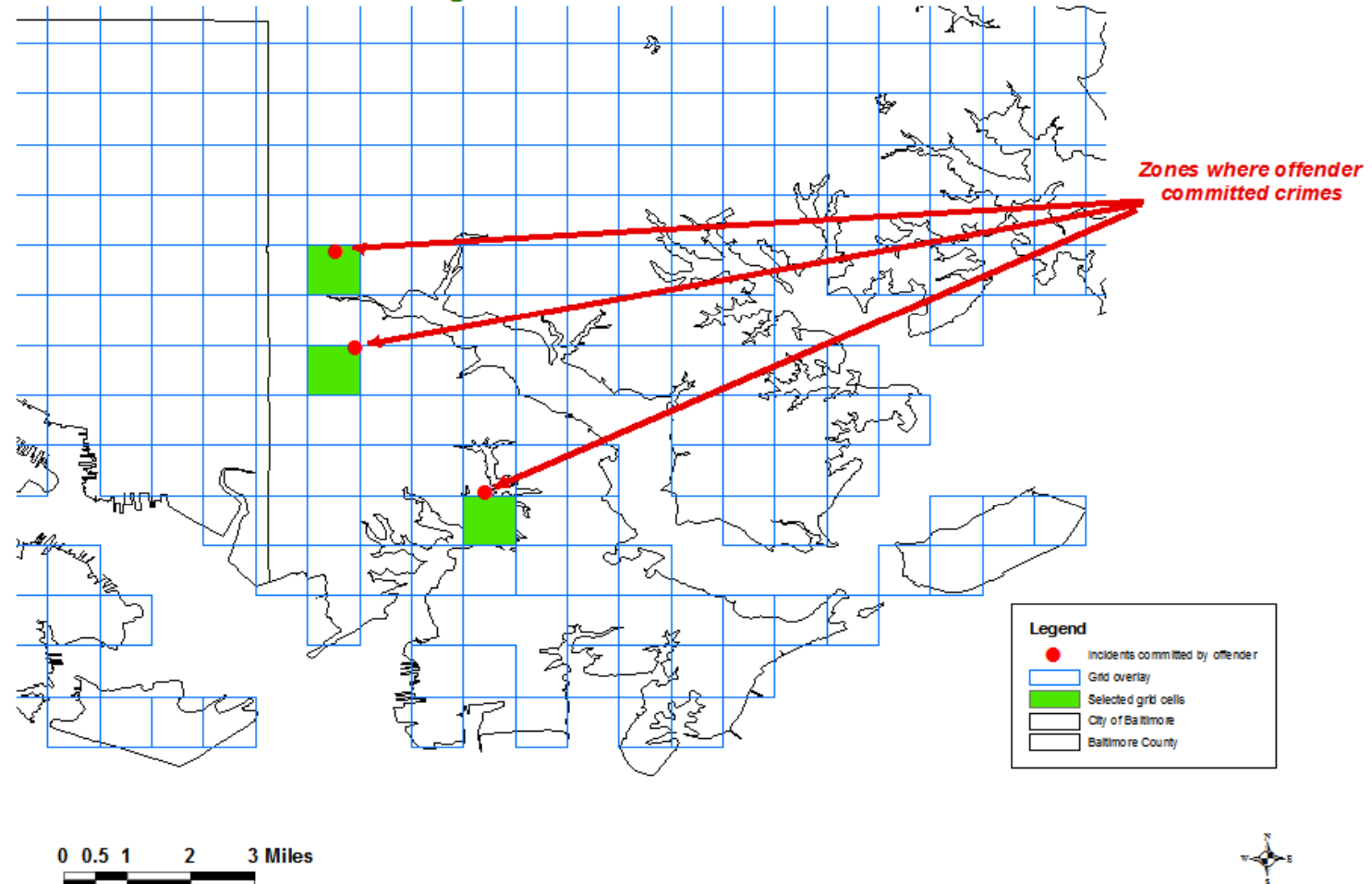
Figure 14.3:
Conditional Origin-Destination Matrix

Destination zones where serial offender committed crimes

Marginal totals for selected zones only

Crime destination zone

Crime origin zone

	1	2	3	4	5	.	.	.	.	N	Σ
1		15		4						12	121
2		**53**		0						15	205
3		9		7						33	65
4		10		**12**						0	35
5		7		2						14	40
.											
.											
.											
M		5		3						**92**	141
Σ		276		99						812	1,597

As mentioned above, it is very difficult to determine the probability of obtaining the data under any circumstance, P(O). Consequently, the Bayesian estimate is usually calculated only with respect to the numerator, the product of the prior probability and the likelihood function, and the result re-scaled so that the probabilities over the study area sum to 1.0.

A very rough approximation to the full Bayesian probability can be obtained by taking the product probability and dividing it by the general probability: It relates the product term (the numerator) to the general distribution of crimes. This will produce a relative risk measure, which is called **Bayesian Risk**:

$$P(Jtc|O) = \frac{P(O|Jtc)*P(Jtc)}{P(O)} \qquad (14.12)$$

In this case, the product probability is being compared to the general distribution of the origins of all offenders irrespective of where they committed their crimes. Note that this measure will correlate with the product term because they both have the same numerator.

The Bayesian Journey-to-crime Estimation Module

The Bayesian Journey-to-crime estimation module is made up of two routines, one for diagnosing which Journey-to-crime method is best and one for applying that method to a particular serial offender. Figure 14.4 show the layout of the module.

Data Preparation for Bayesian Journey-to-crime Estimation

There are four data sets that are required:

1. The incidents committed by a single offender for which an estimate will be made of where that individual lives;

2. A Journey-to-crime travel distance function that estimates the likelihood of an offender committing crimes at a certain distance (or travel time if a network is used);

3. An origin-destination matrix; and

4. A diagnostics file of multiple known serial offenders for which both their residence and crime locations are known (optional for use in diagnostics routine).

Figure 14.4:
Bayesian Journey-to-crime Screen

Serial Offender Data

The first required data set is information on the location of crimes committed by a single serial offender. For each serial offender for whom an estimate will be made of where that person lives, the data set must include the location of the incidents committed by the offender. The data are a series of records in which each represents a single event. On each record, there are X and Y coordinates identifying the location of the incidents this person has committed (Table 14.2). There may be other data on the records, but the X and Y coordinates are essential.

Table 14.2:
Minimum Information Required for Serial Offenders:
Example for Offender Who Committed Seven Incidents

ID	UCR	INCIDX	INCIDY
TS7C	430.00	-76.494300	39.2846
TS7C	440.00	-76.450900	39.3185
TS7C	630.00	-76.460600	39.3157
TS7C	430.00	-76.450700	39.3181
TS7C	311.00	-76.449700	39.3162
TS7C	440.00	-76.450300	39.3178
TS7C	341.00	-76.448200	39.3123

Journey-to-crime Travel Function

The second data set that is required is a journey-to-crime (Jtc) function. The Journey-to-crime travel function is an estimate of the likelihood of an offender traveling a certain distance. Typically, it represents a frequency distribution of distances traveled, though it could be a frequency distribution of travel times if a network was used to calibrate the function with the Journey-to-crime estimation routine. It can come from an *a priori* assumption about travel distances, prior research, or a calibration data set of offenders who have already been caught. The "Calibrate Journey-to-crime function" routine (on the Journey-to-crime page under Spatial modeling) can be used to estimate this function (see Chapter 13).

The BJtc routine can use two different travel distance functions:

1. An already-calibrated distance function; and

2. A mathematical formula.

Either direct or indirect (Manhattan) distances can be used though the default is direct distance (see Measurement Parameters in Chapter 3, p. 3.29). In practice, an empirically-derived

travel function is often as accurate, if not better, than a mathematically-defined one. Given that an origin-destination matrix is also needed, it is easy for the user to estimate the travel function using the "Calibrate Journey-to-crime function".

If the user does not have data to calibrate a journey-to-crime travel function, then a mathematical model should be used. Typically, the negative exponential function is used for this purpose; the default values will work for many distributions.

Origin-destination Matrix

The third required data set is an origin-destination matrix. The origin-destination matrix relates the number of offenders who commit crimes in one of N zones who live (originate) in one of M zones, similar to Figure 14.1 above. It can be created from the "Calculate observed origin-destination trips" routine (on the 'Describe origin-destination trips' page under the Trip distribution module of the Crime Travel Demand model; see Chapter 28).

How many incidents are needed where the origin and destination location are known? While there is no simple answer to this, the numbers ideally should be in the tens of thousands. If there are N destinations and M rows, ideally one would want an average of 30 cases for each cell to produce a reliable estimate. Obviously, that is a huge amount of data that cannot easily found with any real database. For example, if there are 325 destination zones and 532 origin zones (for the Baltimore County example given below), that would be 172,900 individual cells. If the 30 cases or more rule is applied, then that would require 5,187,000 records or more to produce a barely reliable estimate for most cells.

The task becomes even more daunting when it is realized that most of these links (cells) have few or no cases in them as offenders typically travel along certain pathways. Obviously, such a demand for data is impractical even in the largest jurisdictions. Therefore, we recommend that as much data as possible be used to produce the origin-destination (O-D) matrix, at least several years worth. The matrix can be built with what data is available and then periodically updated to produce better estimates.

Diagnostics File for Bayesian Jtc Routine

The fourth data set is an optional diagnostics file. It is used for estimating which of several alternative parameters is best at predicting the residence location of serial offenders. Essentially, it is a set of serial offenders, each record of which has the X and Y coordinates of both the residence location and the crime location. For example, offender T7B committed seven incidents while offender S8A committed eight incidents. The records of both offenders are placed in the same file along with the records for all other offenders in the diagnostics file.

The diagnostics file provides information about which parameter (to be described below) is best at guessing where an offender lives. The assumption is that if a particular parameter was best with the *K* offenders in a diagnostics file in which the residence location was known, then it also will be best for a serial offender for whom the residence location is not known.

How many serial offenders are needed to make up a diagnostics file? Again, there is no simple answer to this although the number is much less than for the O-D matrix. Clearly, the more, the better since the aim is to identify which parameter is most sensitive with a certain level of precision and accuracy. We used 88 offenders in the diagnostics file for Baltimore County (see below). Certainly, a minimum of 10 would be necessary. But, more would certainly be more accurate. Further, the offender records used in the diagnostics file should be similar in other dimensions to the offender that is being tracked. However, this may be impractical. In the example data set, we combined offenders who committed different types of crimes. The results may be different if offenders who had committed only one type of crimes were tested (though Leitner and Kent, 2009, found that using records for all crimes produced more accurate measures than using crime-specific records).

Once the data sets have been collected, they need to be placed in an appended file, with one serial offender on top of another. Each record has to represent a single incident. Further, the records have to be arranged sequentially with all the records for a single offender being grouped together. The routine automatically sorts the data by the offender ID. But, to be sure that the result is consistent, the data should be prepared in this way.

The structure of the records is similar to the example in Table 14.3 below. At the minimum, there needs to be an ID field and the X and Y coordinates of both crime location and the residence location. Thus, in the example, all the records for the first offender (Num 1) are together; all the records for the second offender (Num 2) are together; and so forth. The ID field is any numeric or string variable. In Table 14.3, the ID field is labeled "ID", but any label would be acceptable as long as it is consistent (i.e., all the records of a single offender are together).

In addition to the ID field, the X and Y coordinates of both the crime and residence location must be included on each record. In the example table, the ID variable is called OffenderID, the crime location coordinates are called IncidX and IncidY while the residence location coordinates are called HomeX and HomeY. Again, any label is acceptable as long as the column locations in each record are consistent. As with the Journey-to-crime calibration file, other fields can be included.

Table 14.3:

Example Records in Bayesian Journey-to-crime Diagnostics File

OffenderID	HomeX	HomeY	IncidX	IncidY
Num 1	-77.1496	39.3762	-76.6101	39.3729
Num 1	-77.1496	39.3762	-76.5385	39.3790
Num 1	-77.1496	39.3762	-76.5240	39.3944
Num 2	-76.3098	39.4696	-76.5427	39.3989
Num 2	-76.3098	39.4696	-76.5140	39.2940
Num 2	-76.3098	39.4696	-76.4710	39.3741
Num 3	-76.7104	39.3619	-76.7195	39.3704
Num 3	-76.7104	39.3619	-76.8091	39.4428
Num 3	-76.7104	39.3619	-76.7114	39.3625
Num 4	-76.5179	39.2501	-76.5144	39.3177
Num 4	-76.5179	39.2501	-76.4804	39.2609
Num 4	-76.5179	39.2501	-76.5099	39.2952
Num 5	-76.3793	39.3524	-76.4684	39.3526
Num 5	-76.3793	39.3524	-76.4579	39.3590
Num 5	-76.3793	39.3524	-76.4576	39.3590
Num 5	-76.3793	39.3524	-76.4512	39.3347
Num 6	-76.5920	39.3719	-76.5867	39.3745
Num 6	-76.5920	39.3719	-76.5879	39.3730
Num 6	-76.5920	39.3719	-76.7166	39.2757
Num 6	-76.5920	39.3719	-76.6015	39.4042
Num 7	-76.7152	39.3468	-76.7542	39.2815
Num 7	-76.7152	39.3468	-76.7516	39.2832
Num 7	-76.7152	39.3468	-76.7331	39.2878
Num 7	-76.7152	39.3468	-76.7281	39.2889
.				
.				
.				
.				
Num Last	-76.4320	39.3182	-76.4297	39.3172
Num Last	-76.4880	39.3372	-76.4297	39.3172
Num Last	-76.4437	39.3300	-76.4297	39.3172
Num Last	-76.4085	39.3342	-76.4297	39.3172
Num Last	-76.4083	39.3332	-76.4297	39.3172
Num Last	-76.4082	39.3324	-76.4297	39.3172
Num Last	-76.4081	39.3335	-76.4297	39.3172

Logic of the Routine

The module is divided into two parts (under the "Bayesian Journey-to-crime Estimation" page of "Spatial Modeling"):

1.   Diagnostics for Journey-to-crime methods; and

2.   Estimate likely origin location of a serial offender.

The "diagnostics" routine takes the diagnostics calibration file and estimates a number of methods for each serial offender in the file and tests the accuracy of each parameter against the known residence location. The result is a comparison of the different methods in terms of accuracy in predicting both where the offender lives as well as minimizing the distance between where the method predicts the most likely location for the offender and where the offender actually lives.

The "estimate" routine allows the user to choose one method and to apply it to the data for a *single* serial offender. The result is a probability surface showing the results of the method in predicting where the offender is liable to be living.

Bayesian Journey-to-crime Diagnostics

The following applies to the Bayesian Journey-to-crime (BJtc) Diagnostics routine only.

Data Input

The user inputs the four required data sets.

1.   Any primary file with an X and Y location. A suggestion is to use the file for the one of the serial offenders, but this is not essential;

2.   A grid that will be overlaid on the study area. Use the Reference File under Data Setup to define the X and Y coordinates of the lower-left and upper-right corners of the grid as well as the number of columns (see Chapter 3, p. 3.21);

3.   A Journey-to-crime travel function (Jtc) that estimates the likelihood of an offender committing crimes at various distances or travel times if a network is used (see Chapter 13).

4.   An observed origin-destination matrix (see Chapter 28, p. 28.17); and

5.   A diagnostics file of known serial offenders in which both their residence and crime locations are known (BJtc Diagnostics)

Methods Tested

The BJtc Diagnostics routine compares six methods for estimating the likely location of a serial offender and will include up to four additional methods if filters are used (see below).

1.   The Jtc distance method, P(Jtc);

2.   The general crime distribution based on the origin-destination matrix, P(O). Essentially, this is the distribution of origins irrespective of the destinations;

3.   The distribution of origins in the O-D matrix based only on the incidents in zones that are identical to those committed by the serial offender, P(O|Jtc);

4.   The product of the Jtc estimate (1 above) and the distribution of origins based only on those incidents committed in zones identical to those by the serial offender (3 above), P(Jtc)*P(O|Jtc). This is the numerator of the Bayesian function (Equation 14.8), the product of the prior probability times the likelihood estimate;

5.   The Bayesian Risk estimate as indicated in Equation 14.8 above (method 4 above divided by method 2 above), P(Bayesian). This is a rough approximation to the full Bayesian function in Equation 14.12 above; and

6.   The center of minimum distance, Cmd. Previous research has shown that the center of minimum of distance has the least error in minimizing the distance between the most likely location for the offender and where the offender actually lives (Paulsen, 2006a; Snook, Zito, Bennell, & Taylor, 2005; Levine, 2000).

7.   If filters are used: The product of the Jtc estimate and the filters, P(Jtc)*F1*F2.

8.   If filters are used: The product of the Conditional estimate and the filters, P(O|Jtc)*F1*F2.

9.   If filters are used: The product of the "Product" estimate and the filters, P(Jtc)*P(O|Jtc)*F1*F2.

10.   If filters are used: The product of the Bayesian Risk estimate and the filters, P(Bayesian)*F1*F2.

Interpolated Grid

For each serial offender in turn in the BJtc Diagnostics file and for each method, the routine overlays a grid over the study area. The grid is defined by the Reference File parameters (see Chapter 3). The routine then interpolates each input data set into a probability estimate for each grid cell with the sum of the cells equaling 1.0 (within three decimal places).

14.17

The manner in which the interpolation is done varies by the method:

1.   For the Jtc method, P(Jtc), the routine interpolates the selected distance function to each grid cell to produce a density estimate. The densities are then re-scaled so that the sum of the grid cells equals 1.0 (see Chapter 10 on kernel density interpolation);

2.   For the general crime distribution method, P(O), the routine sums up the incidents by each origin zone from the origin-destination matrix and interpolates that using the normal distribution method of the single kernel density routine (see Chapter 10 on kernel density interpolation). The density estimates are converted to probabilities so that the sum of the grid cells equals 1.0;

3.   For the distribution of origins based only on the incidents committed by the serial offender, from the origin-destination matrix the routine identifies the zone in which the incidents occurred and reads only those origins associated with those destination zones. Multiple incidents committed in the same origin zone are counted multiple times. The routine adds up the number of incidents counted for each zone and uses the single kernel density routine to interpolate the distribution to the grid (see Chapter 10 on kernel density interpolation). The density estimates are converted to probabilities so that the sum of the grid cells equals 1.0;

4.   For the product of the Jtc estimate and the distribution of origins based only on the incidents committed by the serial offender, the routine multiples the probability estimate obtained in 1 above by the probability estimate obtained in 3 above. The probabilities are then re-scaled so that the sum of the grid cells equals 1.0;

5.   For the Bayesian Risk estimate, the routine takes the product estimate (4 above) and divides it by the general crime distribution estimate (2 above). The resulting probabilities are then re-scaled so that the sum of the grid cells equals 1.0; and

6.   For the center of minimum distance estimate, the routine calculates the center of minimum distance for each serial offender in the "diagnostics" file and calculates the distance between this statistic and the location where the offender is actually residing. This is used only for the distance error comparisons.

7.   For the interaction of the Jtc, Conditional, "Product" and Bayesian Risk estimates with the filters, the estimates are obtained by multiplying the filters times these terms and then re-scaling so that the sum of the grid cells equals 1.0.

14.18

Note in all of the probability estimates (excluding 6), the cells are converted to probabilities prior to any multiplication or division. The results are then re-scaled so that the resulting grid is a probability (i.e., all cells sum to 1.0).

Output

For each offender in the BJtc Diagnostics file, the routine calculates three different statistics for **each** of the methods:

1. The estimated probability in the cell where the offender actually lives. It does this by, first, identifying the grid cell in which the offender lives (i.e., the grid cell where the offender's residence X and Y coordinate is found) and, second, by noting the probability associated with that grid cell;

2. The percentile of all grid cells in the entire grid that have to be searched to find the cell where the offender lives based on the probability estimate from 1 above, ranked from those with the highest probability to the lowest. Obviously, this percentage will vary by how large a reference grid is used (e.g., with a very large reference grid, the percentile where the offender actually lives will be small whereas with a small reference grid, the percentile will be larger). But, since the purpose is to compare methods, the actual percentage should be treated as a relative index.

    The result is sorted from low to high so that the smaller the percentile, the better. For example, a percentile of 1% indicates that the probability estimate for the cell where the offender lives is within the top 1% of all grid cells. Conversely, a percentile of 30% indicates that the probability estimate for the cell where the offender lives in within the top 30% of all grid cells; and

3. The distance between the cell with the highest probability and the cell where the offender lives.

These three indices provide information about the accuracy and precision of the method. Table 14.4 illustrates a typical probability output for four of the methods. Only five serial offenders are shown in the table.

Output matrices

The BJtc Diagnostics routine outputs two separate matrices. The probability estimates (numbers 1 and 2 above) are presented in a separate matrix from the distance estimates (number

Table 14.4:
Sample Output of Probability Matrix

Offender	P(Jtc)	Percentile for P(Jtc)	P(O|Jtc)	Percentile for P(O|Jtc)	P(O)	Percentile for P(O)	P(Jtc)*P(O|Jtc)	Percentile for P(Jtc)*P(O|Jtc)
1	0.001169	0.01%	0.000663	0.01%	0.0003	11.38%	0.002587	0.01%
2	0.000292	5.68%	0.000483	0.12%	0.000377	0.33%	0.000673	0.40%
3	0.000838	0.14%	0.000409	0.18%	0.0002	30.28%	0.00172	0.10%
4	0.000611	1.56%	0.000525	1.47%	0.0004	2.37%	0.000993	1.37%
5	0.001619	0.04%	0.000943	0.03%	0.000266	11.98%	0.004286	0.04%

Table 14.5 illustrates a typical distance output for four of the methods. Only five serial offenders are shown in the table.

Table 14.5:
Sample Output of Distance Matrix

Offender	Distance(Jtc)	Distance(O|Jtc)	Distance(O)	Distance for P(Jtc)*P(O|Jtc)
1	0.060644	0.060644	7.510158	0.060644
2	6.406375	0.673807	2.23202	0.840291
3	0.906104	0.407762	11.53447	0.407762
4	3.694369	3.672257	2.20705	3.672257
5	0.423577	0.405526	6.772228	0.423577

3 above). The user can save the total output as a text file or can copy and paste each of the two output matrices into a spreadsheet separately. We recommend the copying-and-pasting method into a spreadsheet as it will be difficult to line up differing column widths for the two matrices and summary tables in a text file.

Which is the Most Accurate and Precise Method?

Accuracy and precision are two different criteria for evaluating a method. With accuracy, one wants to know how close a method comes to a target. The target can be an exact location (e.g., the residence of a serial offender) or it can be a zone (e.g., a high probability area within which the serial offender lives). Precision, on the other hand, refers to the consistency of the method, irrespective of how accurate it is. A more precise measure is one in which the method has a limited variability in estimating the central location whereas a less precise measure has a higher degree of variability. These two criteria - accuracy and precision, often conflict.

The following example is from Jessen (1979). Consider a target that one is trying to 'hit' (Figure 14.5A). The target can be a physical target, such as a dart board, or it can be a location in space, such as the residence of a serial offender. One can think of three different 'throwers' or methods attempting to hit the center of target, the Bulls Eye. The throwers make repeated attempts to hit the target and the 'throws' (or estimates from the method) can be evaluated in terms of accuracy and precision. In Figure 14.5B, the thrower is all over the dartboard. There is no consistency at all. However, if the center of minimum distance (Cmd) is calculated, it is very close to the actual center of the target, the Bulls Eye. In this case, the thrower is accurate but not precise. That is, there is no systematic bias in the thrower's throws, but they are not reliable. This thrower is accurate (or unbiased) but not precise.

In Figure 14.5C, there is an opposite condition. In this case, the thrower is precise but not accurate. That is, there is a systematic bias in the throws even though the throws (or method) are relatively consistent.   Finally in Figure 14.5D, the thrower is both relatively precise and accurate as the Cmd of the throws is almost exactly on the Bulls Eye.

One can apply this analogy to a method. A method produces estimates from a sample. For each sample, one can evaluate how accurate is the method (i.e., how close to the target did it come) and how consistent is it (how much of variability does it produce). Perhaps the analogy is not perfect because the thrower makes multiple throws whereas the method produces a single estimate.   But, clearly, we want a method that is both accurate and precise.
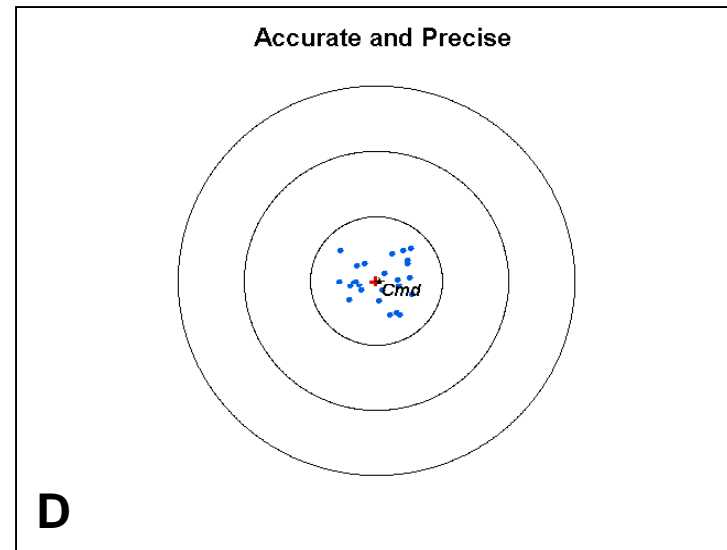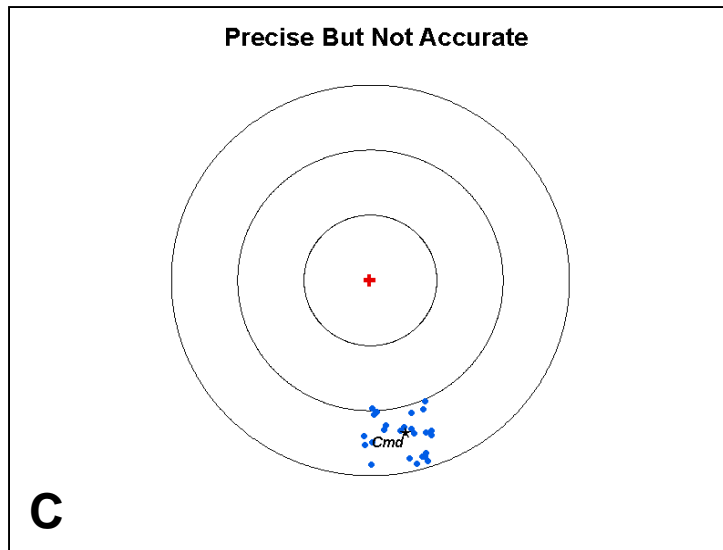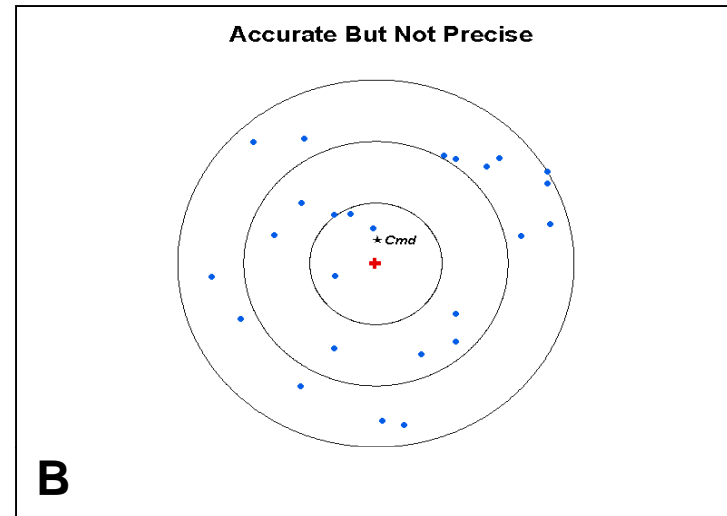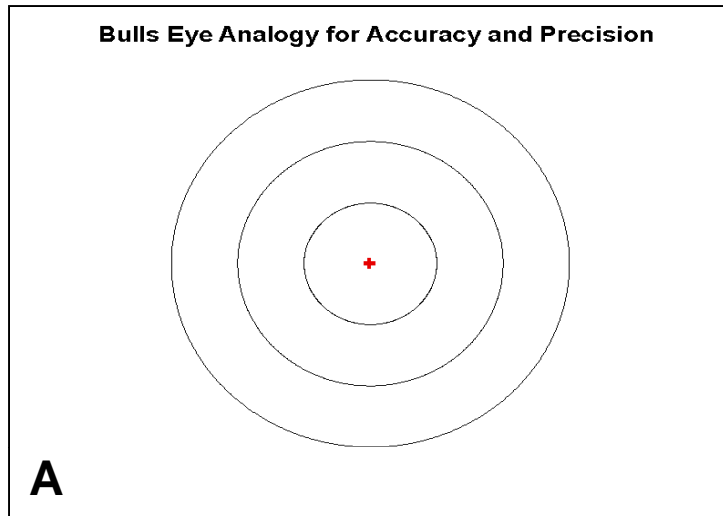
Measures of Accuracy and Precision

Much of the debate in the area of Journey-to-crime estimation has revolved around arguments about the accuracy and precision of the method. Levine (2000) first raised the issue of accuracy by proposing distance from the location with the highest probability to the location where the offender lived as a measure of accuracy, and suggested that simple, centrographic measures were as accurate as more precise Journey-to-crime methods in estimating this. Paulsen (2006a; 2006b) confirmed that centrgraphic methods were more accurate than journey-to-crime method. Snook and colleagues also confirmed this conclusion and showed that human subjects could do as well as any of the algorithms (Snook, Zito, Bennell, & Taylor, 2005; Snook, Taylor & Bennell, 2004).

On the other hand, Canter, Coffey and Missen (2000), Canter (2003), and Rossmo (2000) have argued for an area of highest probability being the criterion for evaluating accuracy, indicating a 'search cost' or a 'hit score' with the aim being to narrow the search area to as small

Figure 14.5:
Accuracy and Precision in Estimates

Bulls Eye Analogy for Accuracy and Precision

A

Accurate But Not Precise

★ Cmd

B

Precise But Not Accurate

Cmd

C

Accurate and Precise

Cmd

D

as possible. Rich and Shivley (2004) compared different Journey-to-crime/geographic profiling software packages and concluded that there were at least five different criteria for evaluating accuracy and precision - error distance, search cost/hit score, profile error distance, top profile area, and profile accuracy.

Rossmo (2005a; b) and Rossmo and Filer (2005) have critiqued these measures as being too simple and have rejected error distance. Levine (2005) justified the use of error distance as being fundamental to statistical error while acknowledging that an area measure is necessary, too.

While the debate continues to develop, practically a distinction can be made between measures of accuracy and measures of precision. Accuracy is measured by how close to the target is the estimate while precision refers to how large or small an area the method produces. The two become identical when the precision is extremely small, similar to a variance converging into a mean as the distance between observations and the mean approach zero.

In evaluating the methods, five different measures are used:

Accuracy Measures

1. **True accuracy** - the probability in the cell where the offender actually lives. The Bayesian Jtc diagnostics routine evaluates the six above mentioned methods on a sample of serial offenders with known residence address. Each of the methods (except for the center of minimum distance, Cmd) has a probability distribution. That method which has the highest probability in the cell where the offender lives is the most accurate.

2. **Diagnostic accuracy** - the distance between the cell with the highest probability estimate and the cell where the offender lives. Each of the methods produces probability estimates for each cell. The cell with the highest probability is the best guess for where the offender lives. The distance from this location to where the offender lives is an indicator of the diagnostic accuracy of the method.

3. **Neighborhood accuracy** - the percent of offenders who reside within the cell with the highest probability. Since the grid cell is the smallest unit of resolution, this measures the percent of all offenders who live at the highest probability cell. This was estimated by those cases where the error distance was smaller than half the grid cell size.

Precision Measures

4. **Search cost/hit score** - the percent of the total study area that has to be searched to find the cell where the offender actually lived after having sorted the output cells from the highest probability to the lowest

5. **Potential search cost** - the percent of offenders who live within a specified distance of the cell with the highest probability. In this evaluation, two distances are used though others can certainly be used:

    A. The percent of offender who live within one mile of the cell with the highest probability.

    B. The percent of offenders who live within one-half mile of the cell with the highest probability ("Probable search area in miles").

Summary Statistics

The "diagnostics" routine will also provide summary information at the bottom of each matrix. There are summary measures and counts of the number of times a method had the highest probability or the closest distance from the cell with the highest probability to the cell where the offender actually lived; ties between methods are counted as fractions (e.g., two tied methods are given 0.5 each; three tied methods are give 0.33 each). For the probability matrix, these statistics include:

1. The mean (probability or percentile);
2. The median (probability or percentile);
3. The standard deviation (probability or percentile);
4. The number of times the P(Jtc) estimate produces the highest probility;
5. The number of times the P(O|Jtc) estimate produces the highest probability;
6. The number of times the P(O) estimate produces the highest probability;
7. The number of times the product term estimate produces the highest probability;
8. The number of times the Bayesian estimate produces the highest probability.
9. If filters are used: The number of times the Jtc, Conditional, Product, and Bayesian Risk estimates times the filters produce the highest probability.

For the distance matrix, these statistics include:

1. The mean distance;
2. The median distance;
3. The standard deviation distance;
4. The number of times the P(Jtc) estimate produces the closest distance;
5. The number of times the P(O|Jtc) estimate produces the closest distance;
6. The number of times the P(O) estimate produces the closest distance;
7. The number of times the product term estimate produces the closest distance;
8. The number of times the Bayesian Risk estimate produces the closest distance; and
9. The number of times the CMD produces the closest distance.
10. If filters are used: The number of times the Jtc, Conditional, Product, and Bayesian Risk estimates times the filters produce the closest distance.

Testing the Routine with Serial Offenders from Baltimore County

To illustrate the use of the Bayesian Jtc diagnostics routine, the records of 88 serial offenders who had committed crimes in Baltimore County, MD, between 1993 and 1997 were compiled into a diagnostics file. The number of incidents committed by these offenders varied from 3 to 33 and included a range of different crime types (larceny, burglary, robbery, vehicle theft, arson, and bank robbery).

Because the methods are interdependent, traditional parametric statistical tests cannot be used. Instead, non-parametric tests have been applied. For the probability and distance measures, two tests were used. First, the Friedman two-way analysis of variance test examines differences in the overall rank orders of multiple measures (treatments) for a group of subjects (Kanji, 1993, 115; Siegel, 1956). This is a Chi-square test and measures whether there are significant differences in the rank orders across all measures (treatments). Second, differences between specific pairs of measures can be tested using the Wilcoxon matched pairs signed-ranks test (Siegel, 1956, 75-83). This examines pairs of methods by not only their rank, but also by the difference in the values of the measurements.

For the percentage of offenders who lived in the same grid cell, within one mile, and within one half-mile of the cell with the peak likelihood, the Cochran Q test for k related samples was used to test differences among the methods (Kanji, 1993, 74; Siegel, 1956, 161-166). This is a Chi-square test of whether there are overall differences among the methods in the percentages, but cannot indicate whether any one method has a statistically higher percentage. Consequently, the method with the highest percentage was tested against the method with the second highest percentage using the Cochran Q test to see whether the best method stood out.

Results: Accuracy

Table 14.6 presents the results three accuracy measures. For the first measure, the probability estimate in the cell where the offender actually lived, the product probability is far superior to any of the others. It has the highest mean probability of any of the measures and is more than double the probability of the Journey-to-crime method. The Friedman test indicates that these differences are significant and the Wilcoxon matched pairs test indicates that the product has a significantly higher probability than the second best measure, the Bayesian Risk, which in turn is significantly higher than the Journey-to-crime measure. At the low end, the general probability has the lowest average and is significantly lower than the other measures.

In terms of the individual offenders, the product probability had the highest probability for 74 of the 88 offenders. For the other methods, the Bayesian Risk measure had the highest probability for 10 offenders, the Journey-to-crime measure for one offender, the conditional probability for two offenders and the general probability for one offender.

Table 14.6:
Accuracy Measures of Total Sample

Method	Mean probability in offender cell[a]	Mean distance from highest probability cell to offender cell (mi)[b]	Percent of offenders whose residence is in highest prob. cell[c]
Journey-to-crime	0.00082	2.78	12.5%
General	0.00025	8.21	0.0%
Conditional	0.00052	3.22	3.4%
Product	**0.00170**	2.65	13.6%
Bayesian Risk	0.00131	3.15	10.2%
Cmd	n.a.	**2.62**	**18.2%**

a   Friedman χ^2 =236.0 ; d.f. = 4; p≤.001; Wilcoxon signed-ranks test at p≤.05: Product > Bayesian Risk > JTC = Conditional > General

b   Friedman χ^2 = 114.2; d.f. = 5; p≤.001; Wilcoxon signed-ranks test at p≤.05: CMD = Product = JTC > Bayesian Risk = Conditional < General

c   Cochran Q=33.9, d.f. =5, p≤.001; Cochran Q of difference between best & second best=1.14, n.s.

Finally, for the third accuracy measure, the percent of offenders residing in the area covered by the cell with the highest probability estimate, the Cmd has the highest percentage (18.2%) followed by the product probability (13.6%), and the Journey-to-crime probability (12.5%). The Cochran Q shows significant differences over all these measures. However, the difference between the measure with the highest percentage in the same grid cell (the Cmd) and the measure with the second highest percentage (the product probability) is not significant.

For accuracy, the product probability appears to be better than the Journey-to-crime estimate and almost as accurate as the Cmd. It has the highest probability in the cell where the offender lived and a lower error distance than the Journey-to-crime method (though not significantly so). Finally, it had a slightly higher percentage of offenders living in the area covered by cell with the highest probability than for the Journey-to-crime.

The Cmd, on the other hand, which had been shown to be the most accurate in previous studies (Paulsen, 2006a; 2006b; Snook, Zito, Bennell, and Taylor, 2005; Snook, Taylor and Bennell, 2004; Levine, 2000), does not appear to be more accurate than the product probability. It has only a slightly lower error distance and a slightly higher percentage of offenders residing in the area covered by the cell with the highest probability. Thus, the product term has equaled the Cmd in terms of accuracy. Both, however, are more accurate than the Journey-to-crime estimate.

For the measure of diagnostic accuracy (the distance from the cell with the highest probability to the cell where the offender lived), the center of minimum distance (Cmd) had the lowest error distance followed closely by the product term. The Journey-to-crime method had a slightly larger error. Again, the general probability had the greatest error, as might be expected. The Friedman test indicates there are overall differences among the six measures in the mean distance. The Wilcoxon signed-ranks test, however, showed that the Cmd, the product, and the Journey-to-crime estimates are not significantly different, though all are significantly lower than the Bayesian Risk measure and the conditional probability which, in turn, are significantly lower than the general probability.

In terms of individual cases, the Cmd produced the lowest average error distance for 30 of the 88 cases while the conditional term (O|Jtc) had the lowest error distance in 17.9 cases (including ties). The product term produced a lower average error distance for 9.5 cases (including ties) and the Jtc estimate produced lower average distance errors in 8.2 cases (again, including ties). In other words, the Cmd will either be very accurate or very inaccurate, which is not surprising given that it is only a point estimate.

Results: Precision

Table 14.7 presents the three precision measures used to evaluate the six different measures. For the first measure, the mean percent of the study area with a higher probability (what Canter called 'search cost' and Rossmo called 'hit score'; Canter, 2003; Rossmo, 2005a, 2005b), the Bayesian Risk measure had the lowest percentage followed closely by the product term. The conditional probability was third followed by the Journey-to-crime probability followed by the general probability. The Friedman test indicates that these differences are significant overall and the Wilcoxon test shows that the Bayesian Risk, product term, conditional probability and Journey-to-crime estimates are not significantly different from each other. The general probability estimate, however, is much worse.

In terms of individual cases, the product probability had either the lowest percentage or was tied with other measures for the lowest percentage in 36 of the 88 cases. The Bayesian Risk and Journey-to-crime measures had the lowest percentage or were tied with other measures for the lowest percentage in 34 of the 88 cases. The conditional probability had the lowest percentage or was tied with other measures for the lowest percentage in 23 of the cases. Finally, the general probability had the lowest percentage or was tied with other measures for the lowest percentage in only 7 of the cases.

Similar results are seen for the percent of offenders living within one mile of the cell with the highest probability and also for the percent living within a half mile. For the percent within one mile, the product term had the highest percentage followed closely by the Journey-to-crime measure and the Cmd. Again, at the low end is the general probability. The Cochran Q test indicates that these differences are significant over all measures though the difference between the best method (the product) and the second best (the Journey-to-crime) is not significant.

Conclusion of the Evaluation

In conclusion, the product method appears to be an improvement over the Journey-to-crime method, at least with these data from Baltimore County. It is substantially more accurate and about as precise. Further, the product probability appears to be, on average, almost as accurate as the Cmd, though the Cmd still is more accurate in assessing the exact location of offenders. That is, the Cmd will identify about one-sixth of all offenders exactly. For a single guess of where a serial offender is living, the center of minimum distance produced the lowest distance error. But, since it is only a point estimate, it cannot point to a search area where the offender might be living. The product term, on the other hand, produced an average distance error almost as small as the center of minimum distance, but produced estimates for other grid cells too. Among all the probability measures, it had the highest probability in the cell where the offender lived and was among the most efficient in terms of reducing the search area.

Table 14.7:
Precision Measures of Total Sample

Method	Mean percent of study area with higher probability[a]	Percent of offenders living within distance of highest probability cell:	
		1 mile[b]	0.5 miles[c]
Journey-to-crime	4.7%	56.8%	44.3%
General	16.8%	2.3%	0.0%
Conditional	4.6%	47.7%	31.8%
Product	4.2%	**59.1%**	**48.9%**
Bayesian Risk	**4.1%**	51.1%	42.0%
Cmd	n.a.	54.5%	42.0%

a   Friedman χ^2 =115.4 ; d.f. = 4; p<.001; Wilcoxon signed-ranks test at p<.05: Bayesian Risk =Product= JTC = Conditional> General

b   Cochran Q = 141.0, d.f. = 5, p<.001; Cochran Q of difference between best and second best = 0.7, n.s.

c   Cochran Q = 112.2, d.f. = 5, p<.001; Cochran Q of difference between best and second best = 2.0, n.s.

In other words, using information about the origin location of other offenders appears to improve the accuracy of the Jtc method. The result is an index (the product term) that is almost as good as the center of minimum distance, but one that is more useful since the center of minimum distance is only a single point.

Of course, each jurisdiction should re-run these diagnostics to determine the most appropriate measure. It is very possible that other jurisdictions will have different results due to the uniqueness of their land uses, street layout, and location in relation to the center of the metropolitan area. Baltimore County surrounds the City of Baltimore on three sides. It has a mixture of neighborhoods including parts of the central city, older suburbs, newer suburbs, separate communities and rural areas. The model results which fit Baltimore County might not fit other places.

Tests with Other Data Sets

The Bayesian Journey-to-crime model was tested in 2009 in several jurisdictions:

1.   In Baltimore County with 850 serial offenders (Leitner & Kent, 2009);

2.   In the Hague, Netherlands with 62 serial burglars (Block & Bernasco, 2009);

3.   In Chicago, with 103 serial robbers (Levine & Block, 2011); and

4.   In Manchester, England with 171 serial offenders (Levine & Lee, 2009).

In all cases, the product probability measure was both more accurate and more precise than the Journey-to-crime measure. In two of the studies (Chicago and the Hague), the product term was also more accurate than the Center of Minimum Distance. In the other two studies (Baltimore County and Manchester), the Center of Minimum Distance was slightly more accurate than the product term.

Among the probability methods, the product term was more accurate than all other measures for three of the studies (Baltimore County, Chicago, Manchester). For the Hague study, however, the conditional estimate was more accurate. This was because the journey-to-crime estimate was very inaccurate due to the small size of the Hague (Block & Bernasco, 2009).

The mathematics of these models has been explored by O'Leary (2009). These studies are presented in a special issue of the *Journal of Investigative Psychology and Offender Profiling*. Introductions are provided by Canter (2009) and Levine (2009).

In short, the product term appears to be almost as good a method as the Center of Minimum Distance and generally the best of the probability methods. However, users should first test whether this conclusion holds for their jurisdiction.

Estimate Likely Origin Location of a Serial Offender

The following applies to the Bayesian Jtc Estimate Likely Origin Location (BJtc) of a Serial Offender routine. Once the BJtc Diagnostic routine has been run and a preferred method selected, the next routine allows the application of that method to a ***single*** serial offender.

Data Input

The user inputs the three required data sets and a reference file grid:

1. The incidents committed by a single offender that we are interested in catching. This ***must*** be the Primary File;

2. A Jtc function that estimates the likelihood of an offender committing crimes at a certain distance (or travel time if a network is used). This can be either a mathematically-defined function or an empirically-derived one (see Chapter 13 on Journey-to-crime Estimation). In general, the empirically-derived function is slightly more accurate than the mathematically-defined one though the differences are not large;

3. An origin-destination matrix; and

4. The reference file also needs to be defined and should include all locations where crimes have been committed (see Reference File).

Selected Method

The BJtc routine interpolates the incidents committed by the serial offender to a grid, yielding an estimate of where the offender is liable to live. There are five standard methods that can be used and ten additional methods if filters are used. However, the user has to choose ***one*** of these:

1. The Jtc distance method, P(Jtc);

2. The general crime distribution based on the origin-destination matrix, P(O). Essentially, this is the distribution of origins irrespective of the destinations;

3. The conditional Jtc distance. This is the distribution of origins based only on the incidents committed by other offenders in the same zones as those committed by the serial offender, P(O|Jtc). This is extracted from the O-D matrix;

4. The product of the Jtc estimate (1 above) and the distribution of origins based only on the incidents committed by the serial offender (3 above), P(Jtc)*P(O|Jtc). This is the numerator of the Bayesian function (Equation 14.8), the product of the prior probability times the likelihood estimate; and

4. The Bayesian Risk estimate as indicated in Equation 14.12 above (method 4 above divided by method 2 above), P(Bayesian).

5.   If one filter is used: the interaction between the Jtc, the Conditional, the Product, and the Bayesian Risk measures with the one filter. Also, the filter by itself can be checked.

6.   If two filters are used: the interaction between the Jtc, the Conditional, the Product, and the Bayesian measures with the two filters. Also, the two filters by themselves can be checked.

As mentioned, however, the user must choose only <u>one</u> of these for estimation.

Interpolated Grid

For the BJtc method that is selected, the routine overlays a grid on the study area. The grid is defined by the reference file parameters (see Chapter 3). The routine then interpolates the input data set (the primary file) into a probability estimate for each grid cell with the sum of the cells equaling 1.0 (within three decimal places). The manner in which the interpolation is done varies by the method chosen:

1.   For the Jtc method, P(Jtc), the routine interpolates the selected distance function to each grid cell to produce a density estimate. The density estimates are converted to probabilities so that the sum of the grid cells equals 1.0 (see Chapter 10 on kernel density interpolation);

2.   For the general crime distribution method, P(O), the routine sums up the incidents by each origin zone and interpolates this to the grid using the normal distribution method of the single kernel density routine (see Chapter 10 on kernel density interpolation). The density estimates are converted to probabilities so that the sum of the grid cells equals 1.0.

3.   For the distribution of origins based only on the incident committed by the serial offender, the routine identifies the zone in which the incident occurs and reads only those origins associated with those destination zones in the origin-destination matrix. Multiple incidents committed in the same origin zone are counted multiple times. The routine then uses the single kernel density routine to interpolate the distribution to the grid (see Chapter 10). The density estimates are converted to probabilities so that the sum of the grid cells equals 1.0;

4.   For the product of the Jtc estimate and the distribution of origins based only on the incidents committed by the serial offender, the routine multiples the probability estimate obtained in <u>1</u> above by the probability estimate obtained in <u>3</u> above. The product probabilities are then re-scaled so that the sum of the grid cells equals 1.0; and

5.   For the full Bayesian estimate as indicated in equation 14.12 above, the routine takes the product estimate (<u>4</u> above) and divides it by the general crime distribution estimate (<u>2</u> above). The resulting density estimates are converted to probabilities so that the sum of the grid cells equals 1.0.

6.   For any of the interactions with filters, the routine takes the filters and interpolates them to a grid and converts the estimates to probabilities. If there is only filter, then this layer is interpolated to the grid and converted into probabilities. If there are two filters, each is first interpolated to the grid and converted into probabilities. Then the two probability interpolations are multiplied by each other. The routine then multiplies the resulting filter probability by the probability estimate of the selected measure (Jtc, Conditional, Product, or Bayesian Risk). The resulting multiplication product is then re-scaled so that sum of the grid cells equals 1.0.

Note in all estimates, the results are then re-scaled so that the resulting grid is a probability (i.e., all cells sum to 1.0).

Output

Once the method has been selected, the routine interpolates the data to the grid cell and outputs it as a 'shp', 'mif/mid', or Ascii file for display in a GIS program. The tabular output shows the probability values for each cell in the matrix and also indicates which grid cell has the highest probability estimate.

Accumulator Matrix

There is also an intermediate output, called the *accumulator matrix* which the user can save. This lists the number of origins identified in each origin zone for the specific pattern of incidents committed by the offender, prior to the interpolation to grid cells. That is, in reading the origin-destination file, the routine first identifies which zone each incident committed by the offender falls within. Second, it reads the origin-destination matrix and identifies which origin zones are associated with incidents committed in the particular destination zones. Finally, it sums up the number of origins by zone ID associated with the incident distribution of the

offender. This can be useful for examining the distribution of origins by zones prior to interpolating these to the grid.

Two Examples of Using the Bayesian Journey-to-crime Routine

Two examples will illustrate the routine. Figure 14.6 presents the probability output for the general origin model, that is for the origins of all offenders irrespective of where they committed their crimes. It is a probability surface in that all the grid cells sum to 1.0. The map is scaled so that each bin covers a probability of 0.0001. The cell with the highest probability is highlighted in light blue.

As seen, the distribution is heavily weighted towards the center of the metropolitan area, particularly in the City of Baltimore. For the crimes committed in Baltimore County between 1993 and 1997 in which both the crime location and the residence location was known, about 40% of the offenders resided within the City of Baltimore and the bulk of those living within Baltimore County lived close to the border with City. In other words, as a general condition, most offenders in Baltimore County live relatively close to the center.

Offender S14A

The general probability output does not take into consideration information about the particular pattern of an offender. Therefore, we will examine specifically a particular offender. Figure 14.7 maps the distribution of an offender who committed 14 offenses between 1993 and 1997 (offender S14A) before being caught and the residence location where the individual lived when arrested

Of the 14 offenses, seven were thefts (larceny), four were assaults, two were robberies, and one was a burglary.   As seen, most of the incidents occurred in the southeast corner of Baltimore County though two incidents were committed more than five miles away from the offender's residence.

The general probability model is not very precise since it assigns the same probability to all grid cells for all offenders. In the case of offender S14A, the error distance between the cell with the highest probability and the cell where the offender actually lived was 7.4 miles.

On the other hand, the Jtc method uses the distribution of the incidents committed by a particular offender and a model of a typical travel distance distribution to estimate the likely origin of the offender's residence. A travel distance estimate based on the distribution of 41,424 offenders from Baltimore County was created using the *CrimeStat* Journey-to-crime calibration routine (see Chapter 13 on Journey-to-crime Estimation).

Figure 14.6:
Bayesian Journey-to-crime Routine
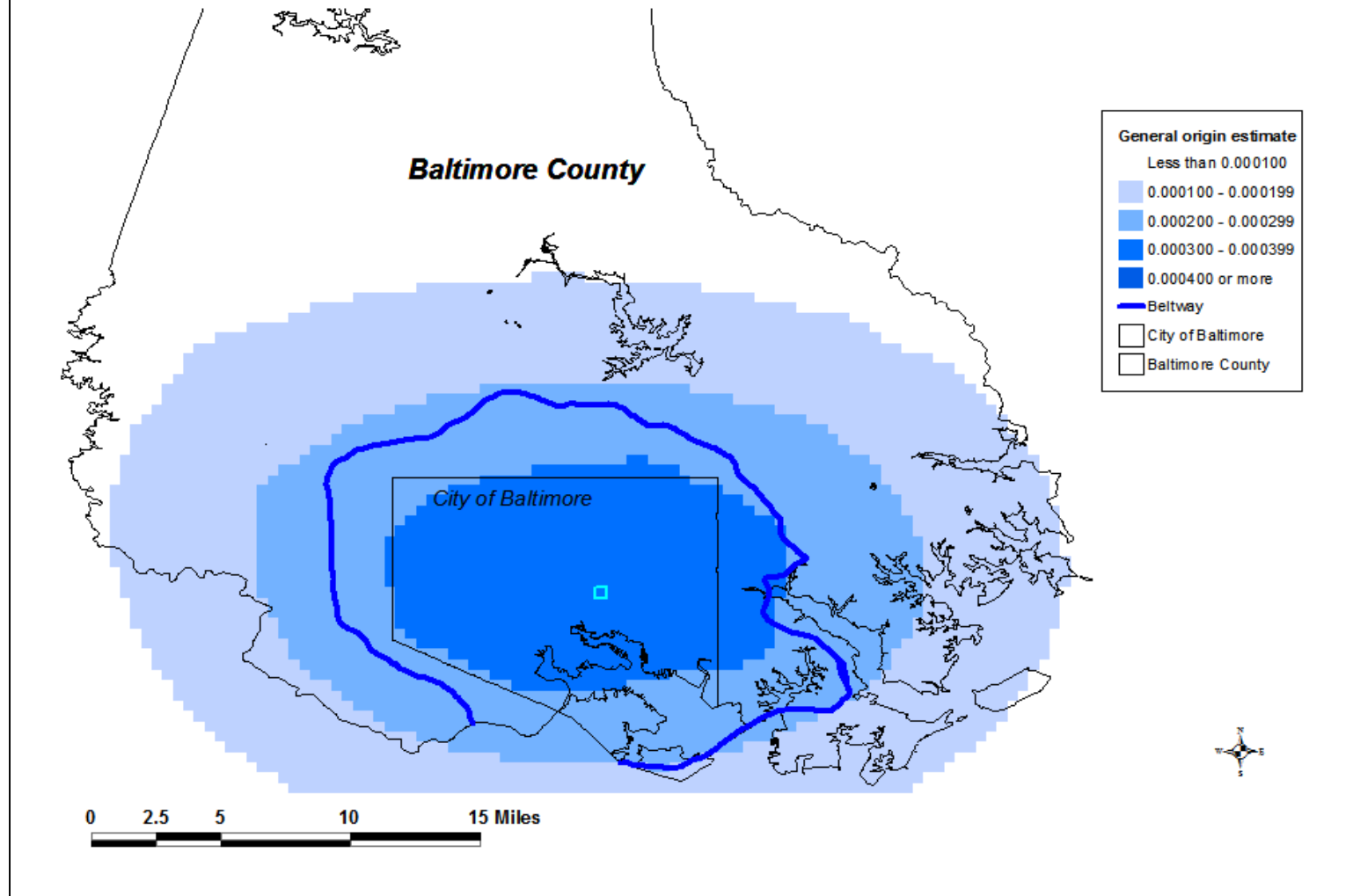General Distribution of Offenders by Residence Location

Figure 14.7:
Bayesian Journey-to-crime Routine
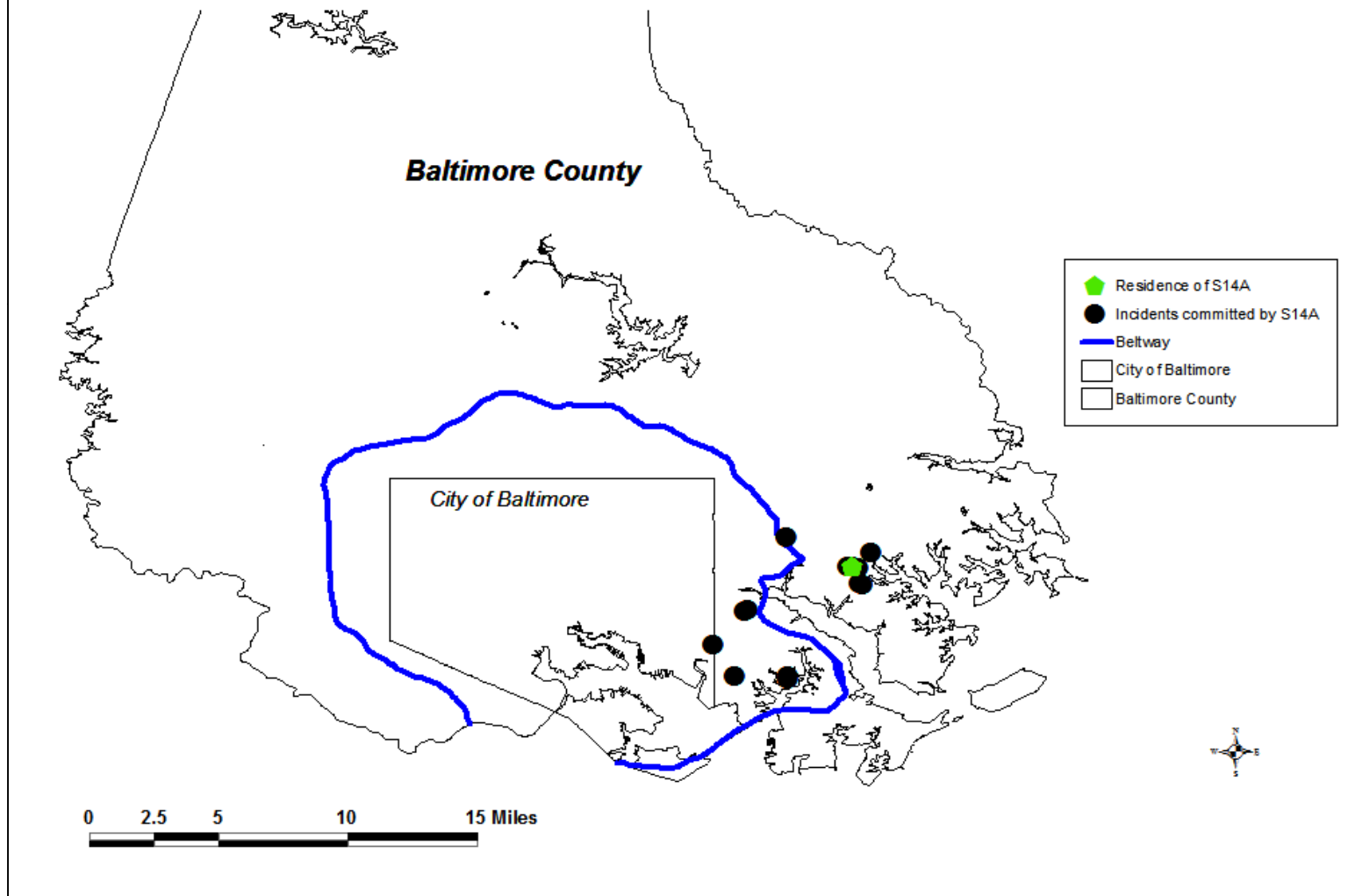Location of Incidents and Residence of Offender S14A

Figure 14.8 shows the results of the Jtc probability output. In this map and the following maps, the bins represent probability ranges of 0.0001. The cell with the highest likelihood is highlighted in light blue. As seen, this cell is very close to the cell where the actual offender lived. The distance between the two cells was 0.34 miles. With the Jtc probability estimate, however, the area with a higher probability (dark red) covers a fairly large area. However, the precision of the Jtc estimate is good since only 0.03% of the cells have higher probabilities that the cell associated with the area where the offender lived. In other words, the Jtc estimate has produced a very good estimate of the location of the offender, as might be expected given the concentration of the incidents committed by this person.

For this same offender, Figure 14.9 show the results of the conditional probability estimate of the offender's residence location, that is the distribution of the likely origin based on the origins of offenders who committed crimes in the same locations as that by S14A. Again, the cell with the highest probability is highlighted (in light green). As seen, this method has also produced a fairly close estimate, with the distance between the cell with the highest probability and the cell where the offender actually lived being 0.18 miles, about half the error distance of the Jtc method. Further, the conditional estimate is more precise than the Jtc with only 0.01% of the cells having a higher probability than the cell associated with the residence of hte offender. Thus, the conditional probability estimate is not only more accurate than the Jtc method, but also more precise (i.e., more efficient in terms of search area).

For this same offender, Figure 14.10 shows the results of the Bayesian product estimate, the product of the Jtc probability and the conditional probability re-scaled to be a single probability (i.e., with the sum of the grid cells equal to 1.0). It is a Bayesian estimate because it updates the Jtc probability estimate with the information on the likely origins of offenders who committed crimes in the same locations (the conditional estimate). Again, the cell with the highest probability is highlighted (in dark tan). The distance error for this method is 0.26 miles, not as accurate as the conditional probability estimate but more accurate than the Jtc estimate. Further, this method is about as precise as the Jtc since 0.03% of the cells having probabilities higher than that associated with the location where the offender lived.

Figure 14.11 shows the results of the Bayesian Risk probability estimate. This method takes the Bayesian product estimate and divides it by the general origin probability estimate. It is analogous to a *risk* measure that relates the number of events to a baseline population. In this case, it is the estimate of the probability of the updated Jtc estimate relative to the probability of where offenders live in general. Again, the cell with the highest likelihood is highlighted (in dark yellow). The Bayesian Risk estimate produces an error of 0.34 miles, the same as the Jtc estimate, with 0.04% of the cells having probabilities higher than that associated with the residence of the offender.

14.37

Figure 14.8:
Bayesian Journey-to-crime Routine
Predicated and Actual Residence Location of Offender S14A

Journey-to-crime Estimate

Baltimore County

City of Baltimore

Residence of S14A
Incidents committed by S14A
Cmd of S14A
Jtc estimate (S14A)
Less than 0.000100
0.000100 – 0.000199
0.000200 – 0.000299
0.000300 – 0.000399
0.000400 or more
Beltway
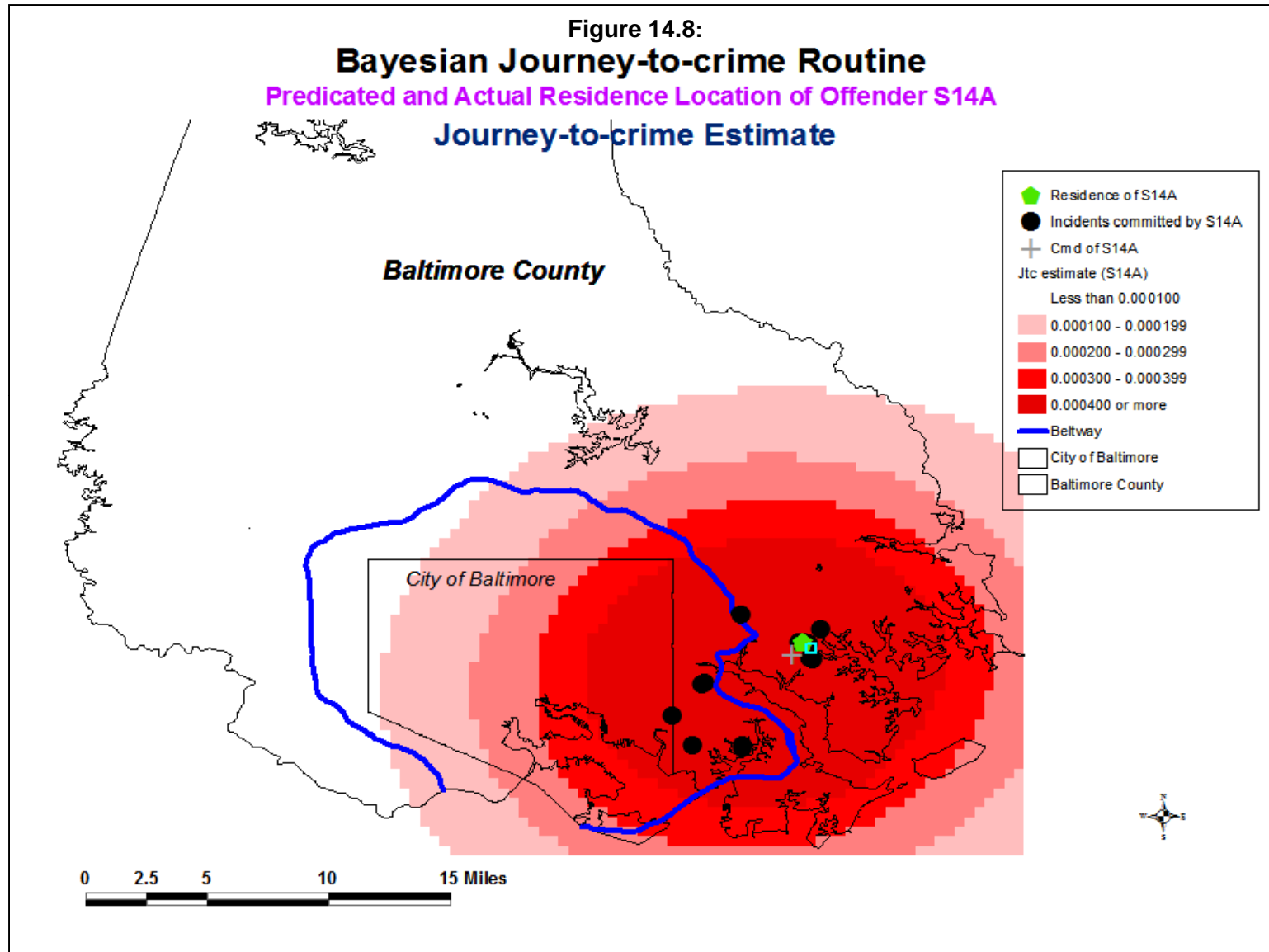City of Baltimore
Baltimore County

0   2.5   5   10   15 Miles

Figure 14.9:
Bayesian Journey-to-crime Routine
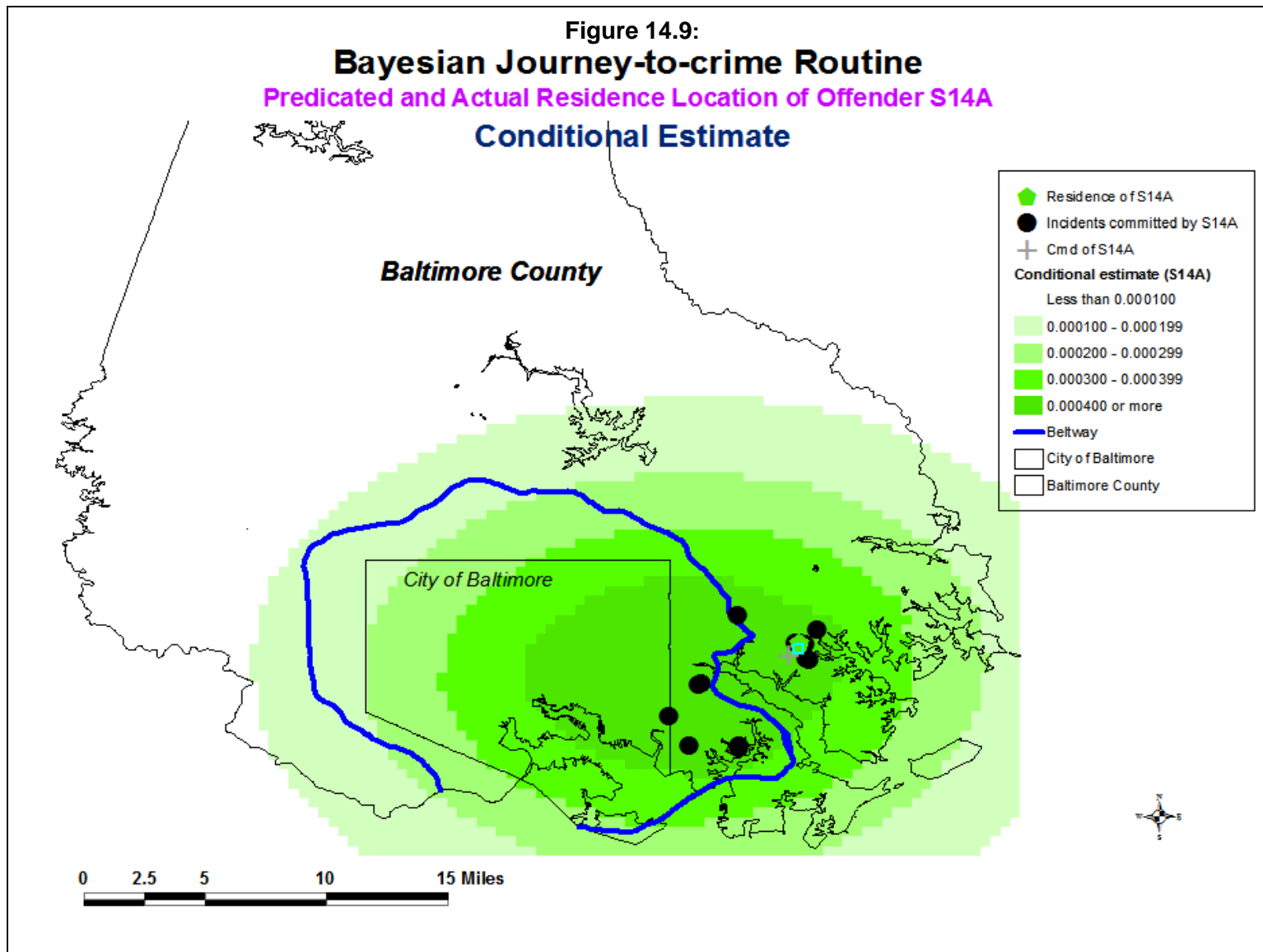Predicated and Actual Residence Location of Offender S14A
Conditional Estimate

Baltimore County

City of Baltimore

Residence of S14A
Incidents committed by S14A
Cmd of S14A
Conditional estimate (S14A)
Less than 0.000100
0.000100 – 0.000199
0.000200 – 0.000299
0.000300 – 0.000399
0.000400 or more
Beltway
City of Baltimore
Baltimore County

0   2.5   5   10   15 Miles

Figure 14.10:
Bayesian Journey-to-crime Routine
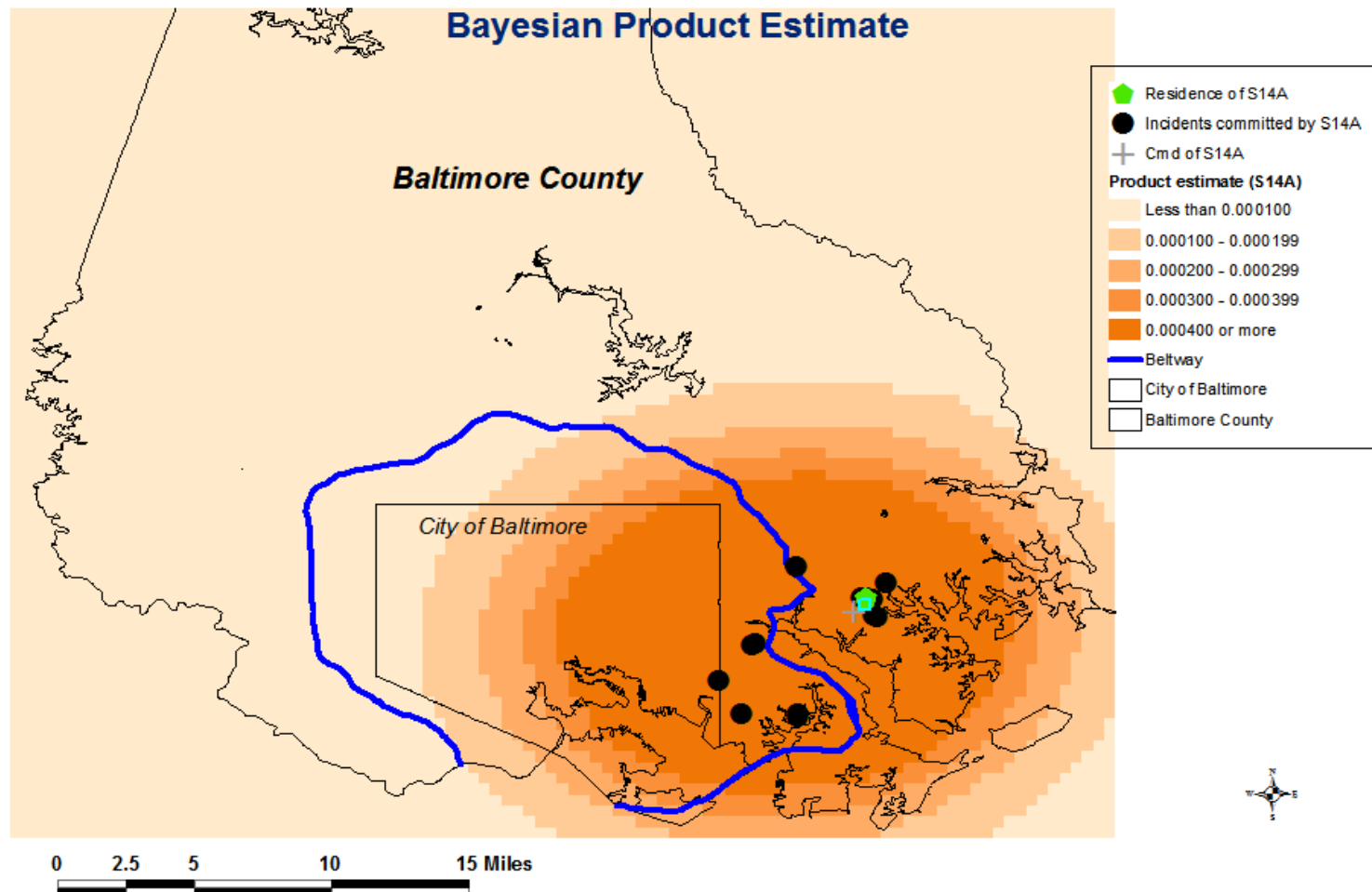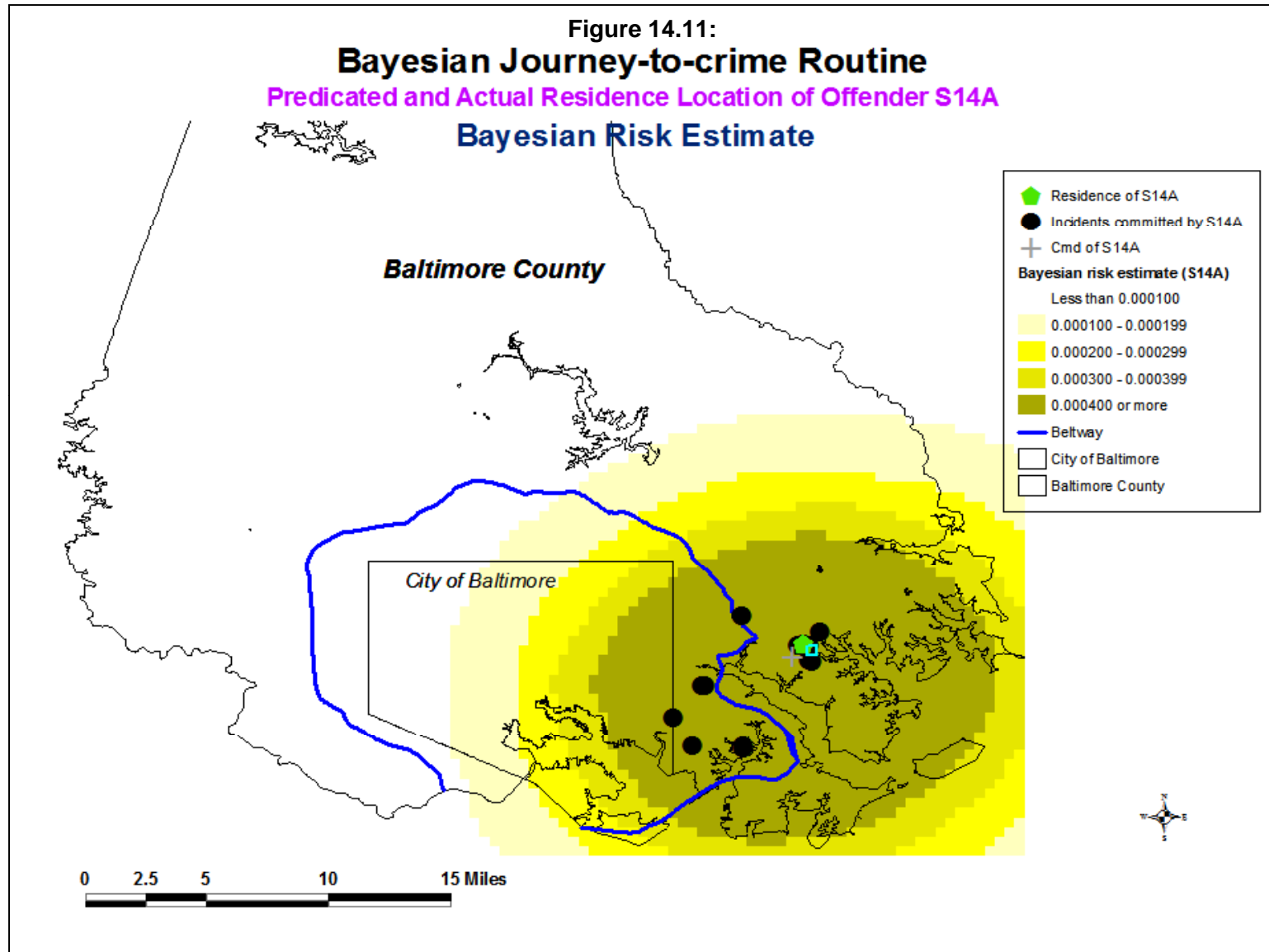Predicated and Actual Residence Location of Offender S14A

Figure 14.11:
Bayesian Journey-to-crime Routine
Predicated and Actual Residence Location of Offender S14A
Bayesian Risk Estimate

Baltimore County

City of Baltimore

Residence of S14A
Incidents committed by S14A
Cmd of S14A
Bayesian risk estimate (S14A)
Less than 0.000100
0.000100 - 0.000199
0.000200 - 0.000299
0.000300 - 0.000399
0.000400 or more
Beltway
City of Baltimore
Baltimore County

0   2.5   5   10   15 Miles

Finally, the center of minimum distance (Cmd) is indicated on each of the maps with a grey cross. In this case, the Cmd is not as accurate as any of the other methods since it has an error distance of 0.58 miles.

In summary, all of the Journey-to-crime estimate methods produced relatively accurate estimates of the location of the offender (S14A). Given that the incidents committed by this person were within a fairly concentrated pattern, it is not surprising that each of the method produced reasonable accuracy. In Canter and Larkin's (1994) terminology, this offender is a 'marauder'.

Offender TS15A

But what happens if an offender who did not commit crimes in the same part of town is selected, what Canter and Larkin (1994) call a 'commuter'? Figure 14.12 shows the distribution of an offender who committed 15 offenses (TS15A). Of the 15 offenses committed by this individual, there were six larceny thefts, two assaults, two vehicle thefts, one robbery, one burglary, and three incidents of arson. Twelve of the offenses are relatively concentrated but two are more than eight miles away.

Only three of the estimates will be shown. The general method produces an error of 4.6 miles. Figure 14.13 show the results of the Jtc method. Again, the map bins are in ranges of 0.0001 and the cell with the highest probability is highlighted. As seen, the cell with the highest probability is located north and west of the actual offender's residence. The error distance is 1.89 miles. The precision of this estimate is good with only 0.08% of the cells having higher probabilities than the cell where the offender lived.

Figure 14.14 show the result of the conditional probability estimate for this offender. In this case, the conditional probability method is less accurate than the Jtc method with an error distance between the cell with the highest probability and the cell where the offender lived being 2.39 miles. However, this method is less precise than the Jtc method with 1.6% of the study area having probabilities higher than that in the cell where the offender lived.

Finally, Figure 14.15 shows the results of the product probability estimate. For this method, the error distance is only 0.47 miles, much less than the Jtc method.   Further, it is smaller than the CMD which has an error distance of 1.33 miles. Again, updating the Jtc estimate with information from the conditional estimate produces a more accurate guess where the offender lives. Further, the product estimate is more precise with only 0.02% of the study area having probabilities higher than the cell covering the area where the offender lived.

Figure 14.12:
Bayesian Journey-to-crime Routine
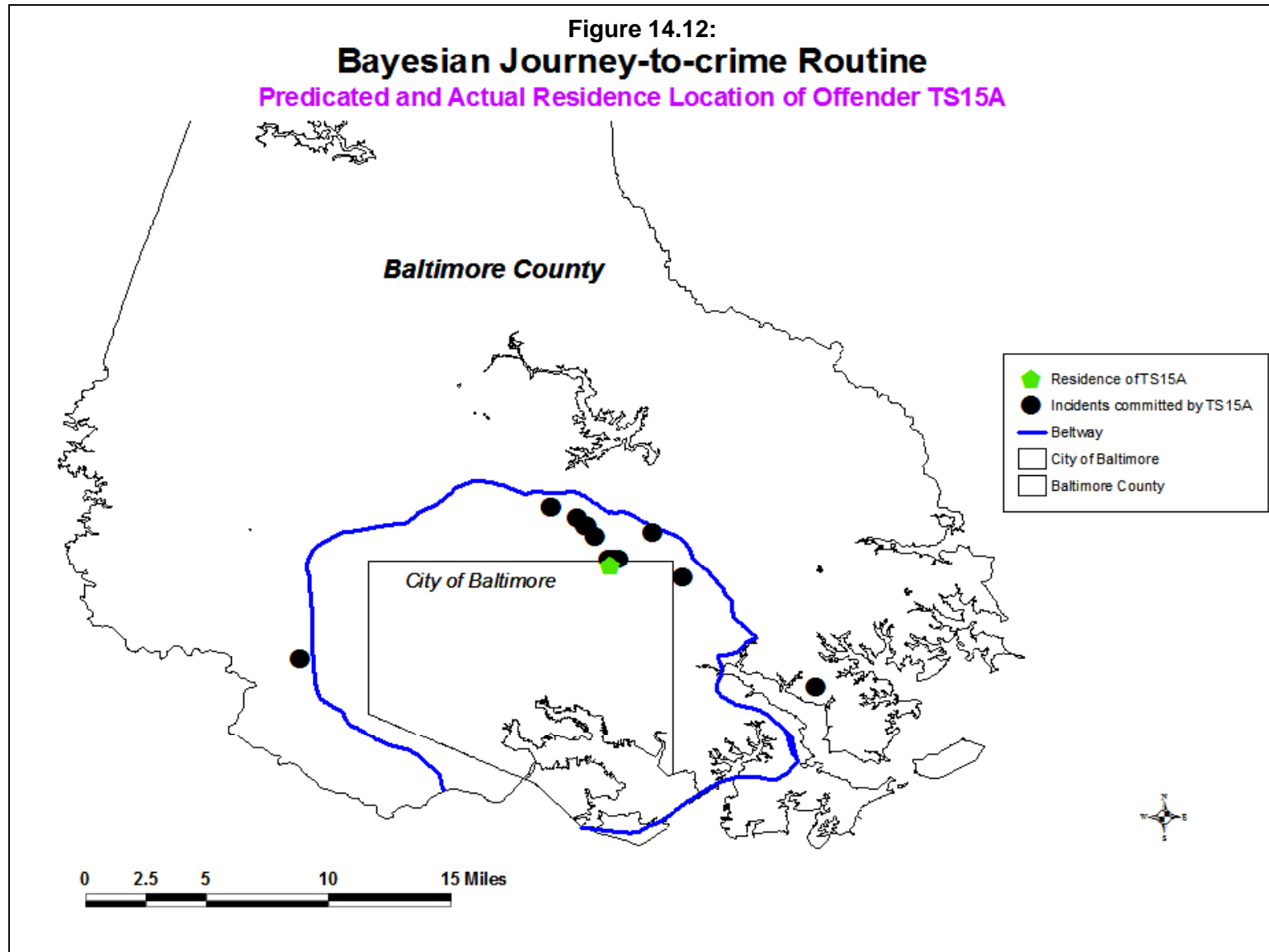## Predicated and Actual Residence Location of Offender TS15A

Baltimore County

Residence of TS15A
Incidents committed by TS15A
Beltway
City of Baltimore
Baltimore County

City of Baltimore

0   2.5   5   10   15 Miles

Figure 14.13:
Bayesian Journey-to-crime Routine
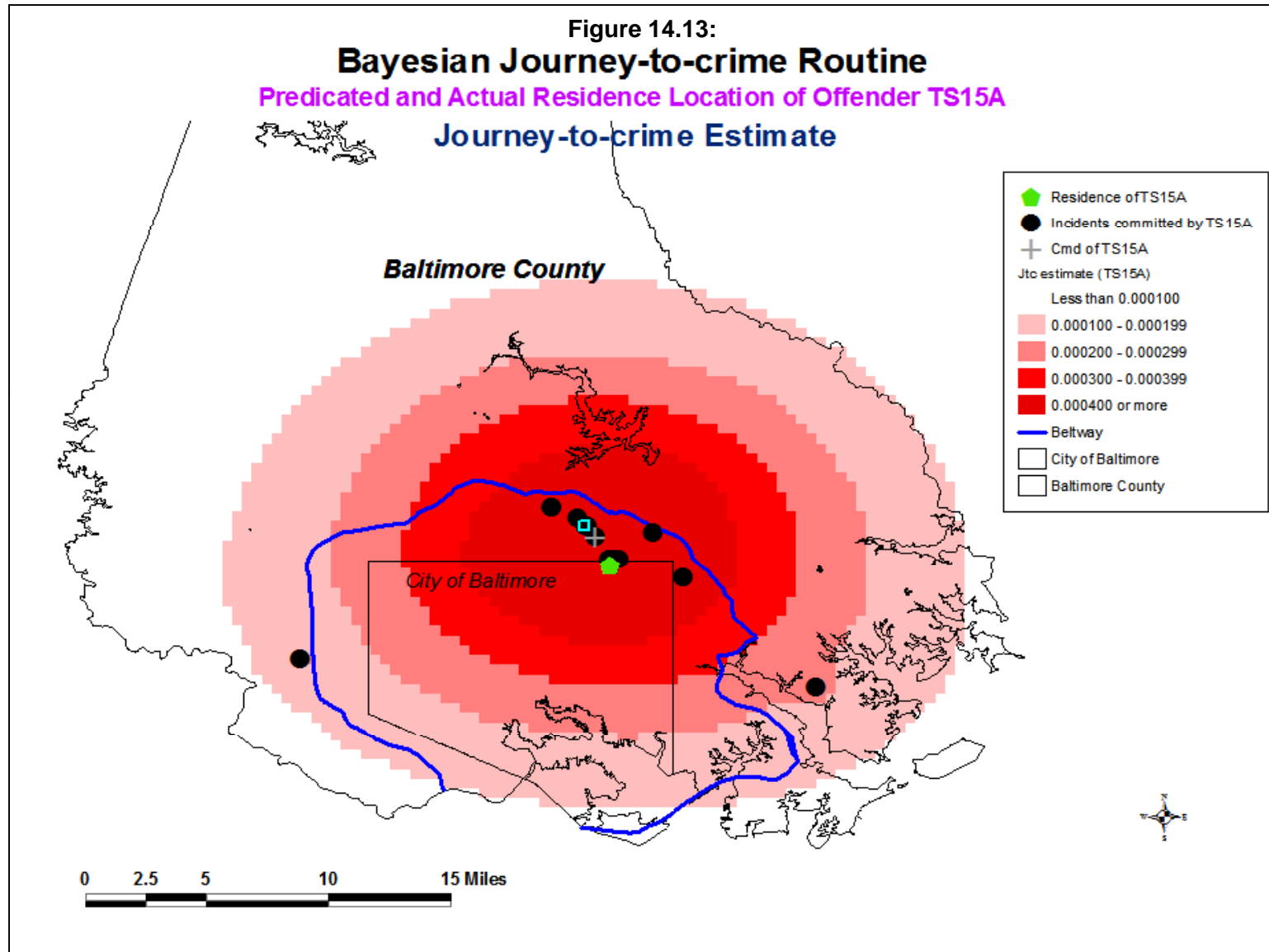**Predicated and Actual Residence Location of Offender TS15A**
Journey-to-crime Estimate

Figure 14.14:
Bayesian Journey-to-crime Routine
Predicated and Actual Residence Location of Offender TS15A
Conditional Estimate

Baltimore County

City of Baltimore

Residence of TS15A
Incidents committed by TS15A
Cmd of TS15A
Conditional estimate (TS15A)
Less than 0.000100
0.000100 – 0.000199
0.000200 – 0.000299
0.000300 – 0.000399
0.000400 or more
Beltway
City of Baltimore
Baltimore County
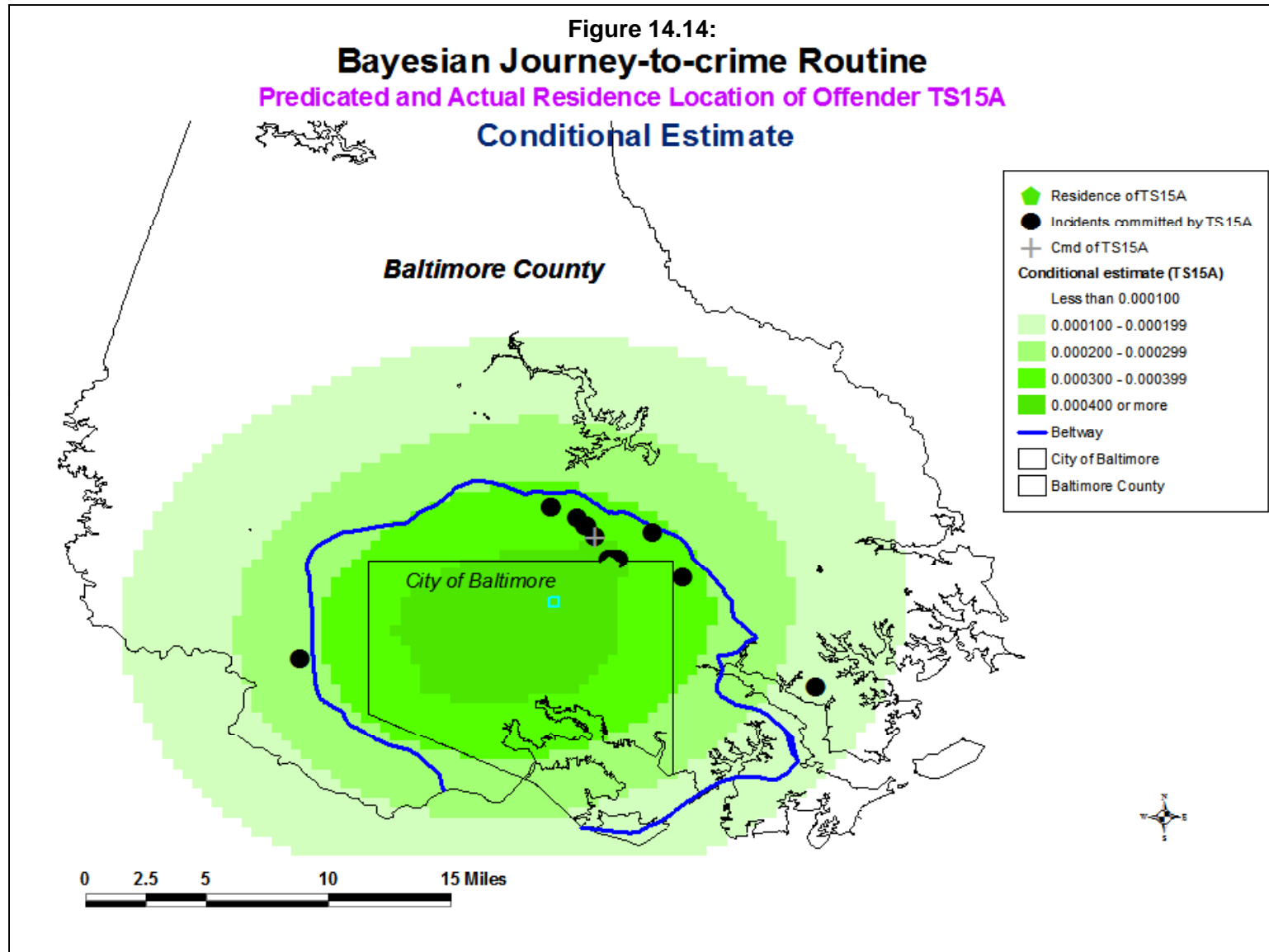
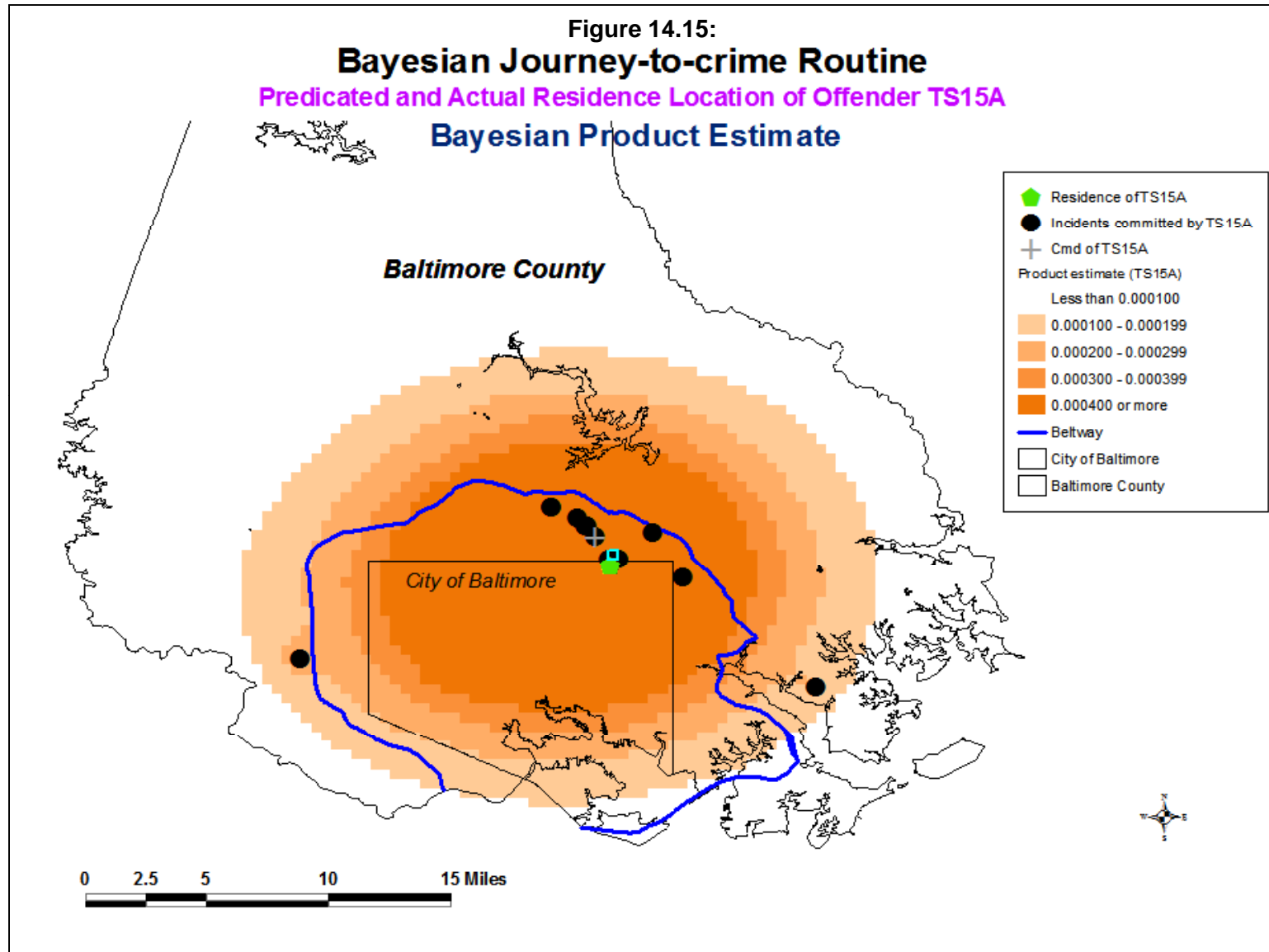0   2.5   5   10   15 Miles

Figure 14.15:
Bayesian Journey-to-crime Routine
Predicated and Actual Residence Location of Offender TS15A

Bayesian Product Estimate

In other words, the BJtc routine allows the estimation of a probability grid based on a **single** selected method. The user must decide which probability method to select and the routine then calculates that estimate and assigns it to a grid. As mentioned above, the BJtc Diagnostics routine should be first run to decide on which method is most appropriate for the jurisdiction in question. In these 88 cases, the Bayesian product estimate was the most accurate of all the probability methods. However, differences in the balance between central-city and suburbs, the road network, and land uses may change the travel patterns of offenders. So far, as mentioned above, in tests in four cities (Baltimore County, Chicago, the Hague, Manchester), the product estimate has consistently been better than the Journey-to-crime estimate and almost as good, if not better, than the center of minimum distance. Further, the product term appears to be more precise than the Journey-to-crime method though in the Hague study, the conditional estimate was more accurate than the product estimate. The center of minimum distance, while generally more accurate than other methods, has no probability distribution; it is simply a point. Consequently, one cannot select a search area from the estimate.

Potential to Add New Information to Improve the Methodology

Further, it should be possible to add more information to this framework to improve the accuracy and precision of the estimates. One obvious dimension that should be added is an opportunity matrix, a distribution of targets that are crime attractions for offenders. Among these are convenience stores, shopping malls, parking lots, and other types of land use that attract offenders. It will be necessary to create a probability matrix for quantifying these attractions. Further, the opportunity matrix would have to be conditional on the distribution of the crimes and on the distribution of origins of offenders who committed crimes in the same location. The Bayesian framework is a conditional one where factors are added to the framework but conditioned on the distribution of earlier factors:

$$P(Jtc|O) \propto P(Jtc) * P(O|Jtc) * P(A|O,Jtc) \qquad (14.13)$$

where A is the attractions (or opportunities), Jtc is the distribution of incidents, and O is the distribution of other offender origins. It will not be an easy task to estimate an opportunity matrix that is conditioned (dependent) upon both the distribution of offences (Jtc) and the origin of other offenders who committed crimes in the same location (O|Jtc) and it may be necessary to approximate this through a series of filters.

Probability Filters

A filter is a probability matrix that is applied to the estimate but is not conditioned on the existing variables in the model. For example, an opportunity matrix that was independent of

the distribution of offences by a single serial offender or the origins of other offenders who committed crimes in the same locations could be applied as an alternative (equation 14.14):

$$P(Jtc|O) \propto P(Jtc) * P(O|Jtc) * P(A) \qquad (14.14)$$

In this case, P(A) is an independent matrix. Another filter that could be applied is residential land use. The vast majority of offenders are going to live in residential areas. Thus, a residential land use filter estimates the probability of a residential land use for every cell, P(R), could be applied to screen out cells that are not residential, such as

$$P(Jtc|O) \propto P(Jtc) * P(O|Jtc) * P(A) \qquad (14.15)$$

In this way, additional information can be integrated into the Journey-to-crime methodology to improve the accuracy and precision of the estimates. Clearly, having additional variables be conditioned upon existing variables in the model would be ideal since that would fit the true Bayesian approach. But, even if independent filters were brought in, the model might be improved.

Defining Filters in the Bayesian Journey-to-crime Routine

The Bayesian Journey-to-crime routine allows filters to be applied. The routine can be run with or without filters and the user has a choice of running the routine with no filters, one filter (called 'Filter 1') and two filters (called 'Filter 1' and 'Filter 2'). See Figure 14.4 above that illustrates how to define the filters on the Bayesian Journey-to-crime page.

For example, one filter could be whether the grid cell is residential or not. Each zone in the filter variable could have a dummy variable indicating whether it is primarily residential (1) or not (0). The criteria for defining residential could be having a minimum number of residential units but also having less than a specified number of persons employed in the zone. A 'pure' residential zone would have only residences and no employment.

Kent and Leitner (2009) showed that the use of residential land covers improved accuracy for Jtc estimates, but did not improve estimates for the Bayesian approach. Nevertheless, it is likely that a subset of residential land cover might improve the precision of an estimate.

Another filter could be the amount of employment in a zone. Zones with many employees (e.g., commercial areas) would have high values on the filter variable while zones with few, if any, employees would have low values on the filter.

A third filter could be the number of businesses of a certain type for crimes of a particular type. For example, to model liquor store robberies, the filter variable could be the number of liquor stores in each zone. Or, to model bank robberies, the filter variable could be the number of banks in each zone.

Whichever variable is used for the filter, the routine interpolates this to the same grid as the Journey-to-crime function, P(O), the conditional probability function, P(O|Jtc), the general function, P(O), the production probability function, P(Jtc)*P(O|Jtc), and the Bayesian Risk function, $\frac{P(Jtc)*P(O|Jtc)}{P(O))}$.

The interpolated filter grid is then multiplied by four functions - the Journey-to-crime grid, P(Jtc), the conditional probability function, P(O|Jtc), the product probability function, P(Jtc)*P(O|Jtc), and the Bayesian Risk function, $\frac{P(Jtc)*P(O|Jtc)}{P(O))}$.

Example of the Use of a Probability Filter

To illustrate this, Figure 14.16 shows the location of 22 crimes committed by a single offender, S22A, and the offender's residence when arrested. The incidents are shown in blue and the residence location in black. The crimes committed were 6 commercial burglaries, 1 residential burglary, 11 vehicle break-ins and 4 vehicle thefts.

Figure 14.17 shows the result of the Journey-to-crime probability estimate. As with the other maps, the center of minimum distance (CMD) is shown as a gray cross. Notice that neither the center of minimum distance nor the journey-to-crime estimates were particularly accurate as the cell with the peak probability was 2.5 miles and the CMD was 2.6 miles respectively away from the actual home location of the offender.

Figure 14.18 shows the conditional probability estimate. With this estimate, the cell with the peak probability was much more accurate, being about 0.5 miles away.

Figure 14.19 shows the product probability estimate which multiplies the Journey-to-crime estimate by the conditional probability estimate and then re-scales the grid to sum to 1.0. This estimate was not particularly accurate as well with the cell having the peak probability being 2.5 miles away. The reason is that the inaccurate Journey-to-crime estimate also made the product estimate inaccurate. In some cases, a more accurate conditional estimate will improve the product probability but in other cases it will not. Block and Bernasco (2009) found that journey-to-crime estimates were inaccurate with serial burglars in The Hague, Netherlands, and that the conditional probability estimate was more accurate than the product probability estimate because the poor journey-to-crime estimates degraded the product estimates. That is why it is

Figure 14.16:
Bayesian Journey-to-crime Routine
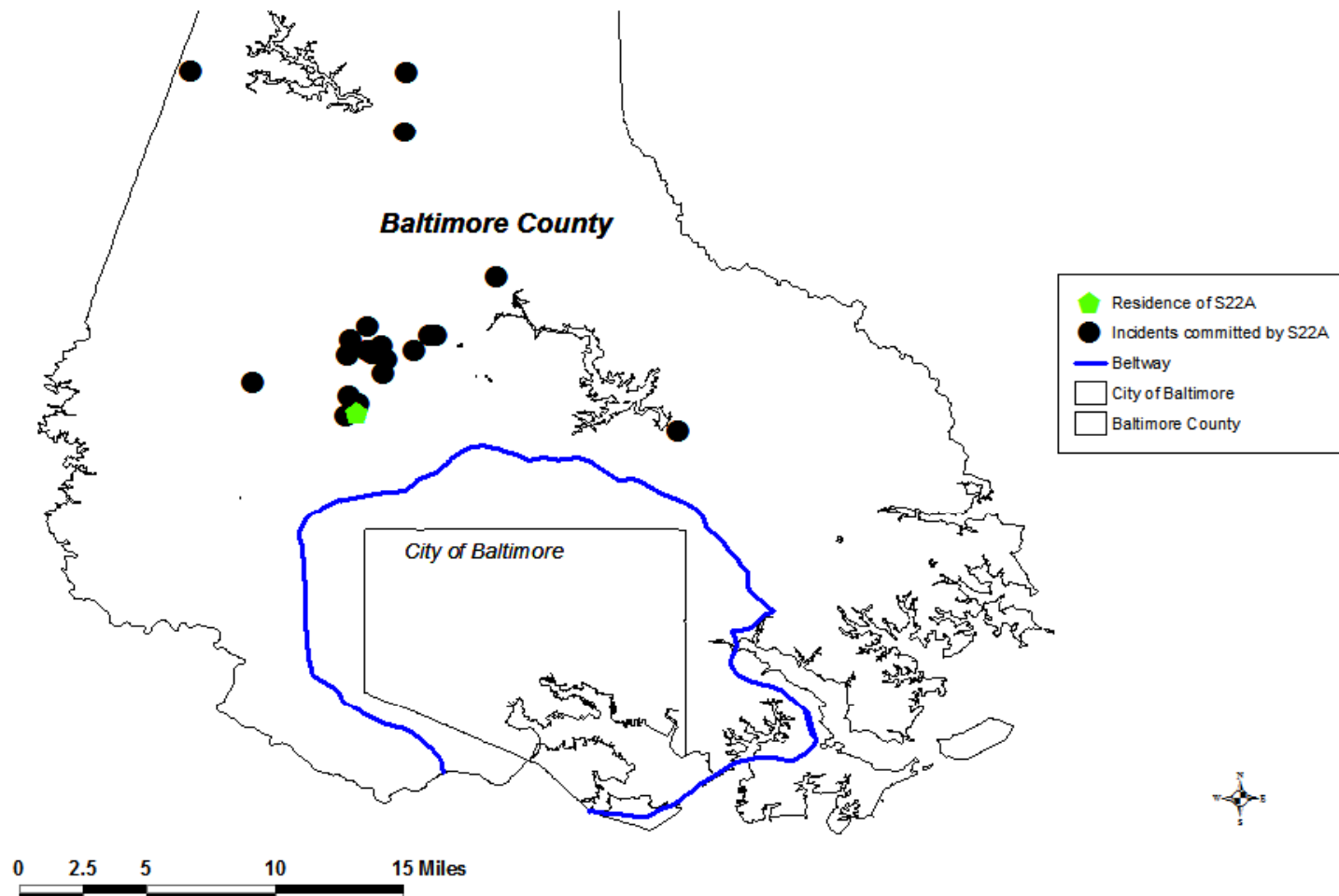Location of Incidents and Residence of Offender S22A

Baltimore County

City of Baltimore

Residence of S22A
Incidents committed by S22A
Beltway
City of Baltimore
Baltimore County

0   2.5   5   10   15 Miles

Figure 14.17:
Bayesian Journey-to-crime Routine
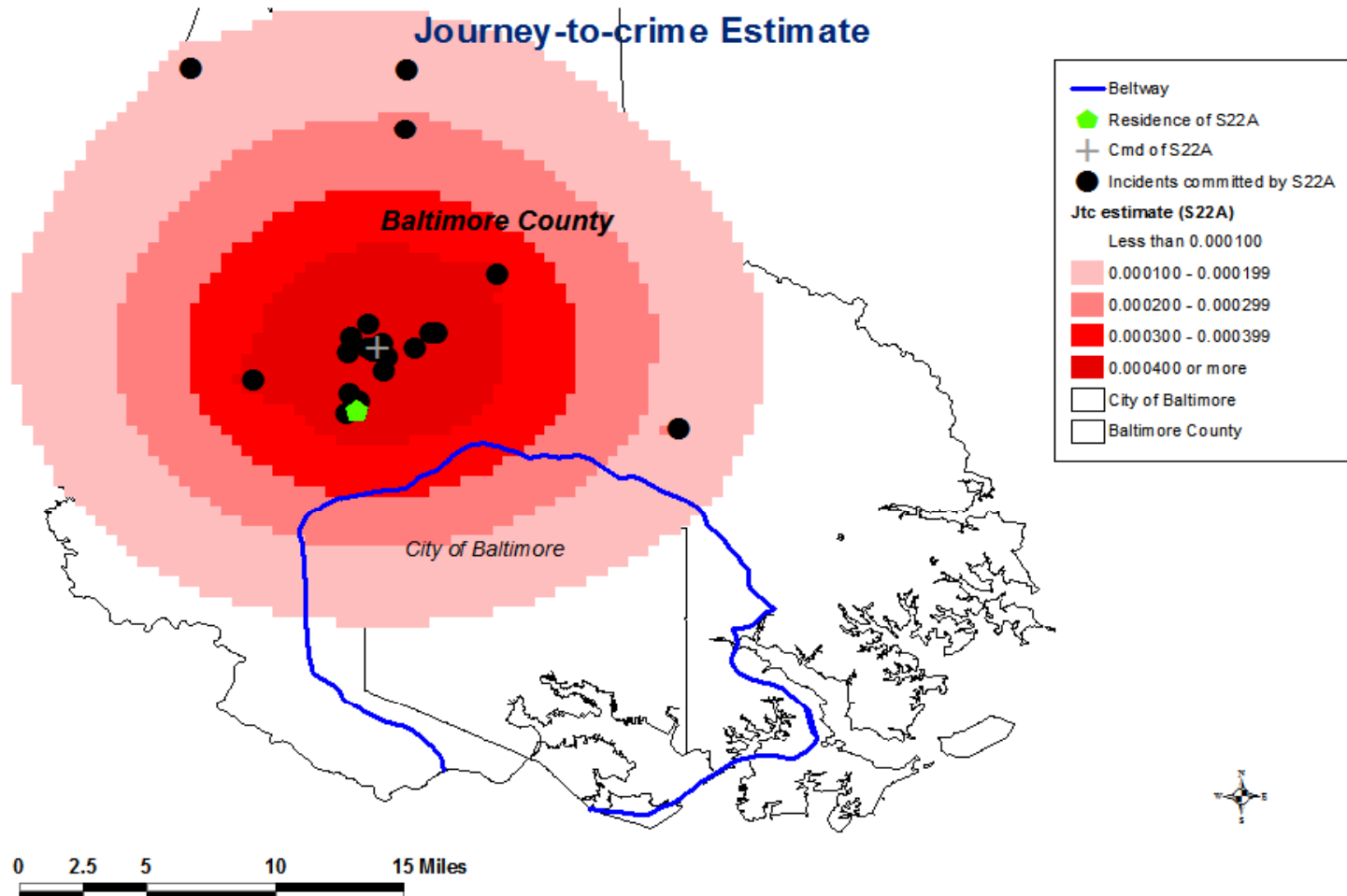Predicted and Actual Residence Location of Offender S22A

Figure 14.18:
Bayesian Journey-to-crime Routine
Predicted and Actual Residence Location of Offender S22A

Conditional Estimate

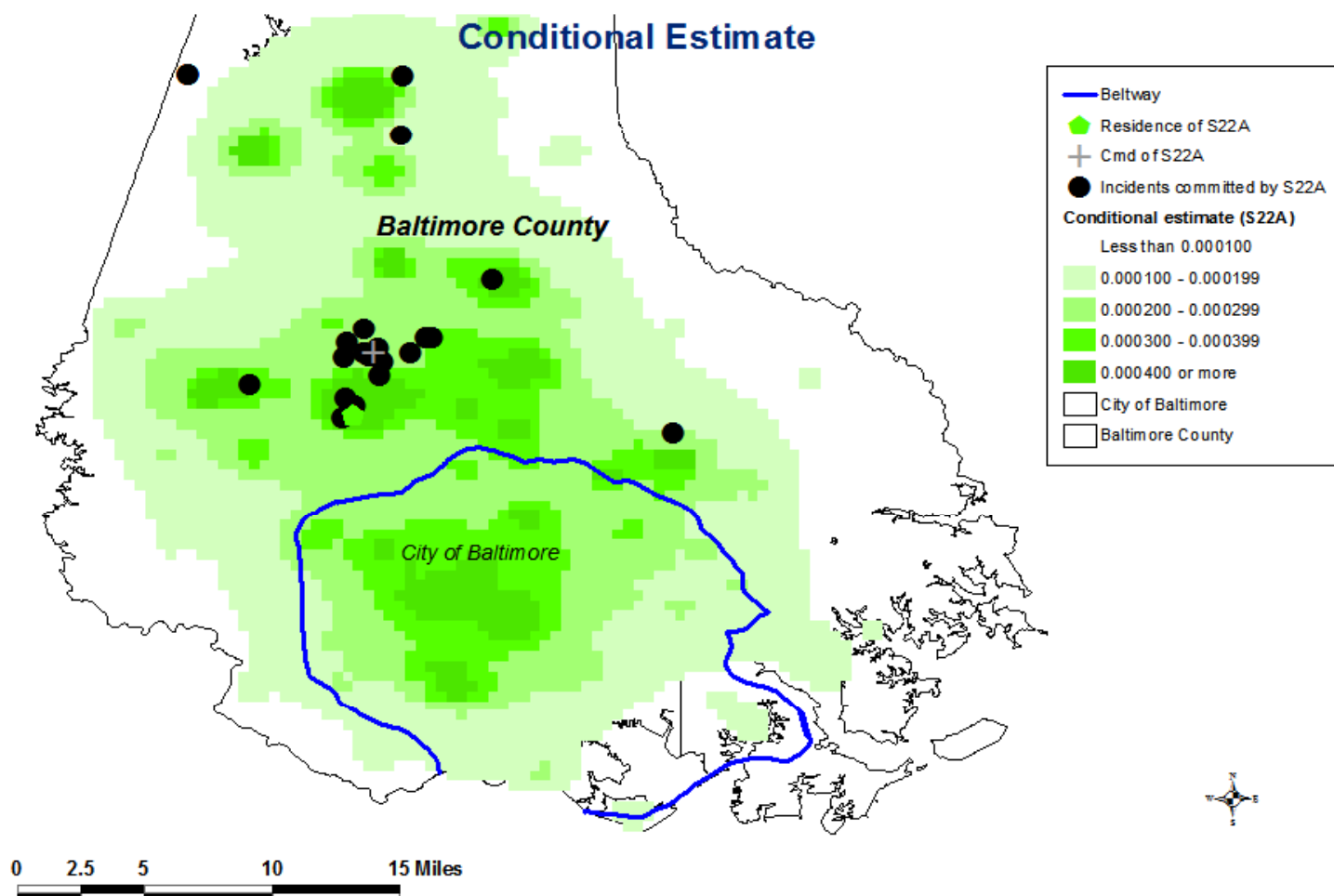Baltimore County

City of Baltimore

Beltway
Residence of S22A
Cmd of S22A
Incidents committed by S22A
Conditional estimate (S22A)
Less than 0.000100
0.000100 - 0.000199
0.000200 - 0.000299
0.000300 - 0.000399
0.000400 or more
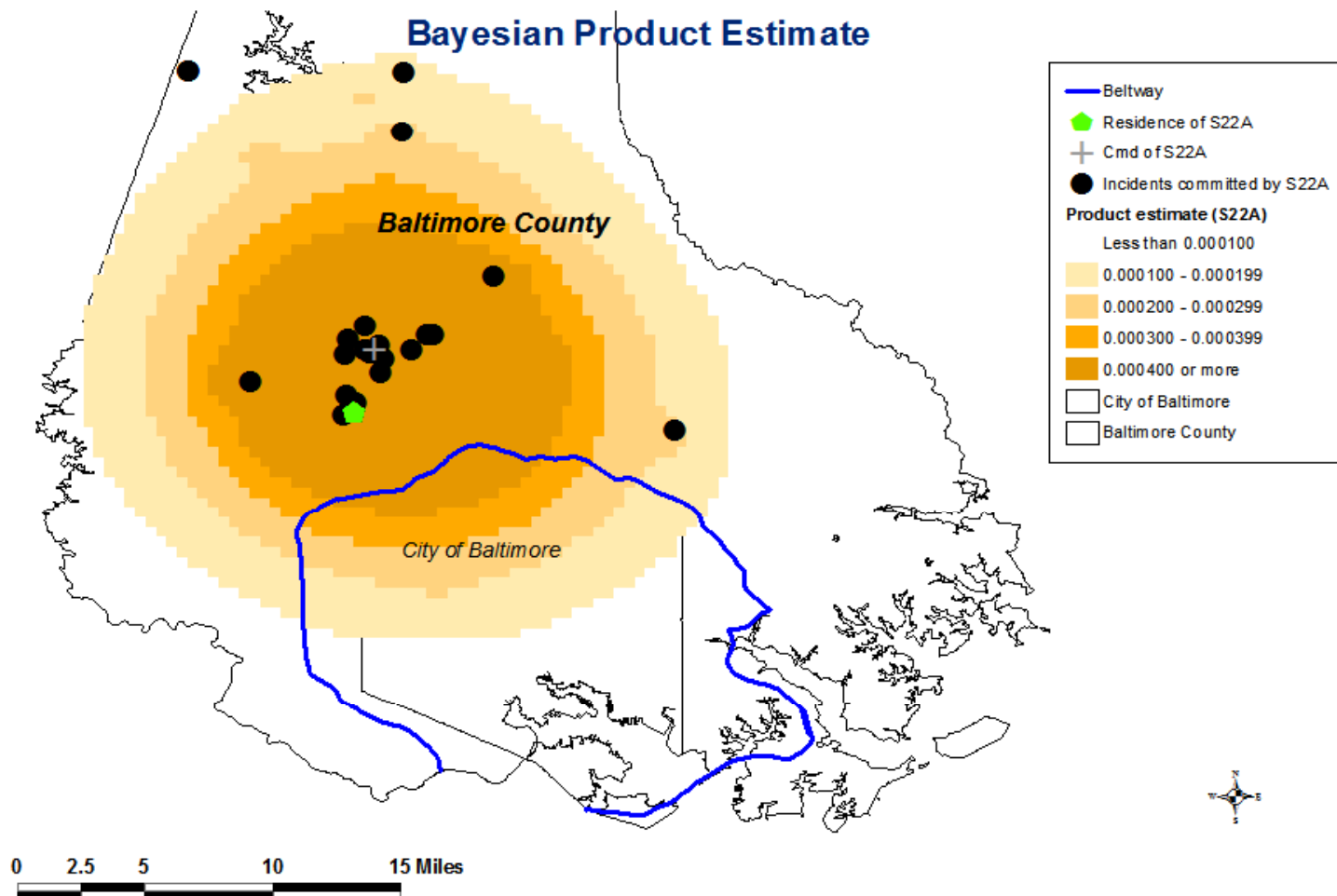City of Baltimore
Baltimore County

0   2.5   5   10   15 Miles

Figure 14.19:
Bayesian Journey-to-crime Routine
Predicted and Actual Residence Location of Offender S22A

Bayesian Product Estimate

Baltimore County

City of Baltimore

Beltway
Residence of S22A
Cmd of S22A
Incidents committed by S22A
Product estimate (S22A)
Less than 0.000100
0.000100 - 0.000199
0.000200 - 0.000299
0.000300 - 0.000399
0.000400 or more
City of Baltimore
Baltimore County

0   2.5   5   10   15 Miles

important to analyze which method is best for a single jurisdiction using the Bayesian journey-to-crime diagnostics routine before applying a particular method a single serial offender.

With offender S22A, a residential land use probability filter was defined by a zonal data base of Traffic Analysis Zones (TAZ) in Baltimore County that included both residential population variables and employment variables. A dummy variable was created by defining TAZ's that had 100 or more persons living in them but 200 or fewer employees working in them. Thus, these TAZ's were primarily residential. When the TAZ layer was interpolated to the grid in the routine, each grid cell had a probability value that varied from 0 to 1 and which indicated the likelihood of the cell residential.

Figure 14.20 shows the result of combining the product probability estimate with this residential land use filter. The results were as accurate as the conditional probability estimate in distance as the cell with the peak probability was about 0.5 miles away. But, more important is the probability estimate in the cell with the peak probability was much higher than the probability estimated for the conditional (0.008842 compared to 0.000345, a ratio that was 25.6 times higher).

In other words, the effect of narrowing the probability estimates of the product probability by discounting cells that were not residential actually improved the accuracy of the product probability estimate. We do not yet know whether using a residential filter will always improve accuracy since we have not tested it on a number of cases yet. It is possible that these filters will improve accuracy but it is also possible that they will make precision worse since they multiply a conditional probability by a matrix that is constant for all offenders.

Until a thorough evaluation is conducted, the filters are provided as tools for users to experiment with in modeling the likely residence location of a serial offender.

Guidelines for Analysts

The following discussion is for analysts wishing to utilize the technique to try to narrow down the geographic areas for particular serial offenders. The hardest part of using the technique is collecting the data and constructing a journey-to-crime (Jtc) function and an origin-destination (O-D) matrix. However, once the data have been assembled and the Jtc function and the O-D matrix constructed, the technique can be used for multiple serial offenders. These estimates need to be only updated every few years in order to account for changes in travel patterns by offenders.

Figure 14.20:
Bayesian Journey-to-crime Routine
Predicted and Actual Residence Location of Offender S22A
Bayesian Product Estimate with Residential Filter
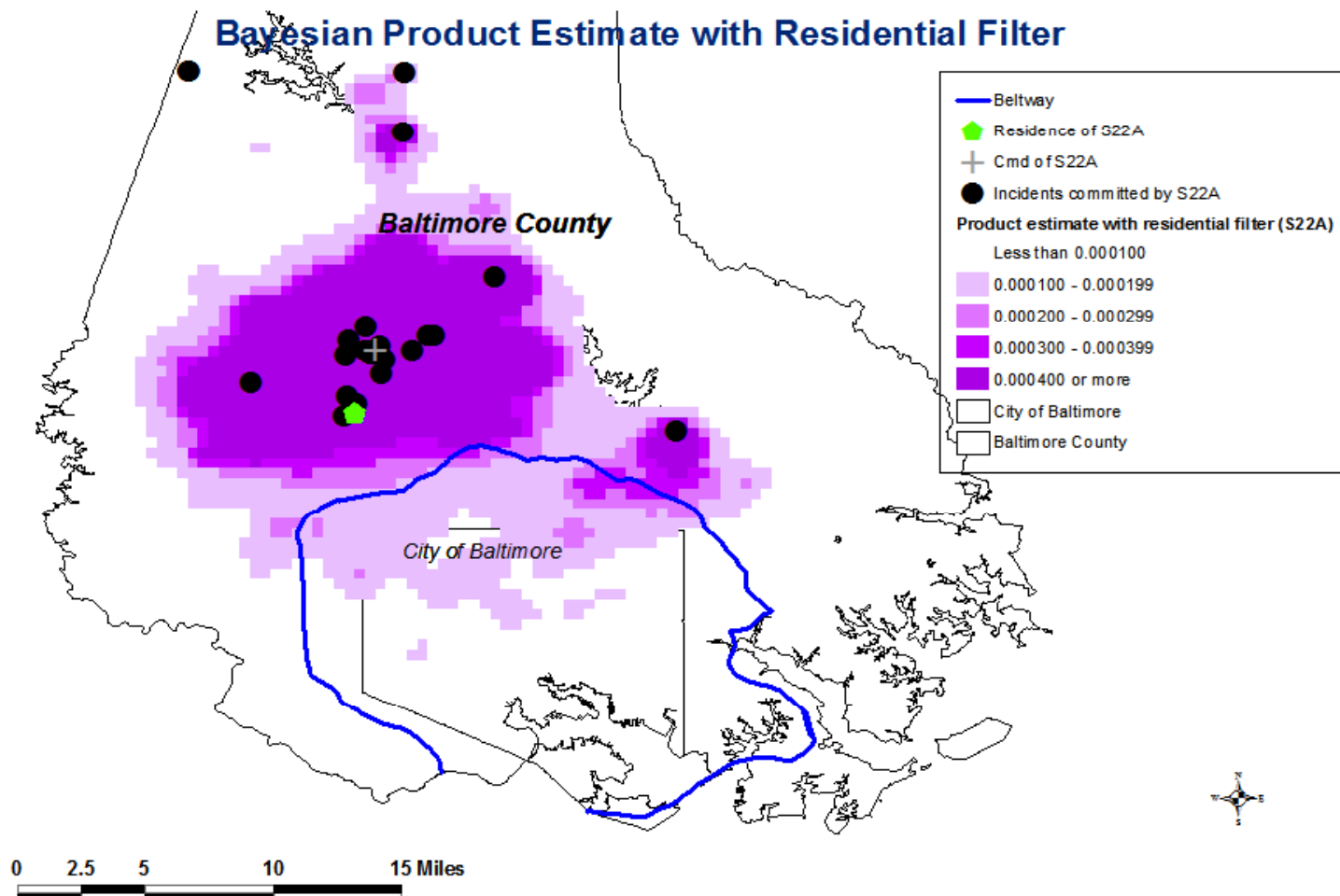
We have argued that an analyst should test which of the different methods produces the best estimate for a particular jurisdiction. However, if an analyst wants to choose a single best technique without testing which method works best in the jurisdiction, we recommend sticking with the Product probability by itself (without filters). We have found that the Product estimate (the product of the Jtc and Conditional probability estimates) generally produces more accurate results than the Jtc function by itself or the Conditional probability by itself, though some exceptions have been noted. The use of filters to improve estimates is still too new a technique and needs to be evaluated further.

To simplify, seven basic steps are required to run the "Estimate likely location of a serial offender" routine and one additional step if the analyst wants to test which method works best for the jurisdiction.

Steps

1. Obtain Required Data

First, the data that will be needed is a large number of records where both the residence location and the crime location are known. Most likely, these will come from arrest records. By large, we mean at least 10,000 cases.

2. Construct Journey-to-crime Function

Second, once these data have been assembled, the user should create a journey-to-crime function using the "Calibrate journey-to-crime function" routine (discussed in Chapter 13).

3. Define Zonal Framework

Third, to construct an origin-destination matrix, the analyst will need a zonal framework for allocating the incidents to both origin and destination locations. Commonly used zones are census tracts or traffic analysis zones, though others can be used. Also, we have found good results by using a grid as the zone structure, especially with small-sized grid cells (e.g., a 100 column x 100 row grid). The single kernel density interpolation tool (discussed in Chapter 10) is a useful tool for creating a grid overlay that can be used as a zone framework.

4. Construct Origin-Destination Matrix

Fourth, using the data on offenders where both the residence location and the crime location are known and the zonal framework, the origin-destination matrix can be constructed using the "Calculate observed origin-destination trips" routine that is discussed in Chapter 28.

The routine reads in an origin file (the zonal framework) and a destination file (also the zonal framework) and a data file (the set of records of offenders where both residence and crime location are known) and outputs the O-D matrix.   The user should save the matrix as a dbf file.

5. Input Jtc Function and O-D Matrix

Fifth, the Jtc function and the O-D matrix are input on the Bayesian Journey-to-crime page.

6. Input Records of Single Serial Offender

Sixth, to estimate the likely origin (residence) location of a serial offender, the records for that serial offender need to be input as the Primary File. As with all Primary File inputs, the coordinate system and the data metrics need to be defined.

7. Estimate Likely Origin Location of a Serial Offender

Seventh, and finally, the user selects <u>one</u> estimation method on the Bayesian Journey-to-crime page and runs the routine. As mentioned above, unless there is contrary information, we recommend using the Product estimate by itself ("Use product of P(Jtc) and P(O|Jtc) estimate").

8. (Optional) Evaluate which Method Produces the Best Results for the Jurisdiction

The Product estimate will generally produce good results for medium-to-large cities. However, for small cities, it may not work well and other measures may work better (e.g., the Conditional or the Bayesian Risk estimate). Therefore, an optional strategy is to evaluate which of the methods works best in the jurisdiction.

To do this, the analyst will have to assemble a diagnostics file on multiple serial offenders where the offender has committed multiple offences (e.g., 5 or more) and where both the residence and crime locations are known. The number of offenders included should be as large as possible, at least 50. The four studies mentioned on page 14.30 all used sizeable data sets (60 or more). The reason is that there needs to be sufficient variability to allow the routine to properly estimate the accuracy and precision of each of the methods.

While this requires even more data to be collected, the advantage is that the best estimation method can be determined for the jurisdiction. As mentioned, the use of the Product estimate may or may not produce the best estimates.

Summary

In summary, the Bayesian Jtc methodology is an improvement over the current Journey-to-crime method and appears to be as good, and more useful, than the center of minimum distance. First, it adds new information to the Journey-to-crime function to yield a more accurate and precise estimate. Second, it can sometimes predict the origin of 'commuter'-type serial offenders, those individuals who do not commit crimes in their neighborhoods (Paulsen, 2007; Canter & Larkin, 1994). The traditional Journey-to-crime function cannot predict the origin location of a 'commuter'-type. Of course, this will only work if there are prior offenders who lived in the same location as the serial offender of interest. If the offender lived in a neighborhood where there were no previous serial offenders that were documented in the origin-destination matrix, the Bayesian approach would not detect that location, either.

Caveat

A caveat should be noted, however. The Bayesian method still has a substantial amount of error. Much of this error reflects the inherent mobility of offenders, especially those living in in suburbs outside of central cities. While adolescent offenders, especially juvenile males, tend to commit crimes within a more circumscribed area (Levine & Lee, 2013), the almost-universal ability adults to own automobiles and to travel outside their residential neighborhoods is turning crime into a much more mobile phenomena than it was, say, 50 years ago when only about half of American households owned an automobile.

Thus, the Bayesian approach to Journey-to-crime estimation must be seen as a tool that produces an incremental improvement in accuracy and precision. Geographic profiling is but one tool in the arsenal of methods that police must use to catch serial offenders.

References

Block, R. & Bernasco, W. (2009). Finding a serial burglar's home using distance decay and conditional origin-destination patterns: A test of Empirical Bayes Journey-to-crime estimation in The Hague. *Journal of Investigative Psychology & Offender Profiling*. 6(3), 187-211.

Canter, D. (2009). Developments in geographical offender profiling: Commentary on Bayesian journey-to-crime modeling. *Journal of Investigative Psychology & Offender Profiling*. 6(3), 161-166.

Canter, D. (2003). *Dragnet: A Geographical Prioritisation Package*. Center for Investigative Psychology, Department of Psychology, The University of Liverpool: Liverpool, UK. http://www.i-psy.com/publications/publications_dragnet.php.

Canter, D., Coffey, T., Huntley, M., & Missen, C. (2000). Predicting serial killers' home base using a decision support system. *Journal of Quantitative Criminology, 16* (4), 457 -- 478.

Canter, D. & Gregory, A. (1994). Identifying the residential location of rapists, *Journal of the Forensic Science Society*, 34 (3), 169-175.

Canter, D. & Larkin, P. (1993). The environmental range of serial rapists, *Journal of Environmental Psychology*, 13, 63-69.

Denison, D.G.T., Holmes, C.C. Mallilck, B. K. & Smith, A. F. M. (2002). *Bayesian Methods for Nonlinear Classification and Regression*. John Wiley & Sons, Ltd: New York.

Gelman, A., Carlin, J. B. Stern, H. S. & Rubin, D. B. (2004). *Bayesian Data Analysis* (second edition). Chapman & Hall/CRC: Boca Raton, FL.

Jessen, R. J. (1979). *Statistical Survey Techniques*. John Wiley & Sons: New York.

Kanji, G. K. (1993). *100 Statistical Tests*. Sage Publications: Thousand Oaks, CA.

Kent, J. & Leitner, M. (2009). Utilizing land cover characteristics to enhance journey-to-crime estimation models. *Crime Mapping*, 1.

Lee, P. M. (2004). *Bayesian Statistics: An Introduction* (third edition). Hodder Arnold: London.

References (continued)

Leitner, M. & Kent, J. (2009). Bayesian Journey-to-crime modeling of single- and multiple crime type series in Baltimore County, MD. *Journal of Investigative Psychology & Offender Profiling*. 6(3), 213-236.

Levine, N. (2009). Introduction to the special issue on Bayesian Journey-to-crime modeling. *Journal of Investigative Psychology & Offender Profiling*. 6(3), 167-185.

Levine, N. (2005). *"The evaluation of geographic profiling software: Response to Kim Rossmo's critique of the NIJ methodology"*. http://www.nedlevine.com/Response to Kim Rossmo Critique of the GP Evaluation Methodology.May 8 2005.doc

Levine, N. (2000). Journey-to-crime Estimation. Chapter 10 of Levine, N. (ed), *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations* (version 1.1). Ned Levine & Associates, Annandale, VA; National Institute of Justice, Washington, DC. August. See archived files at http://www.icpsr.umich.edu/CrimeStat.

Levine, N. & Lee, P. (2013). Crime travel of offenders by gender and age in Manchester, England. Leitner, M. (ed), *Crime Modeling and Mapping Using Geospatial Technologies,* Springer. 145-178.

Levine, N., & Block, R. (2010). Bayesian Journey-to-Crime Estimation: An Improvement in Geographic Profiling Methodology. *The Professional Geographer*. 63(2), 213-229.

Levine, N. & Lee, P. (2009). Bayesian Journey-to-crime modeling of juvenile and adult offenders by gender in Manchester. *Journal of Investigative Psychology & Offender Profiling*. 6(3), 237-251.

O'Leary, M. (2009). The mathematics of geographical profiling. *Journal of Investigative Psychology & Offender Profiling*. 6(3), 253-265.

Paulsen, D. (2007). Improving geographic profiling through commuter/marauder prediction.. *Police Practice and Research*, 8: 347-357

Paulsen, D. (2006a). Connecting the dots: assessing the accuracy of geographic profiling software. *Policing: An International Journal of Police Strategies and Management*. 20 (2), 306-334.

References (continued)

Paulsen, D. (2006b). Human versus machine: A comparison of the accuracy of geographic profiling methods. *Journal of Investigative Psychology and Offender Profiling* 3: 77-89.

Rich, T., & Shively, M. (2004). *A Methodology for Evaluating Geographic Profiling Software*. Final Report for the National Institute of Justice, Abt Associates: Cambridge, MA. http://www.ojp.usdoj.gov/nij/maps/gp.pdf.

Rossmo, D. K. (2005a). Geographic heuristics or shortcuts to failure?: Response to Snook et al. *Applied Cognitive Psychology* 19: 651-654.

Rossmo, D. K. (2005b). Response to NIJ's methodology for evaluating geographic profiling software. http://www.ojp.usdoj.gov/nij/maps/gp.htm.

Rossmo, D. K., & Filer, S. (2005). Analysis versus guesswork. *Blue Line Magazine*, , August / September, 24:26.

Rossmo, D. K. (2000). *Geographic Profiling*. CRC Press: Boca Raton Fl.

Rossmo, D. K. (1995). Overview: multivariate spatial profiles as a tool in crime investigation. In Block, C. R., Dabdoub, M. & Fregly, S., *Crime Analysis Through Computer Mapping*. Police Executive Research Forum: Washington, DC. 65-97.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill: New York.

Snook, B., Zito, M., Bennell, C. & Taylor, P. J. (2005). On the complexity and accuracy of geographic profiling strategies. *Journal of Quantitative Criminology*, 21 (1), 1-26.

Snook, B., Taylor, P. J. & Bennell, C. (2004). Geographic profiling; the fast, fugal and accurate way. *Applied Cognitive Psychology* 18: 105-121.