GEOGRAPHIC PROFILING FROM KINETIC MODELS OF CRIMINAL BEHAVIOR[*]

GEORGE O. MOHLER[†] AND MARTIN B. SHORT[‡]

Abstract. We consider the problem of estimating the probability density of the "anchor point" (residence, place of work, etc.) of a criminal offender given a set of observed spatial locations of crimes committed by the offender. Starting from kinetic models of criminal motion and target selection, we derive the probability density of anchor points using the Fokker–Planck equation and Bayes' theorem. Here, geographic inhomogeneities such as housing densities and geographic barriers (bodies of water, parks, etc.) are naturally incorporated into the probability density estimate, as well as directional bias and distance to crime preferences in offender target selection. The resulting equations are steady state advection-diffusion-reaction PDEs. We test our methodology against crime data provided by the Los Angeles Police Department, and our results highlight the benefits of incorporating these elements of criminal behavior and geographic inhomogeneities into profiling estimates.

Key words. geographic profiling, stochastic differential equations, Bayesian methods, crime

AMS subject classifications. 62M30, 31A25

DOI. 10.1137/100794080

1. Introduction. A classical problem arising in crime science is the estimation of the anchor point $\mathbf{z} \in \mathbb{R}^2$ of a criminal offender, given a set of observed spatial locations $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbb{R}^2$ of crimes assumed to have been committed by the individual. The anchor point could be the offender's legal residence, a friend's or family member's residence, place of work, etc. [16], and can be thought of as a staging point for criminal activity.

In practice, a score function

$$(1.1) \qquad S(\mathbf{z}) = \sum_{i=1}^{N} f(d(\mathbf{x}_i, \mathbf{z}))$$

is often used to determine likely locations for the anchor point [18, 4, 13, 16]. Here, d is a distance metric and f is a kernel that typically decays at long distances (but may be increasing at short distances). One criticism of such an approach is that the score function in (1.1) is not a probability density; O'Leary argues in [16] for modeling the conditional probability density $P(\mathbf{z}|\mathbf{x}_1, \ldots, \mathbf{x}_N)$, instead of the score function.

Another major fault common among methods of the form (1.1) is that f is typically isotropic and does not take into account geographic features such as housing density or observed patterns of criminal behavior. For example, consider the residential burglary scenario depicted in Figure 1.1. Criminals with anchor points located at $\mathbf{z}_1$, $\mathbf{z}_2$, and $\mathbf{z}_3$ would all be assigned equal likelihood of committing the burglary at the red house in the center under a model of the form (1.1), as these points are all the

[†]Department of Mathematics and Computer Science, Santa Clara University, San Jose, CA 95126 (gmohler@scu.edu).

[‡]Department of Mathematics, University of California, Los Angeles, CA 90095 (mbshort@math.ucla.edu).
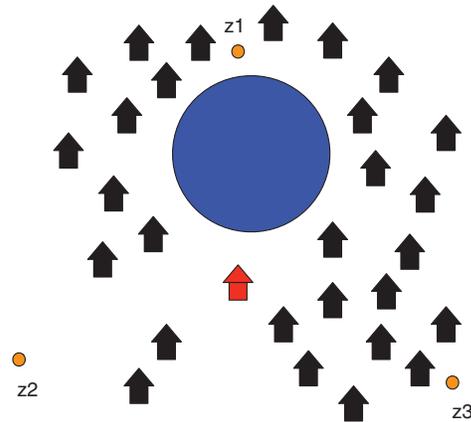
FIG. 1.1. *Residential burglary scenario with geographic heterogeneities.*

same distance from the crime scene. However, the likelihood of the anchor point being at $\mathbf{z}_1$ should be low (in comparison to other points equidistant to the crime scene), because the offender would have to travel around the lake to get there, encountering many attractive targets along the way. Likewise, the likelihood corresponding to $\mathbf{z}_3$ should be lower than $\mathbf{z}_2$, because there are more potential targets in between that could prevent the criminal starting at $\mathbf{z}_3$ from reaching the red house without committing a burglary. Furthermore, a model of the form (1.1) would assign even higher likelihood to points on the lake, as every point on the lake is closer to the red house than $\mathbf{z}_1$, $\mathbf{z}_2$, or $\mathbf{z}_3$.

In this paper, we take an alternative approach to that of (1.1), deriving the probability density of the anchor point, $P(\mathbf{z}|\mathbf{x}_1, \ldots, \mathbf{x}_N)$, using Bayes' theorem as suggested in [16]. This requires a model for the spatial distribution of crimes given an anchor point, as well as a model for the (prior) distribution of anchor points. We derive the former starting from a kinetic description of criminal behavior, along the lines of models introduced in [2, 22, 9, 26] for criminal movement and target selection. Our model takes into account geographic features not typically accounted for in standard geographic profiling models. We then indicate how such a model can be implemented in practice, including computation of the geographic profile, parameter estimation, and modeling of the prior distribution.

The outline of the paper is as follows. In section 2, we introduce a kinetic model of criminal behavior that can be used directly to compute geographic profiles or used as a starting point for more computationally efficient methods. In section 3, we derive equations for the probability density, $P(\mathbf{z}|\mathbf{x}_1, \ldots, \mathbf{x}_N)$, from the kinetic model using the Fokker–Planck equation and Bayes' theorem. Evaluation of the geographic profile density requires the solution of steady state advection-diffusion-reaction equations, and we detail some such solutions in specific cases. In section 4, we fit our model to burglary series data provided by the Los Angeles Police Department. The parameter estimates provide insight into how far criminals travel to commit offenses, the variance of these distances within each series, and directional bias in target selection. In section 5, we illustrate the effectiveness of our methodology in comparison with another common geographic profiling method using the data provided by LAPD.

2. A kinetic model of criminal behavior. Here we develop a kinetic model of criminal behavior that will be used to derive geographic profiling estimates. The

model assumes a foraging behavior for the criminal [2, 22, 11], though in general the type of model should depend on the crime type. As proposed in [22], we start by introducing a (stationary) spatial attractiveness field $A(\mathbf{y}|\mathbf{z}) \geq 0$, reflecting how attractive the target positioned at $\mathbf{y}$ is to a criminal also positioned there, which may depend upon the anchor point $\mathbf{z}$ of the criminal. In the case of residential burglary, A may be proportional to housing density $H(\mathbf{y})$ under the assumption that all houses are equally attractive. The attractiveness field will determine the rate at which criminals commit their crimes.

Letting $\mathbf{y}(t)$ denote the position of a criminal at time t, we model the movement of the criminal by the stochastic differential equation

$$(2.1) \qquad \frac{d\mathbf{y}}{dt} = \boldsymbol{\mu}(\mathbf{y}) + \sqrt{2D}\mathbf{R}_t,$$

where $\mathbf{R}_t$ is a white noise, i.e., $\langle \mathbf{R}_t \rangle = \mathbf{0}$ and $\langle R_t^i R_{t'}^j \rangle = \delta_{ij}\delta(t - t')$, and D is the diffusion parameter. The drift term $\boldsymbol{\mu}$ can be neglected in the case of unbiased motion or could be used to describe more complex criminal behaviors. For instance, it has been suggested that criminals may modify their movements toward regions of higher attractiveness when selecting their targets [22]. This type of behavior could be incorporated into (2.1) through a gradient term of the form $\boldsymbol{\mu} = \chi \nabla A$ or a nonlocal potential involving the attractiveness field.

Since the anchor point can be viewed as the location from which criminals begin their search for a target, we take $\mathbf{y}(0) = \mathbf{z}$ as the initial condition for (2.1). A crime is then committed at $\mathbf{y}(t)$ according to the killing measure $A(\mathbf{y}(t)|\mathbf{z})$, the probability per unit time that the Brownian trajectory starting at $\mathbf{z}$ and given by (2.1) is terminated at the space-time point $\mathbf{y}(t)$ [10, 21].

At this point, geographic profiles could in principle be computed using Monte Carlo simulation for given values of D, $\boldsymbol{\mu}(\mathbf{y})$, and $A(\mathbf{y}|\mathbf{z})$, assuming an estimate exists for the density of criminal anchor points $P(\mathbf{z})$. Given the popularity of agent-based models for crime applications [22, 9, 26], such an approach may be appropriate for complex agent models that do not lend themselves to mathematical analysis. However, for models of the form (2.1), we show in the next section that the geographic profiles can be computed more efficiently.

3. Derivation of geographic profiling densities. Given that a criminal starts his random walk governed by (2.1) from the anchor point $\mathbf{z}$, the transition (survival) probability density $\rho(\mathbf{x}, t|\mathbf{z})$ [10] of the position of the criminal satisfies the Fokker–Planck equation

$$(3.1) \qquad \frac{d\rho}{dt} = \nabla \cdot (D\nabla\rho) - \nabla \cdot (\boldsymbol{\mu}(\mathbf{x})\rho) - A(\mathbf{x}|\mathbf{z})\rho,$$

$$(3.2) \qquad \rho_0 = \delta(\mathbf{x} - \mathbf{z}),$$

where ∇ is with respect to the variable $\mathbf{x}$ [10, 21].

Integrating (3.1)–(3.2) in time, the probability density of where the crime is committed is then determined by

$$(3.3) \qquad P(\mathbf{x}|\mathbf{z}) = A(\mathbf{x}|\mathbf{z})\rho(\mathbf{x}|\mathbf{z}),$$

where $\rho(\mathbf{x}|\mathbf{z}) = \int_0^\infty \rho(\mathbf{x}, t|\mathbf{z})dt$ solves the elliptic partial differential equation

$$(3.4) \qquad -\nabla \cdot (D\nabla\rho) + \nabla \cdot (\boldsymbol{\mu}(\mathbf{x})\rho) + A(\mathbf{x}|\mathbf{z})\rho = \delta(\mathbf{x} - \mathbf{z}).$$

We will refer to this equation as the "forward" equation.

3.1. Case 1. At this point, it is illustrative to examine some particular solutions to (3.4) for various values of $\boldsymbol{\mu}(\mathbf{x})$ and $A(\mathbf{x}|\mathbf{z})$. The simplest case to consider is that of homogeneous D and A and zero drift with $\mathbf{z} = 0$; we refer to this as Case 1. Here, $P(\mathbf{x}|\mathbf{z})$ is a radially symmetric function $P(\mathbf{r})$ and is given by

$$(3.5) \qquad\qquad P(\mathbf{r}) = \frac{\mathrm{K}_0(r/\lambda)}{2\pi\lambda^2},$$

where $\lambda \equiv \sqrt{D/A}$ is a lengthscale that serves as the only effective parameter and $\mathrm{K}_0(x)$ is the zeroth order modified Bessel function of the second kind. It is important to note that $\mathrm{K}_0(x)$ diverges logarithmically at the origin, making this particular example somewhat ill-defined. However, since the divergence is only logarithmic, the density remains integrable over all of two-dimensional space, making this divergence less worrying. Furthermore, the *distance to crime* (DTC, sometimes referred to as *journey to crime*) probability distribution $P(r) = 2\pi r P(\mathbf{r})$ is well-defined, with the mean DTC given by $\overline{r} = \pi\lambda/2$.

Case 1 may work reasonably well for criminals that tend to stay very close to their home and commit crimes isotropically around their anchor point; such criminals have typically been referred to as "marauders" in previous work [5]. However, it is also noted in [5] that some criminals appear to be "commuters," traveling relatively large distances to commit their crimes but often remaining within localized regions of the city. Examples of a marauder and a commuter for real burglary series in Los Angeles are depicted in section 5. Note that for commuters, the variance in DTC within a series is typically much smaller than the mean DTC for that series. This is something that Case 1 cannot incorporate in any way, as Case 1 possesses only one parameter that sets both the mean and variance in DTC; here, the variance is $v_r = 4\lambda^2 - \overline{r}^2$. Furthermore, commuters (and occasionally marauders) also display a strong anisotropy in crime locations, typically committing crimes along what appears to be a preferred direction, another behavior not seen in Case 1.

In [5], marauders and commuters are mathematically defined (respectively) based upon whether the criminal residence is located within the smallest circle containing the two most widely separated crimes in the series. With geographic heterogeneity, however, it is possible that under this definition a marauding data series could be mistaken for commuting if the attractiveness field A biases the marauder in a particular direction (for example, the criminal at $\mathbf{z}_2$ in Figure 1.1). Additionally, this sort of definition is binary and does not allow for any real measure of how strongly an individual may be associated with either behavior.

3.2. Case 2. To see how our kinetic model may begin to address commuters in a more general way, we first examine a simple source of anisotropy by considering the case in which D and A are homogeneous (as in Case 1) but where there is a constant drift $\boldsymbol{\mu} = \mu\hat{u}$, where $\hat{u}$ is a unit vector in any direction desired; this will be referred to as Case 2. This drift may represent the general direction that a particular criminal travels often, perhaps on his way to work or other important point along his daily routine. We assume here that no geographic barriers exist, so that a constant $\boldsymbol{\mu}$ is possible, but we note that the introduction of barriers to movement in principle can be handled on the numeric level. With $\mathbf{z} = 0$, the crime density in Case 2 is given by

$$(3.6) \qquad\qquad P(\mathbf{r}) = \frac{e^{\alpha r \cos\phi/\lambda}}{2\pi\lambda^2} \mathrm{K}_0\left(\frac{r\sqrt{1+\alpha^2}}{\lambda}\right),$$
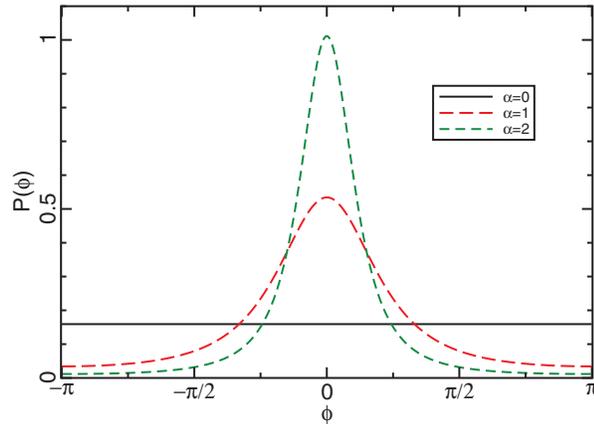
FIG. 3.1. *The distribution of criminal offenses as a function of angle ϕ as measured from the $\hat{u}$ direction for varying values of α in Case 2. A higher value of advection speed μ causes greater and greater anisotropy.*
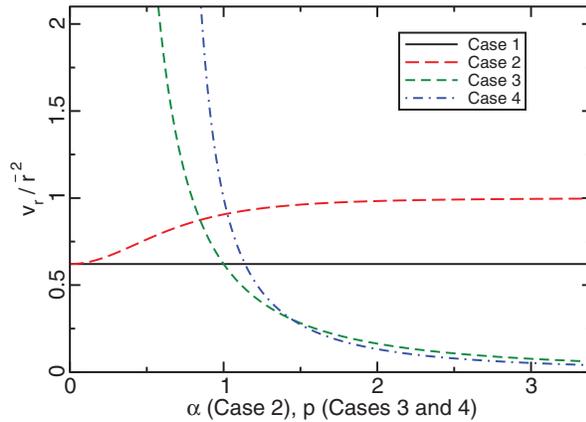

FIG. 3.2. *Comparing the ratio $v_r/\overline{r}^2$ for Cases 1–4. For Case 1, this ratio takes on the constant value $16/\pi^2 - 1$. For Case 2, the ratio is plotted versus the anisotropy parameter α; we see that all values of the ratio are higher than in Case 1. For Cases 3 and 4 (assuming large γ), the ratio is plotted versus p, and we find that all possible values of the ratio are obtained as p varies.*

where ϕ is the polar angle measured from the $\hat{u}$ direction and

$$(3.7) \qquad\qquad \alpha \equiv \frac{\mu\lambda}{2D}$$

is a parameter new to Case 2 that affects the anisotropy of the crime distribution. This anisotropy is illustrated in Figure 3.1, where we plot the crime probability density as a function of the angle ϕ for varying values of α and observe that higher α values lead to a tighter peak around $\phi = 0$, collapsing to a delta function as $\alpha \to \infty$.

Though Case 2 incorporates anisotropy in target selection in a natural and simple way, it does not fully address the issue of variance versus mean in DTC. As a two-parameter distribution, Case 2 does allow for some independence between mean DTC and variance in DTC, but only insofar as it allows the variance to be *greater* than that of Case 1 for a given value of $\overline{r}$ (see Figure 3.2). This is because the mean DTC

in Case 2 is given by

$$(3.8) \qquad \overline{r} = \frac{\lambda}{\sqrt{1+\alpha^2}} \left[2(1+\alpha^2)E_E \left[\frac{\alpha^2}{1+\alpha^2} \right] - E_K \left[\frac{\alpha^2}{1+\alpha^2} \right] \right],$$

where E_K and E_E are the complete elliptic integrals of the first and second kind, respectively, while the variance in Case 2 is given by

$$(3.9) \qquad v_r = 4\lambda^2(1+2\alpha^2) - \overline{r}^2.$$

Hence, the ratio $v_r/\overline{r}^2$ for Case 2 is purely a function of α that is always greater than or equal to $16/\pi^2 - 1$, which is the value that this ratio takes on in Case 1. Furthermore, Case 2 continues to exhibit the divergence in $P(\mathbf{r})$ at the origin that was noted in Case 1, making it also somewhat ill-defined.

3.3. Case 3. Therefore, we look to another mechanism that will allow for greater control in DTC variance and hopefully remove the weakly divergent behavior at the origin. For this, we propose an example in which the attractiveness $A(\mathbf{x}|\mathbf{z})$ is no longer homogeneous but is rather a function of $|\mathbf{x} - \mathbf{z}|$, the distance from the target to the criminal's anchor point. Specifically, we choose

$$(3.10) \qquad A(\mathbf{x}|\mathbf{z}) = A \left(\frac{|\mathbf{x} - \mathbf{z}|}{\ell} \right)^k,$$

where ℓ is a lengthscale and $k > -2$ (see Appendix A for a justification of this restriction). This specific form for $A(\mathbf{x}|\mathbf{z})$ allows for criminals of all types: those that strongly desire to commit crimes very close to home ($k < 0$), those that have no inherent preference ($k = 0$, as in Case 1), and those that prefer to commit crimes further away from their residence ($k > 0$). Of course, other forms may be used for $A(\mathbf{x}|\mathbf{z})$, such as functions that saturate to a limiting value as $\mathbf{x}$ becomes very far from $\mathbf{z}$, but here we focus on this simple case. Case 3 will therefore assume homogeneous D, no drift (to separate the effects of anisotropy from the issue at hand), and the $A(\mathbf{x}|\mathbf{z})$ of (3.10). For $k > -2$, the isotropic crime distribution $P(\mathbf{r})$ in this case is given by

$$(3.11) \qquad P(\mathbf{r}) = \frac{1}{2\pi p \beta^2} \left(\frac{r}{\beta} \right)^k K_0 \left[\frac{1}{p} \left(\frac{r}{\beta} \right)^p \right],$$

where $p = 1 + k/2$ and the lengthscale β is given by

$$(3.12) \qquad \beta \equiv \sqrt[p]{\lambda \ell^{p-1}}.$$

The distribution (3.11) allows for the variance in DTC to take any value for a given $\overline{r}$, with values greater than those of Case 1 occurring for $0 < p < 1$ and values smaller than those of Case 1 occurring for $p > 1$ (see Figure 3.2). Here, the mean DTC is given by

$$(3.13) \qquad \overline{r} = \beta(2p)^{1/p}\Gamma^2 \left[\frac{2p+1}{2p} \right],$$

and the variance in DTC is given by

$$(3.14) \qquad v_r = \beta^2(2p)^{2/p}\Gamma^2 \left[\frac{p+1}{p} \right] - \overline{r}^2.$$

Furthermore, (3.11) no longer diverges at the origin for values of $k > 0$, eliminating that issue.[1]

3.4. Case 4. Finally, by combining the constant advection of Case 2 with the spatially varying attractiveness of Case 3, we create Case 4, which captures all the desired qualities of $P(\mathbf{x}|\mathbf{z})$: free variance versus mean of DTC (see Figure 3.2), variable anisotropy, and variable lengthscale. To begin in this case, we first assume $\mathbf{z} = 0$, $\boldsymbol{\mu} = \mu\hat{x}$ ($\hat{x}$ is the unit vector in the x-direction, chosen without loss of generality), D is homogeneous, and $A(\mathbf{x}|\mathbf{z})$ is given by (3.10). We furthermore assume for the remainder of Case 4 that $k > -1$, as we are interested only in such cases here. With these choices, (3.1) becomes

$$(3.15) \qquad D\nabla^2\rho - \mu\frac{\partial\rho}{\partial x} - A\left(\frac{r}{\ell}\right)^k \rho = \frac{\partial\rho}{\partial t}.$$

We now rewrite (3.15) in dimensionless form by scaling spatial variables by a lengthscale L and time by a timescale T, though for simplicity of notation we still refer to space and time variables by their original symbols. This gives

$$(3.16) \qquad \frac{\beta}{2\gamma L}\nabla^2\rho - \frac{\partial\rho}{\partial x} - \frac{1}{2\gamma}\left(\frac{L}{\beta}\right)^c r^{c-1}\rho = \frac{L}{\mu T}\frac{\partial\rho}{\partial t},$$

where β is defined in (3.12), $c = k + 1$, and

$$(3.17) \qquad \gamma \equiv \frac{\mu\beta}{2D},$$

which is the equivalent of α from Case 2. We now choose L and T to be

$$(3.18) \qquad L = \beta\sqrt[c]{2\gamma c}, \; T = L/\mu,$$

which are the natural advective lengthscales and timescales for Case 4. With these choices, (3.16) becomes

$$(3.19) \qquad \xi\nabla^2\rho - \frac{\partial\rho}{\partial x} - cr^{c-1}\rho = \frac{\partial\rho}{\partial t},$$

where

$$(3.20) \qquad \xi \equiv \frac{1}{2\gamma\sqrt[c]{2\gamma c}}.$$

Let us now focus on the case $\gamma \gg 1$ so that $\xi \ll 1$ and we ignore the diffusive term in (3.19). Here, since the initial condition for ρ is a delta function at the origin, and our simplified (3.19) contains only advection along the x direction with speed 1, we have $r = x = t$. The solution in this limiting case is therefore

$$(3.21) \qquad \rho(\mathbf{x}, t) \propto e^{-t^c}\delta(x - t)\delta(y).$$

Upon integrating over all t to find $\rho(\mathbf{x}|\mathbf{z})$, multiplying by the attractiveness to find $P(\mathbf{r})$, and converting back to dimensional variables, we find

$$(3.22) \qquad P(\mathbf{r}) = \frac{c}{Lr}\left(\frac{r}{L}\right)^{c-1}e^{-(r/L)^c}\delta(\phi).$$

---

[1]Of course, for $k < 0$, the divergence at the origin is even stronger than in Case 1. However, if the behavior associated with negative k is desired, other functional forms for $A(\mathbf{x}|\mathbf{z})$ may be used instead of (3.10) that do not diverge at the origin but still preserve the same general behavior.

GEORGE O. MOHLER AND MARTIN B. SHORT

Hence, the DTC distribution is

$$(3.23) \qquad P(r) = \frac{c}{L}\left(\frac{r}{L}\right)^{c-1} e^{-(r/L)^c}.$$

Interestingly, then, for $\gamma \gg 1$, the DTC distribution $P(r)$ of Case 4 approaches a Weibull distribution with shape parameter $c = k+1$ (which is equivalent to a Rayleigh distribution for $k = 1$) and lengthscale L given above. This is because in this case, advection dominates the criminal's motion, sending him mostly in a straight line along which he commits crimes at a rate that is proportional to the distance he has traveled to the power k. The Weibull distribution, meanwhile, describes the "time to failure" of an item whose failure rate is proportional to the time elapsed to the power k. Hence, a Weibull distribution should be expected in this case, as the distance a criminal moves before committing a crime can be interpreted as a "distance to failure," since we assume the criminal is no longer active after committing one crime in a given outing.

Going back now to (3.19), we wish to determine an approximate solution when the diffusive term is not completely ignored but is still considered small. We first integrate (3.19) over all dimensional t, so that we are now finding $\rho(\mathbf{x}|\mathbf{z}) = \rho(\mathbf{r})$. Note that the advective term in (3.19) and the dimensions of $\rho(\mathbf{r})$ can then be removed via the substitution

$$(3.24) \qquad \rho(\mathbf{r}) = \frac{T}{L^2} e^{x/2\xi} g(r),$$

so that the radially symmetric, dimensionless function $g(r)$ satisfies

$$(3.25) \qquad \xi\nabla^2 g - \left[\frac{1}{4\xi} + cr^{c-1}\right] g = -\delta(\mathbf{r}),$$

where r is still dimensionless. Now, whatever the solution to (3.25) may be, we know that the distribution $P(\mathbf{r})$ is given, in dimensional variables, by

$$(3.26) \qquad P(\mathbf{r}) = \frac{c}{L^2}\left(\frac{r}{L}\right)^{c-1} e^{\gamma r \cos\phi/\beta} g(r/L).$$

The DTC distribution, then, is given by

$$(3.27) \qquad P(r) = \frac{2\pi cr}{L^2}\left(\frac{r}{L}\right)^{c-1} I_0(\gamma r/\beta) g(r/L).$$

At this point, we make the assumption that the DTC distribution will be well approximated by the Weibull distribution (3.23) as long as $\gamma \gg 1$, even if we do not completely ignore the diffusive term in (3.19), for the reasons outlined above. Using this assumption, then, we equate (3.27) and (3.23), solve for $g(r/L)$, and substitute back into (3.26). We then make use of the fact that for large arguments (as when $\gamma \gg 1$),

$$(3.28) \qquad I_0(x) \approx \frac{e^x}{\sqrt{2\pi x}}.$$

Thus, an approximation for the exact $P(\mathbf{r})$ is

$$(3.29) \qquad P(\mathbf{r}) \approx \frac{c}{L}\sqrt{\frac{q}{2\pi r}}\left(\frac{r}{L}\right)^{c-1} e^{-(r/L)^c} e^{-qr[1-\cos\phi]},$$
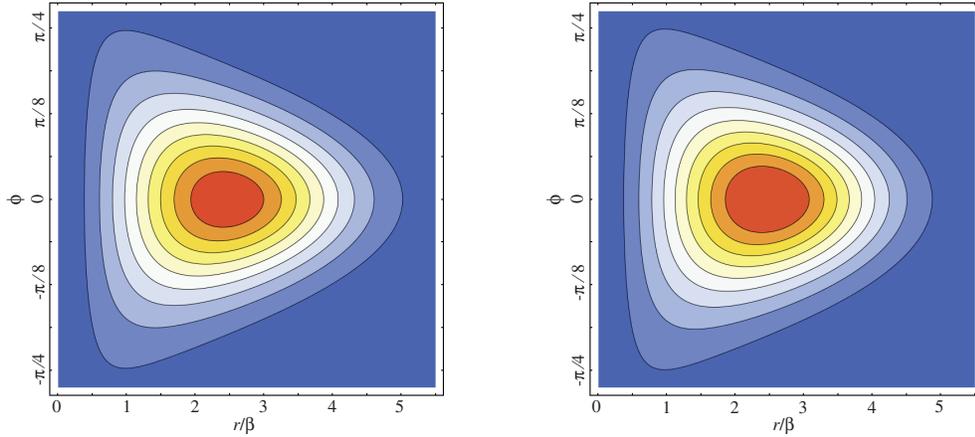
where $q = \gamma/\beta$.

FIG. 3.3. *The distribution of criminal offenses $P(\mathbf{r})$ as a function of scaled distance r/β and angle ϕ (measured from the x-axis) for the exact solution of Case 4 (left) and an approximation to this solution (right) as given in (3.29), using $k = 2$ and $\gamma = 5$. The contour lines begin at a value 0.02 (the largest contour) and proceed to the value 0.18 (the smallest contour) in steps of 0.02.*

Note that the distribution (3.29), which we shall use later as the "solution" for Case 4, should not be interpreted as a rigorous solution following from (3.15). Instead, one may consider it a distribution *inspired by* the true solution to Case 4, exhibiting the same general qualitative behavior at "large" γ but being much easier to evaluate than the true solution, which must in general be solved for numerically. However, to illustrate the similarity between our approximate form and the true solution for Case 4, we note that an exact solution to (3.25) is available for the special case $k = 2$. This solution is

$$(3.30) \qquad g(r/L) \propto \mathrm{Exp}\left[-\frac{1}{2}\left(\frac{r}{\beta}\right)^2\right] \mathrm{U}\left[\frac{1}{2}+\left(\frac{\gamma}{2}\right)^2, 1, \left(\frac{r}{\beta}\right)^2\right],$$

where $\mathrm{U}[a, b, c]$ is the confluent hypergeometric function. In Figure 3.3, the distribution $P(\mathbf{r})$ generated from this exact solution is compared to that generated by (3.29) in the case $\gamma = 5$; the agreement is reasonably good, even for this relatively low γ value.

Finally, we note that the distribution (3.29) allows for the variance in DTC to take any value for a given $\overline{r}$, as was true in Case 3. Since the DTC distribution in this case is Weibull, the mean DTC is given by

$$(3.31) \qquad \overline{r} = L\Gamma\left[\frac{c+1}{c}\right],$$

and the variance in DTC is given by

$$(3.32) \qquad v_r = L^2\Gamma\left[\frac{c+2}{c}\right] - \overline{r}^2.$$

3.5. A remark on (3.3). In general, of course, $\int_0^\infty \rho(\mathbf{x}, t|\mathbf{z})dt$ is not necessarily well-defined for arbitrary drift $\boldsymbol{\mu}(\mathbf{x})$ and attractiveness $A(\mathbf{x}|\mathbf{z})$. However, as shown above, the integral is well-defined when we consider Case 3 with $k > 0$ and in Case 4.

In practice, (3.1) and (3.4) need to be considered on a finite domain to facilitate numerical approximation. Here, either physically realistic boundary conditions need to be imposed or the domain should be taken large enough such that the event of the random walker given by (2.1) reaching the boundary without committing a crime has low probability. In this paper, we consider a 140 km^2 domain containing Los Angeles, and in numerical experiments with Dirichlet and Neumann boundary conditions we find that $\rho(\mathbf{x}, t|\mathbf{z})$ decays sufficiently fast for the integral to be well-defined.

3.6. Adjoint equation. Given the solution of (3.4), along with a prior distribution of criminal anchor points $P(\mathbf{z})$, the geographic profiling distribution of a single criminal event can then be determined using Bayes' theorem,

$$(3.33) \qquad P(\mathbf{z}|\mathbf{x}) = \frac{P(\mathbf{x}|\mathbf{z})P(\mathbf{z})}{\int_{\mathbb{R}^2} P(\mathbf{x}|\mathbf{z}')P(\mathbf{z}')d\mathbf{z}'} = \frac{A(\mathbf{x}|\mathbf{z})\rho(\mathbf{x}|\mathbf{z})P(\mathbf{z})}{\int_{\mathbb{R}^2} A(\mathbf{x}|\mathbf{z}')\rho(\mathbf{x}|\mathbf{z}')P(\mathbf{z}')d\mathbf{z}'}.$$

A similar procedure can be carried out for multiple crimes. Assuming event independence and that all crimes were committed by the same person, the geographic profiling density for multiple events is given by

$$(3.34) \qquad P(\mathbf{z}|\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{P(\mathbf{x}_1|\mathbf{z}) \cdots P(\mathbf{x}_N|\mathbf{z})P(\mathbf{z})}{\int_{\mathbb{R}^2} P(\mathbf{x}_1|\mathbf{z}') \cdots P(\mathbf{x}_N|\mathbf{z}')P(\mathbf{z}')d\mathbf{z}'}.$$

Implementing (3.34) is highly inefficient, though, as in principle one must solve (3.4) for each potential anchor point on a numeric grid (which may have tens of thousands of points), only to evaluate that solution at the N locations where crimes actually occurred.

However, since $\rho(\mathbf{x}|\mathbf{z})$ is the Green's function corresponding to the linear operator on the left side of (3.4), for fixed $\mathbf{x}$ and varying $\mathbf{z}$ the function $f(\mathbf{z}) = \rho(\mathbf{x}|\mathbf{z})$ solves the "backward" or "adjoint" [1, 7, 12] equation

$$(3.35) \qquad -\nabla \cdot (D\nabla f) - \boldsymbol{\mu}(\mathbf{z}) \cdot \nabla f + A(\mathbf{z}|\mathbf{x})f = \delta(\mathbf{z} - \mathbf{x}),$$

where ∇ is now with respect to the variable $\mathbf{z}$. Here, the point mass on the right-hand side is located at the scene of the crime, rather than at the potential anchor point. We also note that the first order derivative term changes sign going from the forward equation (3.4) to the backward equation (3.35). This has practical implications, for if criminals move up gradients of attractiveness then police investigations starting from the scene of the crime should move down gradients of attractiveness. Thus, the geographic profiling density can be efficiently computed in practice by solving the backward equation given by (3.35) only N times and is given by

$$(3.36) \qquad P(\mathbf{z}|\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{P(\mathbf{z}) \prod_{i=1}^{N} A(\mathbf{x}_i|\mathbf{z}) f_i(\mathbf{z})}{\int_{\mathbb{R}^2} P(\mathbf{z}') \prod_{i=1}^{N} A(\mathbf{x}_i|\mathbf{z}') f_i(\mathbf{z}')d\mathbf{z}'},$$

where $f_i(\mathbf{z})$ is the solution to (3.35) with $\mathbf{x} = \mathbf{x}_i$.

Of course, the solution to (3.35) requires the specification of a vector of parameters $\boldsymbol{\theta}$. Hence, a prior distribution $\pi(\boldsymbol{\theta})$ for the parameter vector $\boldsymbol{\theta}$ can be incorporated into the modeling framework, so that

$$(3.37) \qquad P(\mathbf{z}|\mathbf{x}_1, \ldots, \mathbf{x}_N) \propto \int P(\mathbf{x}_1|\mathbf{z}, \boldsymbol{\theta}) \cdots P(\mathbf{x}_N|\mathbf{z}, \boldsymbol{\theta})P(\mathbf{z})\pi(\boldsymbol{\theta})d\boldsymbol{\theta},$$

as discussed in [16]. The prior distribution π provides a means of incorporating information from historic crime series into the model.

4. Fitting the forward model to crime data. For a given crime series with anchor point $\mathbf{z}$ and crime locations $\mathbf{x}_1, \ldots, \mathbf{x}_N$, the parameters of the model need to be estimated before a geographic profile can be computed. In the absence of drift, (3.4) can be efficiently solved using Multigrid [3] for a given set of parameters. The best fit parameter vector, $\hat{\boldsymbol{\theta}}$, can then be estimated by maximizing the log-likelihood function

$$(4.1) \qquad \hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{N} \log\left(P(\mathbf{x}_i|\mathbf{z}, \boldsymbol{\theta})\right).$$

However, including drift significantly complicates numerical solutions of (3.4) and (3.35). Hence, in these cases we restrict ourselves to the regime of Case 4 from section 3.4, for which the approximate solution (3.29) is available. In (3.29), we assumed the preferred drift direction of the criminal was the x-axis, but we now allow the preferred direction to be any vector $\hat{u}$ at angle ϕ_0 from the x-axis, giving us the distribution

$$(4.2) \qquad P(\mathbf{r}) \approx \frac{c}{L}\sqrt{\frac{q}{2\pi r}}\left(\frac{r}{L}\right)^{c-1} e^{-(r/L)^c} e^{-qr[1-\cos(\phi-\phi_0)]}.$$

The advantage of this form of the model is that it can be easily evaluated (without numerically solving a PDE) and can be fit easily to crime series data. Using maximum likelihood estimation, the parameters c and L can be found by fitting the DTC data to a Weibull distribution using standard techniques [17]. Likewise, the parameter ϕ_0 is found to be

$$(4.3) \qquad \phi_0 = \arctan(\overline{y}/\overline{x}),$$

where $\overline{x}$ and $\overline{y}$ are the average x and y distances from the anchor point to the crimes in the given series. Finally, the parameter q is found to be

$$(4.4) \qquad q = \frac{1}{2\left[\overline{r} - \overline{x}\cos(\phi_0) - \overline{y}\sin(\phi_0)\right]},$$

where $\overline{r}$ is again the average DTC for the series.

For illustration, we fit (4.2) to residential burglary data collected by the Los Angeles Police Department from 2003 through 2008. The data consists of 221 solved burglary series in Los Angeles, including the geocoded locations where the burglaries occurred and the corresponding residences of the offenders (presumed to be their anchor points). Each crime series includes between 3 and 32 burglaries (the average is 3.76), where at least two of the crimes occur at different locations. For each series, we estimate L, c, q, and ϕ_0 as indicated above, and in Figure 4.1 we plot the resulting parameter distributions across the 221 crime series. In the top left, we plot the estimated distribution of ϕ_0, which provides a measure of the directional bias of burglars in Los Angeles. We note that the most prevalent directions for criminals to travel are northwest and southeast, the directions that the major freeways US-101, I-5, and I-405 are oriented. In the top right and lower left, we plot the distributions of the estimated anisotropy and shape parameters q and c, respectively. High values of these parameters correspond to a higher degree of commuting, whereas lower values correspond to marauding. As discussed above, these parameters allow for a spectrum of commuting, as opposed to a binary definition.
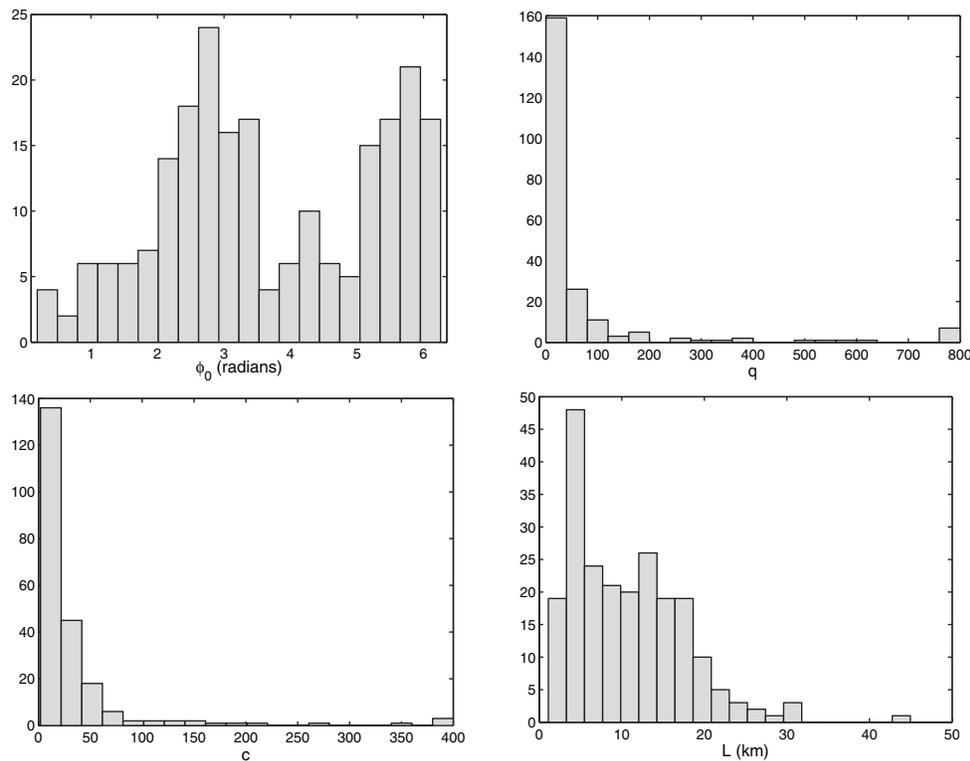
FIG. 4.1. *Clockwise from top left to lower left: frequency histograms of estimated parameters* ϕ_0, q, L, *and* c *of* (4.2) *for the* 221 *crime series in the Los Angeles data set.*

5. Geographic profiling of residential burglars in Los Angeles. Next, we illustrate the implementation of our methodology for the purpose of geographic profiling using the 221 crime series from section 4. The first model we consider makes use of the density given by (4.2), an approximation to the solution of the steady state advection-diffusion-reaction equation (3.4) under the assumptions of Case 4 in section 3.4. We estimate the parameters as outlined in section 4, though to prevent numerical under- or overflow when computing the profiles we impose $c \leq 40$ and $q \leq 200$ as constraints in the maximization of the likelihood. For a given crime series m with crime locations $\{\mathbf{x}_i^m\}_{i=1}^{N_m}$ and anchor point $\mathbf{z}^m$, we compute the geographic profile using the leave-one-out method,

$$P(\mathbf{z}|\mathbf{x}_1^m, \ldots, \mathbf{x}_{N_m}^m) \propto \int P(\mathbf{x}_1^m|\mathbf{z}, \boldsymbol{\theta}) \cdots P(\mathbf{x}_{N_m}^m|\mathbf{z}, \boldsymbol{\theta}) P(\mathbf{z})\pi(\boldsymbol{\theta})d\boldsymbol{\theta},$$

$$(5.1) \qquad \approx \frac{P(\mathbf{z})}{T-1} \sum_{\substack{j=1 \\ j \neq m}}^{T} \prod_{i=1}^{N_m} P(\mathbf{x}_i^m|\mathbf{z}, L^j, \phi_0^j, q^j, c^j),$$

where T is the total number of crime series and we are assuming a uniform (atomic) prior over the estimated parameters. Additionally, we assume that criminals are uniformly distributed amongst all residences within the city and set $P(\mathbf{z})$ to be proportional to housing density $H(\mathbf{z})$. It should be noted that our housing density information is given on a 128×128 grid of Los Angeles (140^2 km) and thus the geographic

profile is also a grid function. We will refer to this geographic profiling algorithm as the forager inspired distribution method, case α (FIND$_\alpha$).

Though method FIND$_\alpha$ incorporates many aspects of criminal motion and target selection, it assumes that any spatial variation in $A(\mathbf{x}|\mathbf{z})$ is due only to the distance between $\mathbf{x}$ and $\mathbf{z}$, ignoring any intrinsic heterogeneity of crime targets. The second model we consider will incorporate such intrinsically heterogeneous attractiveness but will ignore advection, allowing for quick numerical solutions of (3.35). For simplicity, we assume that all houses are equally attractive, so that

$$\tag{5.2} A(\mathbf{x}|\mathbf{z}) = AH(\mathbf{x}),$$

similar to Case 1 from section 3.1. As with method FIND$_\alpha$, $P(\mathbf{z})$ is taken to be proportional to housing density $H(\mathbf{z})$ as well. As seen in the discussion of Case 1 above, this method possesses one effective parameter: a lengthscale λ. We then proceed similarly to FIND$_\alpha$, first estimating λ via maximum likelihood for each crime series as outlined in section 4 when no drift is present. In particular, the negative log-likelihood is minimized using the MATLAB routine fminbnd with the bounds chosen to be $10^{-4} \le \lambda \le 10^2$. To compute $P(\mathbf{x}|\mathbf{z}, \lambda)$, the backward equation (3.35) with no drift and with the attractiveness field (5.2) is solved using Multigrid on a 140 km^2 domain with a 128×128 resolution and Dirichlet boundary conditions. We then compute the geographic profiles using (5.1) with L^j, ϕ_0^j, q^j, c^j replaced by λ^j. We will refer to this geographic profiling method as FIND$_\beta$.

Because the prior parameter distributions, $\pi(\boldsymbol{\theta})$, are approximated by uniform (atomic) distributions over our empirical parameters, the resulting geographic profiles of FIND$_\alpha$ and FIND$_\beta$ are noisier than geographic profiles used in practice. To improve visualization and create slightly smoother probability surfaces, we apply a Gaussian filter to the geographic profiles as a last step in the algorithm; the bandwidth of the filter is chosen to be small, the length of one grid cell.

We compare FIND$_\alpha$ and FIND$_\beta$ with the widely used criminal geographic targeting (CGT) algorithm outlined in [16, 18, 20]. CGT uses a score function of the form (1.1), where

$$\tag{5.3} f(d) = \begin{cases} \frac{k}{d^h}, & d > B, \\ \frac{kB^{g-h}}{(2B-d)^g}, & d \le B, \end{cases}$$

and d is the Manhattan distance metric. In this study, we use the parameters $h = g = 1.2$ and B is taken to be $1/2$ the average nearest neighbor distance between the crimes in the series, as recommended in [19, 20].

In Figure 5.1, density plots of geographic profiles are exhibited for a marauder series and a commuter series using FIND$_\alpha$, FIND$_\beta$, and CGT. By construction, the Bayesian models take into account geographic heterogeneity of potential anchor points and place probability mass only in valid regions of the city (residential areas). In comparison, the CGT algorithm places some probability mass in commercial zones, mountains, and the ocean. In comparison to FIND$_\alpha$, the probability mass of the profiles of CGT and FIND$_\beta$ are tightly clustered around the crimes of each series. In the case of CGT (and other methods using a score function similar to (1.1)), the profile can actually be interpreted as a kernel density estimate of the *next crime*, and thus the unknown criminal is estimated to live near the crime sites; i.e., the criminal is assumed to be a marauder. In the case of FIND$_\beta$, the assumptions of event independence and no drift lead to the clustering of the density near the crime sites.
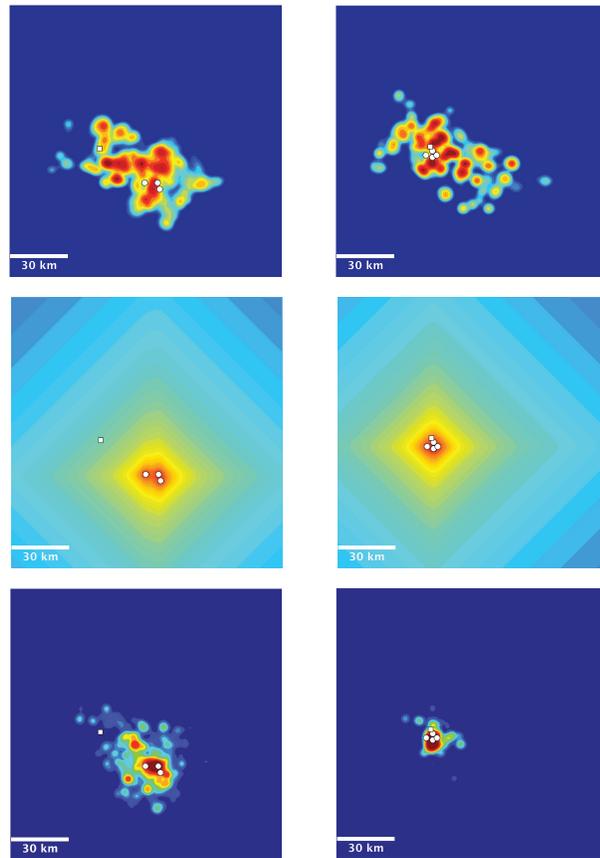
Fig. 5.1. *Geographic profile densities (on a logarithmic scale) of a commuter series (left) and marauder series (right) using FIND$_\alpha$ (top), CGT (center), and FIND$_\beta$ (bottom). Crime locations are shown as small circles, and the anchor points are shown as small squares.*

To compare the efficacy of CGT with FIND$_\alpha$ and FIND$_\beta$, we use a metric similar to the "hit score" often used in geoprofiling [19]. Here, the city is first divided into square cells of size $(140/128)^2$ km^2. For each burglary series in the data set, we then construct a geoprofile using each of the three methods as described above. For each method, the cells of the geoprofile are ordered based upon the probability mass (or score, in the case of CGT) located in each cell, from lowest to highest, and the top F fraction of these ordered cells are "flagged" as being most likely to contain the anchor point. Because CGT can result in multiple cells having the same score, we assign a ranking to each of these nonunique cells such that the true anchor point is located in the highest ranked cell (thus we do not underestimate the accuracy of CGT). We then determine whether the actual anchor point lies within one of these flagged cells for each method. After performing this procedure for each of the series in our data, we determine the overall fraction of anchor points that did indeed lie within a flagged cell for each of the three methods.

In Figure 5.2, we plot the fraction of criminals (anchor points) correctly located versus varying values of F. For small fractions of the city flagged, the algorithms that emphasize marauding, CGT and FIND$_\beta$, slightly outperform FIND$_\alpha$. For intermediate fractions, FIND$_\alpha$ is the most accurate, as this method is better suited for ranking
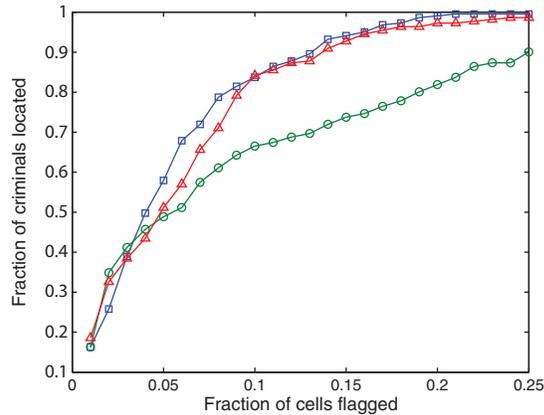
FIG. 5.2. *Fraction of criminal residences predicted by model vs. fraction of the city flagged for* FIND$_\alpha$ *(blue squares),* FIND$_\beta$ *(red triangles), and CGT (green circles).*

the cells where commuters are likely to live. Eventually all methods plateau at 100% accuracy at $F < 1$, since housing density is nonzero in only a fraction of the entire domain.

6. Discussion. We developed a new framework for geographic profiling based upon Bayes' theorem and kinetic descriptions of criminal behavior. In the future it may prove fruitful to consider more general models for the forward equation of the form

$$(6.1) \qquad -L\rho + \nabla \cdot (\boldsymbol{\mu}(\mathbf{x})\rho) + A(\mathbf{x}|\mathbf{z})\rho = \delta(\mathbf{x} - \mathbf{z}).$$

Here L could correspond to fractional diffusion [25], as a heavy tailed distribution may be more appropriate in the case of more serious crimes (for which geographic profiling is typically used). Another possibility would be to model the diffusion parameter in (2.1) as a stochastic process, along the lines of stochastic volatility models used in financial modeling [8]. In both cases, however, care would be required so that any probability distributions found via these methods would be well-defined.

It also may be advantageous to incorporate heterogeneous criminal mobility, i.e., $L = \nabla \cdot (D(\mathbf{x})\nabla)$, into geographic profiling estimates. For example, in the city of Los Angeles, traffic could play an important role in determining the location of anchor points. Anchor points equidistant (spatially) from the crime, but with varying *time* distances from the crime, should have different probability weights. Furthermore, geographic obstacles such as parks and bodies of water could also be included in such estimates, as criminals must diffuse around these objects (not through) to reach targets on the other side.

For other types of crime where housing density does not play a role in target selection, such as person-to-person crimes or auto theft, the attractiveness field can be estimated from historical crime data. First, the marginal probability density of crime $P_c(\mathbf{x})$, which for a given model satisfies

$$(6.2) \qquad P_c(\mathbf{x}) = \int_{\mathbb{R}^2} A(\mathbf{x}|\mathbf{z})\rho(\mathbf{x}|\mathbf{z})P(\mathbf{z})d\mathbf{z},$$

can be inferred from aggregate spatial crime data using density estimation techniques for spatial point processes [15, 24]. Then, for a given estimate $\hat{P}_c(\mathbf{x})$, the attractiveness

field A could be estimated by minimizing

$$(6.3) \qquad \int_{\mathbb{R}^2} \left(\hat{P}_c(\mathbf{x}) - \int_{\mathbb{R}^2} A(\mathbf{x}|\mathbf{z})\rho(\mathbf{x}|\mathbf{z})P(\mathbf{z})d\mathbf{z} \right)^2 d\mathbf{x}$$

over A, subject to positivity and regularity constraints.

Additionally, it has been shown that the attractiveness field is nonstationary over timescales of several days to several months, stemming from self-exciting effects of recent criminal activity [14, 22, 23]. Here, the estimated marginal density of crime $\hat{P}_c(\mathbf{x})$ could be reconstructed at the time of an offense using self-exciting point processes fit to crime data [14] from the preceding several weeks or months.

Appendix. We now justify the restriction $k > -2$ for (3.4) and (3.10). Take first the case $k = -2$. Here, for $r > 0$, a solution $\rho(r)$ to (3.4) satisfies the Euler equation

$$(A.1) \qquad r^2\rho''(r) + r\rho'(r) - \frac{\ell^2}{\lambda^2}\rho(r) = 0;$$

hence, $\rho(r)$ is given by

$$(A.2) \qquad \rho(r) = Cr^{\ell/\lambda} + Er^{-\ell/\lambda}.$$

Though the above is necessary, it is not sufficient in order for (A.2) to be a solution of (3.4). That is, it must also be the case that

$$(A.3) \qquad \int_{\mathbb{R}^2} \left[D\nabla\rho\nabla\phi + \frac{A\ell^2}{r^2}\rho\phi \right] dx = \phi(0)$$

for all $\phi \in C_0^\infty(\mathbb{R}^2)$. Upon substituting (A.2) into (A.3), we see that $E = 0$ must hold, as the term

$$\frac{EA\ell^2}{r^2}r^{-\ell/\lambda}\phi$$

is not integrable over $\mathbb{R}^2$ if $\phi(0) \neq 0$, due to the strong divergence at the origin. To investigate the term with coefficient C in (A.2), we fix $\epsilon > 0$ and choose ϕ in (A.3) so that $\phi = 1$ in the ball B_ϵ and $\phi = 0$ outside the ball $B_{\epsilon+\delta}$. Upon sending $\delta \downarrow 0$, we see that (A.3) implies that ρ must satisfy

$$(A.4) \qquad \int_{r=\epsilon} D(\nabla\rho \cdot \mathbf{n}) \, d\sigma + \iint_{r<\epsilon} \frac{A\ell^2}{r^2}\rho \, dx = 1.$$

Again substituting (A.2), this time with $E = 0$, we see that

$$\iint_{r<\epsilon} A\frac{\ell^2}{r^2}\rho \, dx = 2\pi AC\ell^2 \int_0^\epsilon r^{\ell/\lambda-1} \, dr = 2\pi AC\ell\lambda\epsilon^{\ell/\lambda} \xrightarrow{\epsilon\downarrow 0} 0,$$

while

$$\int_{r=\epsilon} D(\nabla\rho \cdot \mathbf{n}) \, d\sigma = \frac{2\pi DC\ell}{\lambda}\epsilon\left(\epsilon^{\ell/\lambda-1}\right) \xrightarrow{\epsilon\downarrow 0} 0.$$

Thus, (A.4), and thereby (A.3), cannot be satisfied in the $k = -2$ case, so no solution exists here that is valid in all of $\mathbb{R}^2$.

The situation is essentially the same for cases where $k < -2$. Here, for $r > 0$, the solution to (3.4) is

$$(A.5) \qquad \rho(r) = C \mathrm{K}_0 \left[\frac{1}{|p|} \left(\frac{r}{\beta} \right)^p \right] + E \mathrm{I}_0 \left[\frac{1}{|p|} \left(\frac{r}{\beta} \right)^p \right]$$

with β and p the same as given above and $\mathrm{I}_0(x)$ the zeroth order modified Bessel function of the first kind. Since $k < -2$, we have $p < 0$, so that when (A.5) is substituted into (A.3), we again see that $E = 0$ must be satisfied, as the term proportional to $E r^k \mathrm{I}_0(r^p)$ is too heavily divergent at the origin to be integrable. To investigate the term with coefficient C, we do as above, finding that

$$\iint_{r<\epsilon} A \left(\frac{r}{\ell} \right)^k \rho \, dx = \frac{2\pi AC}{\ell^k} (\beta\epsilon)^p \mathrm{K}_1 \left[\frac{1}{|p|} \left(\frac{\epsilon}{\beta} \right)^p \right] \xrightarrow{\epsilon \downarrow 0} 0$$

and

$$\int_{r=\epsilon} D(\nabla\rho \cdot \mathbf{n}) \, d\sigma = 2\pi DC\epsilon \left(\frac{\epsilon}{\beta} \right)^{p-1} \mathrm{K}_1 \left[\frac{1}{|p|} \left(\frac{\epsilon}{\beta} \right)^p \right] \xrightarrow{\epsilon \downarrow 0} 0.$$

Therefore, the situation here is the same as seen in the $k = -2$ case: one of the two potential solutions is too strongly divergent at the origin to be integrable, while the other solution is not divergent enough to replicate the delta function present in (3.4). Thus, the restriction $k > -2$.

Acknowledgments. The authors would like to thank Lincoln Chayes, Andrea Bertozzi, Jeff Brantingham, and George Tita for helpful discussions and Sean Malinowski and Nathan Ong of the Los Angeles Police Department for providing the burglary data used in this study. The authors especially acknowledge Mike O'Leary for sharing his expertise on distance-to-crime patterns, which led to significant improvements in the manuscript.

REFERENCES

[1] G. BARTON, *Elements of Green's Functions, Waves, and Propagation: Potentials, Diffusion, and Waves*, Clarendon Press, Oxford, UK, 1989.
[2] P.J. BRANTINGHAM AND G. TITA, *Offender mobility and crime pattern formation from first principles*, in Artificial Crime Analysis Systems, L. Liu and J. Eck, eds., IGI Global, Hershey, PA, 2008.
[3] W.L. BRIGGS, V. EMDEN HENSON, AND S.F. MCCORMICK, *A Multigrid Tutorial*, SIAM, Philadelphia, 2000.
[4] D. CANTER, T. COFFEY, M. HUNTLEY, AND C. MISSEN, *Predicting serial killers' home base using a decision support system*, J. Quant. Criminol., 16 (2000), pp. 457–278.
[5] D. CANTER AND P. LARKIN, *The environmental range of serial rapists*, J. Environmental Psychology, 13 (1993), pp. 63–69.
[6] R.V. CLARKE AND M. FELSON, eds., *Routine Activity and Rational Choice: Advances in Criminological Theory*. vol. 5, Transaction Books, New Brunswick, NJ, 1993.
[7] D. ESTEP, *A Short Course on Duality, Adjoint Operators, Green's Functions, and A Posteriori Error Analysis*, www.math.colostate.edu/~estep/research/preprints/adjointcourse_final.pdf (2004).
[8] J.-P. FOUQUE, G. PAPANICOLAOU, AND K.R. SIRCAR, *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press, UK, 2000.
[9] E. GROFF, *Characterizing the spatio-temporal aspects of routine activities and the geographic distribution of street robbery*, in Artificial Crime Analysis Systems, L. Liu and J. Eck, eds., IGI Global, Hershey, PA, 2008.
[10] D. HOLCMAN, A. MARCHEWKA, AND Z. SCHUSS, *Survival probability of diffusion with trapping in cellular neurobiology*, Phys. Rev. E, 72 (2005), p. 031910.

[11] S.D. JOHNSON, L. SUMMERS, AND K. PEASE, *Offender as forager? A direct test of the boost account of victimization*, J. Quant. Criminol., 25 (2009), pp. 181–200.

[12] A. KEATS, E. YEE, AND F.-S. LIEN, *Bayesian inference for source determination with applications to a complex urban environment*, Atmos. Environment, 41 (2007), pp. 465–479.

[13] N. LEVINE, *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v. 3.1)*, Ned Levine and Associates, Houston, TX, National Institute of Justice, Washington, D.C., 2007.

[14] G. MOHLER, M. SHORT, P. BRANTINGHAM, F. SCHOENBERG, AND G. TITA, *Self-exciting point process modeling of crime*, J. Amer. Statist. Assoc., 106 (2011), pp. 100–108.

[15] G. MOHLER, A. BERTOZZI, T. GOLDSTEIN, AND S. OSHER, *Fast TV regularization for 2D maximum penalized likelihood estimation*, J. Comput. Graph. Statist., 20 (2011), pp. 479–491.

[16] M. O'LEARY, *The mathematics of geographic profiling*, J. Investigative Psychology Offender Profiling, 6 (2009), pp. 253–265.

[17] R. HORST, *The Weibull Distribution: A Handbook*, CRC Press, Boca Raton, FL, 2008.

[18] K. ROSSMO, *Geographic Profiling*, CRC Press, Boca Raton, FL, 2000.

[19] N.E. RAINE, D.K. ROSSMO, AND S.C. LE COMBER, *Geographic profiling applied to testing models of bumble-bee foraging*, J. Roy. Soc. Interface, 6 (2009), pp. 307–319.

[20] D.K. ROSSMO, *Expert System Method of Performing Crime Site Analysis*, U.S. Patent 5,781,704, 1998.

[21] A. SCHUSS, *Theory and Applications of Stochastic Differential Equations*, Wiley Ser. Probab. Stat., Wiley-Interscience, New York, 1980.

[22] M.B. SHORT, M.R. D'ORSOGNA, V.B. PASOUR, G.E. TITA, P.J. BRANTINGHAM, A.L. BERTOZZI, AND L. CHAYES, *A statistical model of criminal behavior*, M3AS, 18 (2008), pp. 1249–1267.

[23] M.B. SHORT, ET AL., *Measuring and modeling repeat and near-repeat burglary effects*, J. Quant. Criminol., 25 (2009), p. 325.

[24] B.W. SILVERMAN, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, UK, 1986.

[25] C. TADJERAN, M.M. MEERSCHAERT, AND H.-P. SCHEFFLER, *A second-order accurate numerical approximation for the fractional diffusion equation*, J. Comput. Phys., 213 (2006), pp. 205–213.

[26] X. WANG, L. LIU, AND J. ECK, *Crime simulation using GIS and artifical intelligence*, in Artificial Crime Analysis Systems, L. Liu and J. Eck, eds., IGI Global, Hershey, PA, 2008.