**Title**
Incorporating Spatial Information into Density Estimates and Street Gang Models



**Author**
Smith, Laura

**Publication Date**
2012

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

# Incorporating Spatial Information into Density Estimates and Street Gang Models

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

**Laura Michelle Smith**

2012

Abstract of the Dissertation

# Incorporating Spatial Information into Density Estimates and Street Gang Models

by

**Laura Michelle Smith**

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2012

Professor Andrea L. Bertozzi, Chair

The spatial features within a region influence many processes in human activity. Mountains, lakes, oceans, rivers, freeways, population densities, housing densities, and road networks are examples of geographical factors that impact spatial behaviors. Separated into two parts, the work presented here incorporates this information into both density estimation methods and models of street gang rivalries and territories.

Part I discusses methods for producing a probability density estimate given a set of discrete event data. Common methods of density estimation, such as Kernel Density Estimation, do not incorporate geographical information. Using these methods could result in non-negligible portions of the support of the density in unrealistic geographic locations. For example, crime density estimation models that do not take geographic information into account may predict events in unlikely places such as oceans, mountains, etc. To obtain more geographically accurate density estimates, a set of Maximum Penalized Likelihood Estimation methods based on Total Variation norm and $H_1$ Sobolev semi-norm regularizers in conjunction with *a priori* high resolution spatial data is proposed. These methods are applied to a residential burglary data set of the San Fernando Valley using geographic features obtained from satellite images of the region and housing density information.

Part II addresses the behaviors and rivalries of street gangs and how the spatial characteristics of the region affect the dynamics of the system. Gangs typically claim a specific

territory as their own, and they tend to have a set space, a location they use as a center for their activities within the territory. The spatial distribution of gangs influences the rivalries that develop within the area. One stochastic model and one deterministic model are proposed, providing different types of outputs. Both models incorporate important geographical features from the region that would inhibit movement, such as rivers and large highways. In the stochastic method, an agent-based model simulates the creation of street gang rivalries. The movement dynamics of agents are coupled to an evolving network of gang rivalries, which is determined by previous interactions among agents in the system. Basic gang data, geographic information, and behavioral dynamics suggested by the criminology literature are integrated into the model. The deterministic method, derived from a stochastic approach, modifies a system of partial differential equations from a model for coyotes. Territorial animals and street gangs often exhibit similar behavioral characteristics. Both groups have a home base and mark their territories to distinguish claimed regions. To analyze the two methods, the Hollenbeck policing division of the Los Angeles Police Department is used as a case study.

The dissertation of Laura Michelle Smith is approved.

Luminita Vese

P. Jeffrey Brantingham

Chris Anderson

Andrea L. Bertozzi, Committee Chair

University of California, Los Angeles

2012

Table of Contents

# List of Figures

# List of Tables

# Acknowledgments

Looking back over my path to finishing the doctoral degree, there are many people that have supported me, believed in me, and helped me along the way. I am forever grateful to all of them, and I would like to acknowledge a few that have exceedingly made a difference in my life.

I would first like to thank my advisor and mentor, Professor Andrea Bertozzi, for her support and guidance throughout my doctoral program. She gave me a summer project that led to my first publication, and the rest is history. I owe her much gratitude and appreciation.

I would also like to acknowledge the members of my doctoral committee, Professors Chris Anderson, Luminita Vese, and P. Jeffrey Brantingham, for having significant roles in my education and research while at UCLA.

It has been a great pleasure to work with many individuals in both the mathematical and criminological community. The research found in Part I is from a collaboration with Matthew Keegan, George Mohler, Todd Wittman, and Andrea Bertozzi. This work is from the article, "Improving density estimation by incorporating spatial information," in EURASIP Journal on Advances in Signal Processing, article ID 265631, volume 2010, doi:10.1155/2010/265631. This article is published through both Hindawi and SpringerOpen. This work was supported by NSF Grant BCS-0527388, NSF Grant DMS-0914856, ARO MURI Grant 50363-MA-MUR, ARO MURI Grant W911NS-09-1-0559, ONR Grant N000140810363, ONR Grant N000141010221, and the Department of Defense. I would also like to thank George Tita, P. Jeffrey Brantingham, Martin Short, and the IPAM RIPS program for data sets used in this research.

Part II of this thesis is comprised of research found in two articles. The work in Chapter 4 is the published article, "Geographical influences of an emerging network of gang rivalries," in Physica A (2011), doi:10.1016/j.physa.2011.05.040. It was done in collaboration with Rachel Hegemann, Alethea Barbaro, Shannon Reid, George Tita, and Andrea Bertozzi. This work was supported by NSF grant DMS-0907931, NSF grant DMS-0968309, AFOSR MURI grant FA9550-10-1-0569, ARO grant 58344-MA, ONR grant N000141010221, and ARO MURI

grant 50363-MA-MUR. The second article used in Part II is to appear in Volume 32 of Discrete and Continuous Dynamical Systems (DCDS-A) in September 2012. Titled, "Adaptation of an ecological territorial model to street gang spatial patterns in Los Angeles," it is divided into Chapters 3 and 5. This work was done in collaboration with Andrea Bertozzi, P. Jeffrey Brantingham, George Tita, and Matthew Valasik. This project was supported by AFOSR MURI grant FA9550-10-1-0569, NSF grants DMS-0968309 and DMS-0907931, ARO grant W911NF1010472, reporting number 58344-MA, ONR grant N000141010221, and ARO-MURI award W911NF-11-1-0332. I would also like to acknowledge Martin Short for his helpful conversations on this project.

Many faculty members of Western Washington University and the Applied Mathematical Sciences Summer Institute were great mentors that helped encourage me to pursue a Ph.D. I would not be here without the support and guidance of Erika Camacho, Angela Gallegos, Lily Khadjavi, Ed Mosteig, Yun-Qiu Shen, Randy Swift, David Uminsky, and Stephen Wirkus.

I would never have been able to complete this degree without the support of God, family, and friends. In particular, I would especially like to thank my wonderful parents, Kevin and Dorothy Smith, for always believing in me and pushing me to do my best. I love them dearly and greatly appreciate everything they have done for me. Monica Walker, my favorite sister, is the best friend anyone could ask for. I'm truly grateful for her support and friendship. My Aunt Margaret Barnard and cousin Janelle Varner have been important parts of my life throughout my graduate programs. They have encouraged me through the difficult times and have been wonderful friends and roommates. I am incredibly grateful to my boyfriend, Benjamin Moreno, for his support, encouragement, love, and most importantly, patience. The friendships that have developed over these years are near and dear to me, and I would especially like to thank Alethea Barbaro, Katherine Foston, Rachel Hegemann, Hayden Schaeffer, and Melissa Tong for being so caring and uplifting, and James von Brecht for his writing assistance and wit during the final year.

# Vita

| | |
|---|---|
| 1984 – 2002 | Traveled the world. |
| 2003–2005 | Mathematics Fellow, University Math Center Tutor, Western Washington University (WWU), Bellingham, WA. |
| 2005 | B.S. (Mathematics), *magna cum laude*, WWU.<br>Outstanding Mathematics Graduate, WWU.<br>Outstanding Honors Graduate, WWU. |
| 2005–2007 | Teaching Assistant, Mathematics Department, WWU.<br>Elias Bond Graduate Fellowship, WWU. |
| 2006–2007 | Research Assistant, Applied Mathematical Sciences Summer Institute (AMSSI), Loyola Marymount University and Cal Poly Pomona. |
| 2007 | M.S. (Mathematics), WWU.<br>Richard Greene Award for Teaching Excellence, WWU.<br>Richard Greene Award for Academic Excellence, WWU. |
| 2007–2011 | National Science Foundation (NSF) Vertical Integration of Research and Education in the Mathematical Sciences (VIGRE) Fellowship, University of California, Los Angeles (UCLA).<br>Teaching Assistant, Mathematics Department, UCLA. |
| 2010 | Research Assistant, Mathematics Department, UCLA.<br>Sorgenfrey Distinguished Teaching Assistant Award, UCLA. |
| 2011 | Graduate Student Instructor, UCLA.<br>Research in Industrial Projects for Students (RIPS) Academic Mentor, IPAM, UCLA. |
| 2011–2012 | Dissertation Year Fellowship, UCLA. |

# Publications

L. Smith, A. Bertozzi, P.J. Brantingham, G. Tita, M. Valasik. "Adaptation of an Ecological Territorial Model to Street Gang Spatial Patterns in Los Angeles." To appear in Discrete and Continuous Dynamical Systems, Volume 32, Number 9, 2012.

R. Hegemann, L. Smith, A. Barbaro, S. Reid, G. Tita, A. Bertozzi. "Geographical Influences of an Emerging Network of Gang Rivalries." Physica A (2011), doi:10.1016/j.physa.2011.05.040.

L. Smith, M. Keegan, T. Wittman, G. Mohler, A. Bertozzi. "Improving Density Estimation by Incorporating Spatial Information." EURASIP J. on Advances in Signal Processing, Article ID 265631, Volume 2010.

Part I

# Using Spatial Information to Improve Density Estimates

# CHAPTER 1

# Introduction to Density Estimation

Suppose $n$ events are observed at the positions $\mathbf{x}_1,\ \mathbf{x}_2,\ \ldots, \mathbf{x}_n\ \in\ \mathbb{R}^d$. These events are assumed to be samples of a probability density function, $f(\mathbf{x})$. A density estimate, $u(\mathbf{x})$, is a function that approximates $f(\mathbf{x})$ from the set of events.

## 1.1 Existing Methods

A common method for creating a probability density estimate is Kernel Density Estimation [117, 118], which approximates the true density by a sum of kernel functions,

$$u(\mathbf{x}) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right),$$

where $h$ is the bandwidth. A larger bandwidth will produce a smoother estimate. A popular choice for the kernel is the Gaussian distribution, also known as the normal distribution. This function is smooth, radially-symmetric, has non-compact support, and is given by

$$k(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}}\ e^{\frac{-|\mathbf{x}|^2}{2}}.$$

Another approach for producing a probability density estimate is through the variational method Maximum Penalized Likelihood Estimation (MPLE) [35, 46, 118]. This appropriately named method takes the estimate that maximizes the log-likelihood function, subject to a roughness penalty. To further explain this approach, we first consider the log-likelihood

for a function $u$ as the underlying probability density for the set of data,

$$\sum_{i=1}^{n} \log(u(\mathbf{x}_i)).$$

If we were to maximize this log-likelihood, then we would obtain a set of delta functions at the locations of the observed points. To remedy this, a term that relates the roughness of the density is included in the maximization through a penalty functional, $P(u)$. With a parameter $\eta$ to indicate the amount of desired smoothing, we now maximize

$$\sum_{i=1}^{n} \log(u(\mathbf{x}_i)) - \eta P(u). \tag{1.1}$$

Requiring that $u$ also be a probability density, we obtain the two constraints,

$$\int_{\Omega} u \; d\mathbf{x} = 1, \quad 0 \leq u(\mathbf{x}). \tag{1.2}$$

Thus, our final density estimate is obtained by finding the argument that maximizes Equation (1.1) subject to the constraints of Equation (1.2), given by

$$u(\mathbf{x}) = \underset{\int_{\Omega} u \; d\mathbf{x}=1, \quad 0 \leq u(\mathbf{x})}{\operatorname{argmax}} \left\{ \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) - \eta P(u) \right\}.$$

We change this to a minimization problem by rewriting it as

$$u(\mathbf{x}) = \underset{\int_{\Omega} u \; d\mathbf{x}=1, \quad 0 \leq u(\mathbf{x})}{\operatorname{argmin}} \left\{ P(u) - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) \right\}.$$

The parameter $\mu$ determines how strongly weighted the maximum likelihood term is compared to the penalty functional, replacing the $\frac{1}{\eta}$ term. A range of penalty functionals have been proposed, including $P(u) = \int_{\Omega} |\nabla \sqrt{u}|^2 \; d\mathbf{x}$ [46, 35] and $P(u) = \int_{\Omega} |\nabla^3 (\log(u))|^2 \; d\mathbf{x}$ [117, 35]. More recently, variants of the Total Variation (TV) functional [105], $P(u) = \int_{\Omega} |\nabla u| \; d\mathbf{x}$, have been proposed for MPLE [65, 107, 78].

There are many other probability density estimation techniques, including the taut string

and logspline methods [28, 66]. The authors of [107] demonstrated the Total Variation MPLE provided better results for nonsmooth densities. In this thesis, we will be producing estimates for nonsmooth densities, and so we will not discuss these models here.

## 1.2 Motivation for Incorporating External Spatial Data

High resolution and hyperspectral satellite images, city and county boundary maps, census data, and other types of geographical data provide much information about a given region. It is desirable to integrate this knowledge into models defining geographically dependent data. Often, it is unreasonable for events to occur in certain regions. For example, residential burglaries and other types of crimes are unlikely to occur in oceans, mountains, and other areas. Such regions can be determined using aerial images or other external spatial data, and we denote these improbable locations as the *invalid region*. Ideally, the support of the density estimate should be contained in the *valid region*.

Section 1.1 introduced existing techniques for density estimation. These methods do not explicitly incorporate the information that can be obtained from the external spatial data, although some note the need to allow for various domains. Consequently, the density estimate typically has some nonzero probability of events occurring in the invalid region. Figure 1.1 demonstrates these problems with the current methods. Located in the middle of the image are two disks where events cannot occur, depicted in Figure 1.1(a). From the region outside the disks, events were sampled from a predefined probability density that is provided in Figure 1.1(b). The 4,000 events chosen are shown in Figure 1.1(c). With a variance of $\sigma = 2.5$, the Gaussian Kernel Density Estimation in Figure 1.1(d) predicts that events may occur in our invalid region.

Figure 1.1(e) displays the density estimate for an MPLE method using the Total Variation penalty functional (TV MPLE). Even though the TV functional can produce sharp gradients, the boundaries of the constant regions do not necessarily agree with the boundaries within the image. This is exemplified in Figure 1.1(e). This method predicts events in the invalid region with non-negligible estimates. This method also performs poorly when the data is

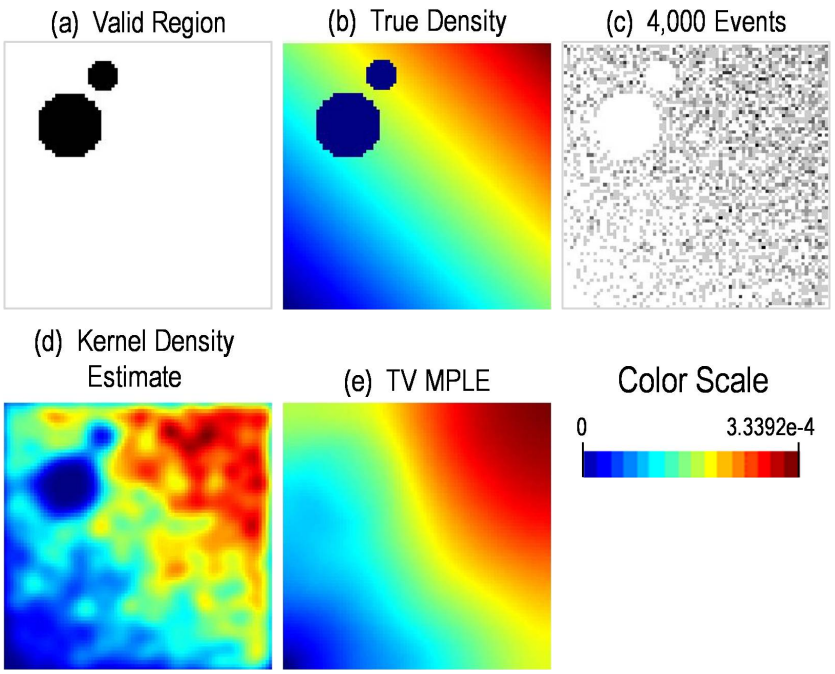


Figure 1.1: This is a motivating example that demonstrates the problem with existing methods. Figures (a) and (b) give the valid region to be considered and the true density for the example. Figure (c) gives the 4000 events sampled from the true density. Figures (d) and (e) show two of the current methods used. The color scale represents the relative probability of an event occurring in a given pixel. The images are 80 pixels by 80 pixels.

too sparse, as the density is smoothed to have equal probability almost everywhere, which will be demonstrated in Section 2.4.3.

### 1.2.1 Overview of Proposed Methods

In Chapter 2 we propose a novel set of non-parametric methods that restrict the support of the density estimate to the valid region and ensure realistic behavior. The methods use the Maximum Penalized Likelihood Estimation framework. The density estimate is calculated as the minimizer of some predefined energy functional. The novelty of our approach is in the way we define the energy functional with explicit dependence on the valid region such that the density estimate obeys our assumptions of its support. The results from our methods for the simple example of Figure 1.1 are illustrated in Figure 1.2.

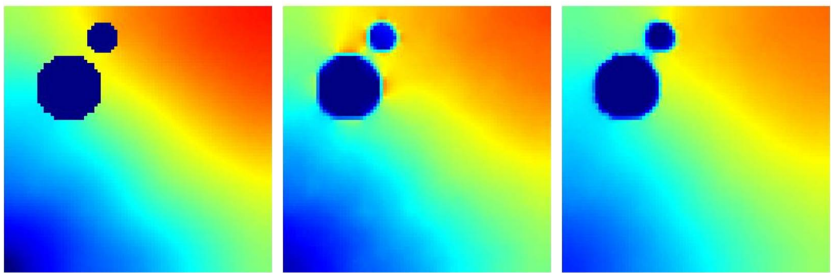

Figure 1.2: This continues the example in Figure 1.1, demonstrating how our methods will improve density estimates. The figures give the estimates for the Modified TV MPLE (left), the Weighted $H_1$ MPLE (center), and the Weighted TV MPLE (right), which will be presented in Chapter 2.

## 1.3 Background on Bregman Iteration and Split Bregman

Solving a maximum penalized likelihood estimation (MPLE) formulation requires minimizing a functional subject to constraints. One method to enforce the equality constraints is to use Bregman Iteration, which we will discuss in Section 1.3.1. Additionally, if we have a total variation functional present in the MPLE formulation, then it is useful to apply Split Bregman [44, 78]. The Split Bregman approach is introduced in Section 1.3.2 for $l_1$ minimization problems and applied to the total variation functional in Section 1.3.3.

### 1.3.1 Bregman Iteration for Constrained Optimization

To explain this approach, we consider a general formulation of the variational problem,

$$\mathbf{v} = \underset{\mathbf{v}}{\operatorname{argmin}} \{H(\mathbf{v})\} \qquad \text{subject to } A\mathbf{v} = \mathbf{b}, \tag{1.3}$$

where $A$ is a linear operator and $H$ is a convex functional. We first change this to an unconstrained optimization problem by adding the constraint in as the penalty $\frac{\lambda}{2}||A\mathbf{v} - \mathbf{b}||_2^2$. This gives the unconstrained problem

$$\mathbf{v} = \underset{\mathbf{v}}{\operatorname{argmin}} \left\{ H(\mathbf{v}) + \frac{\lambda}{2}||A\mathbf{v} - \mathbf{b}||_2^2 \right\}. \tag{1.4}$$

We then implement an iterative method that uses a modified version of this functional to include the "Bregman distance," [21]

$$D_H^{\mathbf{p}}(\mathbf{u}, \mathbf{v}) := H(\mathbf{u}) - H(\mathbf{v}) - \langle \mathbf{p}, \mathbf{u} - \mathbf{v} \rangle .$$

Here, $\mathbf{p} \in \partial H(\mathbf{v})$, meaning that the vector $\mathbf{p}$ is in the set of subdifferentials of $H$ at $\mathbf{v}$. If $H$ is differentiable, then $\mathbf{p} = \nabla H(\mathbf{v})$. Note that this is not a strict distance metric as

$$D_H^{\mathbf{p}}(\mathbf{u}, \mathbf{v}) \neq D_H^{\mathbf{p}}(\mathbf{v}, \mathbf{u})$$

for all $\mathbf{u}$ and $\mathbf{v}$.

To use this Bregman distance, we replace $D_H^{\mathbf{p}^k}(\mathbf{v}, \mathbf{v}^k)$ for $H(\mathbf{v})$ in Equation (1.4), where $\mathbf{v}^k$ is the solution for the $k^{\text{th}}$ iteration. This results in the iterative procedure

$$\begin{aligned} \mathbf{v}^{k+1} &= \underset{\mathbf{v}}{\operatorname{argmin}} \left\{ D_H^{\mathbf{p}^k}(\mathbf{v}, \mathbf{v}^k) + \frac{\lambda}{2} ||A\mathbf{v} - \mathbf{b}||_2^2 \right\} \\ &= \underset{\mathbf{v}}{\operatorname{argmin}} \left\{ H(\mathbf{v}) - H(\mathbf{v}^k) - \left\langle \mathbf{p}^k, \mathbf{v} - \mathbf{v}^k \right\rangle + \frac{\lambda}{2} ||A\mathbf{v} - \mathbf{b}||_2^2 \right\} \\ &= \underset{\mathbf{v}}{\operatorname{argmin}} \left\{ H(\mathbf{v}) - \left\langle \mathbf{p}^k, \mathbf{v} \right\rangle + \frac{\lambda}{2} ||A\mathbf{v} - \mathbf{b}||_2^2 \right\}, \end{aligned} \tag{1.5}$$

where $\mathbf{p}^k \in \partial H(\mathbf{v}^k)$. The optimality conditions indicate that

$$\mathbf{0} = \mathbf{p}^{k+1} - \mathbf{p}^k + \lambda A^T \left( A\mathbf{v}^{k+1} - \mathbf{b} \right).$$

This gives our iterative update formulation,

$$\begin{aligned} \mathbf{v}^{k+1} &= \underset{\mathbf{v}}{\operatorname{argmin}} \left\{ H(\mathbf{v}) - \left\langle \mathbf{p}^k, \mathbf{v} \right\rangle + \frac{\lambda}{2} ||A\mathbf{v} - \mathbf{b}||_2^2 \right\}, \\ \mathbf{p}^{k+1} &= \mathbf{p}^k - \lambda A^T \left( A\mathbf{v}^{k+1} - \mathbf{b} \right). \end{aligned} \tag{1.6}$$

The authors of [144] have shown that this update is equivalent to the following scheme

$$\begin{aligned} \mathbf{v}^{k+1} &= \underset{\mathbf{v}}{\operatorname{argmin}} \left\{ H(\mathbf{v}) + \frac{\lambda}{2} ||A\mathbf{v} - \mathbf{b}^k||_2^2 \right\}, \\ \mathbf{b}^{k+1} &= \mathbf{b}^k - \left( A\mathbf{v}^{k+1} - \mathbf{b} \right). \end{aligned} \tag{1.7}$$

A convergence proof is provided in [44], demonstrating that the iterates $\mathbf{v}^k \rightarrow \mathbf{v}^*$, where $\mathbf{v}^*$ is the solution to the original constrained problem of Equation (1.3). More theoretical results can be found in [44, 97, 144].

### 1.3.2 Split Bregman for $l_1$ Variational Methods

We now consider an unconstrained variational method with an $l_1$ minimization,

$$\mathbf{v} = \underset{\mathbf{v}}{\operatorname{argmin}} \left\{ ||J(\mathbf{v})||_1 + H(\mathbf{v}) \right\}, \tag{1.8}$$

where $H(\mathbf{v})$ is convex, and $J$ is a linear operator. Following the splitting approach of [44, 144], we make a substitution for $J(\mathbf{v})$ and create an equality constraint, giving

$$(\mathbf{v}, \mathbf{d}) = \underset{\mathbf{v},\mathbf{d}}{\operatorname{argmin}} \left\{ ||\mathbf{d}||_1 + H(\mathbf{v}) \right\}, \qquad \text{subject to } J(\mathbf{v}) = \mathbf{d}. \tag{1.9}$$

Now we have a constrained optimization problem similar to that of Equation (1.3). Thus, we can apply Bregman iteration as detailed in Section 1.3.1, using $(\mathbf{d} - J(\mathbf{v}))$ in place of $A\mathbf{v}$ and $\mathbf{0}$ for $\mathbf{b}$. This results in the iterative scheme

$$\begin{aligned} \left(\mathbf{v}^{k+1}, \mathbf{d}^{k+1}\right) &= \underset{\mathbf{v},\mathbf{d}}{\operatorname{argmin}} \left\{ ||\mathbf{d}||_1 + H(\mathbf{v}) + \frac{\lambda}{2} || \left(\mathbf{d} - J(\mathbf{v})\right) - \mathbf{b}^k||_2^2 \right\}, \\ \mathbf{b}^{k+1} &= \mathbf{b}^k - \left( \mathbf{d}^{k+1} - J\left(\mathbf{v}^{k+1}\right) \right). \end{aligned} \tag{1.10}$$

The advantage of using this splitting technique can be seen in the optimization of the $\mathbf{d}$

term. Keeping $\mathbf{u}$ constant, minimizing Equation (1.10) with respect to $\mathbf{d}$ gives

$$\mathbf{d}_i^{k+1} = \text{shrink}\left(J(\mathbf{v}^{k+1})_i + \mathbf{b}_i^k,\ \frac{1}{\lambda}\right), \tag{1.11}$$

where

$$\text{shrink}\,(x,\alpha) := \text{sign}(x)\max\{|x|-\alpha,0\} = \begin{cases} x-a & \text{if } y<\alpha \\ 0 & \text{if } -\alpha \le y \le \alpha\,. \\ x+a & \text{if } y>-\alpha \end{cases} \tag{1.12}$$

The update for $\mathbf{v}^{k+1}$ is given by

$$\mathbf{v}^{k+1} = \underset{\mathbf{v}}{\text{argmin}}\left\{H(\mathbf{v}) + \frac{\lambda}{2}||\left(\mathbf{d}^k - J(\mathbf{v})\right) - \mathbf{b}^k||_2^2\right\}. \tag{1.13}$$

### 1.3.3 Applying Split Bregman to Total Variation

Optimizing a functional that contains a Total Variation term can be quickly implemented with the Split Bregman technique. More specifically, if we are solving the variational problem

$$\mathbf{v} = \underset{\mathbf{v}}{\text{argmin}}\left\{\int_\Omega |\nabla \mathbf{v}| + H(\mathbf{v})\right\},$$

then we have

$$||J(\mathbf{v})||_1 = ||\nabla \mathbf{v}||_1.$$

Using the iterative procedure of Section 1.3.2, we obtain

$$\begin{aligned} \mathbf{v}^{k+1} &= \underset{\mathbf{v}}{\text{argmin}}\left\{H(\mathbf{v}) + \frac{\lambda}{2}||\left(\mathbf{d}^k - \nabla\mathbf{v}\right) - \mathbf{b}^k||_2^2\right\} \\ \mathbf{d}_i^{k+1} &= \text{shrink}\left(\nabla \mathbf{v}_i^{k+1} + \mathbf{b}_i^k,\ \frac{1}{\lambda}\right) \\ \mathbf{b}^{k+1} &= \mathbf{b}^k + \nabla\mathbf{v}^{k+1} - \mathbf{d}^{k+1}. \end{aligned}$$

The Bregman iteration and Split Bregman techniques will be applied to the methods developed in Chapter 2.

## 1.4 Background of Useful Image Processing Methods

Maximum penalized likelihood estimation (MPLE) methods are formulated by optimizing a functional. The log-likelihood term, $-\mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i))$, is a data fidelity term that encourages the estimated density to match the locations of known events. The penalty functional smooths the estimated density so that it does not just reproduce the original data. This approach with data fidelity and penalty functionals is seen in variational image processing techniques. There are many problems in image processing, including registration, segmentation, denoising, deblurring, inpainting, pattern recognition, pan-sharpening, and so on. A few are highlighted here to introduce the penalty functionals that will be incorporated into the proposed MPLE methods of Chapter 2.

### 1.4.1 Mumford-Shah Functional

Segmentation is the process of partitioning the image into regions that correspond to objects or divisions within the scene it represents. We consider an image $f$ that attempts to capture a scene $u$. The image may have some degradation from transmission, blurring, aging, noise, missing portions, or other issues. Mumford and Shah proposed the following energy functional to minimize [86],

$$E_{MS}(u, \Gamma) = \alpha \int \int_R (u - f)^2 \; dx \; dy + \int \int_{R-\Gamma} |\nabla u|^2 \; dx \; dy + \beta |\Gamma|. \tag{1.14}$$

Here, $R$ is the domain of the image and $\Gamma$ is the set of discontinuities. Thus, $\Gamma$ is the edge set that partitions the image into its respective parts. In this minimization problem, the first term, $\alpha \int \int_R (u - f)^2 \; dx \; dy$ is the data fidelity term that encourages $u$ to approximate $f$. The second term, $\int \int_{R-\Gamma} |\nabla u|^2 \; dx \; dy$ is a penalty functional that smooths $u$ away from the edges within the image. Finally, the $\beta|\Gamma|$ term measures the length of the set of curves

that makes up the edges, giving the shortest boundaries that could partition the regions.

The Mumford-Shah model has been modified for many other approaches and problems, including inpainting. Image inpainting is the process of filling in missing or corrupted regions of an image ($R$ – $D$) based on information obtained from outside those regions ($D$). One method that alters the functional of Equation (1.14) is Mumford-Shah-Euler image model of [36]. This new method incorporates Euler's elastica curve, replacing the third term of Equation (1.14), giving

$$\begin{aligned} E_{MSE}(u,\Gamma) &= \alpha \int\int_D (u-f)^2 \; dx \; dy + \int\int_{R-\Gamma} |\nabla u|^2 \; dx \; dy + e(\Gamma) \qquad (1.15) \\ e(\Gamma) &= \int_\Gamma \left(a + b\kappa^2\right) \; ds = a\text{length}(\Gamma) + b\int_\Gamma \kappa^2 ds, \end{aligned}$$

where $\kappa$ is the curvature. Without this curvature term, the inpainted region tends to use straight lines to connect the level sets, creating unnatural edges. With the term, it produces smoother level curves. The only other difference between this model and the Mumford-Shah segmentation model is that $u$ is encouraged to approximate $f$ on $D$ instead of $R$.

### 1.4.2 Total Variation Penalty Functional

While the Mumford-Shah penalty functional uses the $L^2$ norm of $\nabla u$, variational methods using the total variation norm have also been used for many image processing applications with much success. The ROF model is one such model applied to denoising problems [105]. Assuming that $u$ is the true image and $f$ is the noisy observed image, denoising attempts to reconstruct $u$. This particular model is given by taking the minimizing argument of the following functional,

$$E_{ROF}(u) = \int\int_R |\nabla u| \; dx \; dy + \frac{\lambda}{2}\int\int_R (u-f)^2 \; dx \; dy. \qquad (1.16)$$

Using this approach allows for the solution to have discontinuities. Additionally, the total variation penalty discourages the optimal $u$ from containing oscillations.

### 1.4.3 Variational Pan-Sharpening Methods

Pan-sharpening is the process of making a high-resolution multispectral image by fusing a low-resolution multispectral image and a high-resolution panchromatic grayscale image. We let $M_n$ be the $n^{\text{th}}$ band of the low-resolution multispectral image, and we let $P$ be the high-resolution panchromatic image. The $n^{\text{th}}$ band of the optimal solution will be denoted as $u_n$. There are two variational approaches that will be highlighted here. The first is of [9] called P+XS, which minimizes the functional

$$\begin{aligned} E_{p+xs}(u) &= \sum_{n=1}^{N} \gamma_n \int_\Omega \left(|\nabla u_n| + div\theta \cdot \nabla u_n\right)\, dx + \mu \sum_{n=1}^{N} \int_\Omega \Pi_S \left((k_n * u_n) - M_n\right)^2\, dx \\ &\quad + \lambda \int_\Omega \left(\sum_{n=1}^{N} \alpha_n u_n - P\right)^2\, dx. \end{aligned}$$

Without going into too much detail about this functional and all of the variables, we briefly explain each term. Here, $\theta$ is the vector field of the unit normal vectors to the level sets of the high-resolution panchromatic image, $P$. We would like our optimal image to have the same edges as this high-resolution image, and so we want to align the level curves of $u_n$ with the level curves of $P$, i.e.

$$\theta = \frac{\nabla P}{|\nabla P|} = \frac{\nabla u_n}{|\nabla u_n|}.$$

This geometry matching results in first term, $\sum_{n=1}^{N} \gamma_n \int_\Omega \left(|\nabla u_n| + div\theta \cdot \nabla u_n\right)\, dx$.

The second term, $\mu \sum_{n=1}^{N} \int_\Omega \Pi_S \left((k_n * u_n) - M_n\right)^2\, dx$, is included from the fact that the low-resolution multispectral bands are obtained by applying a low-pass filter to high-resolution multispectral bands and then subsampled. The third term,

$$\lambda \int_\Omega \left(\sum_{n=1}^{N} \alpha_n u_n - P\right)^2\, dx,$$

comes from an assumption that the grayscale image equals a linear combination of the spectral bands. This does not appear to be a true assumption, and the next method we discuss proposes an alternate approach.

The second pan-sharpening technique we introduce here comes from [77]. Refered to as Variational Wavelet Pan-Sharpening (VWP), this method minimizes an energy composed of the sum of four functionals,

$$E_{vwp}(u) = E_g + E_w + E_c + E_s.$$

The first of these, $E_g$, is a modified version of the geometry matching term from the P+XS method above, namely

$$E_g = \sum_{n=1}^{N} \left[ \gamma \int_\Omega |\nabla u_n| \; dx + \eta \int_\Omega div\theta \cdot \nabla u_n \; dx \right].$$

The second functional, $E_w$, transforms both the panchromatic image and the upsampled multispectral images ($\uparrow M_n$) into a wavelet decomposition. A new decomposition is created by taking the approximation coefficients from the $\uparrow M_n$ and detail coefficients from $P$. Then the wavelet decomposition of the optimal image $u$ is assumed to approximate this new decomposition, resulting in

$$E_w = \sum_n c_0 \left( a_L^i[n] - \alpha_L^i[n] \right)^2 \phi_{j,n}^2 + \sum_n \sum_{j=1}^{L} \sum_{k=1}^{3} c_j \left( d_{\{k,j\}}[n] - \beta_{\{k,j\}}^i[n] \right)^2 \psi_{j,n}^k.$$

The third term is a color preserving term, matching the upsampled multispectral image away from the edges within the region, denoted as $\Lambda$. This results in the functional

$$E_c = \nu \sum_n \int_{\Omega - \Lambda} \left( u_n - \uparrow M_n \right)^2 \; dx.$$

The final term is designed to create a spectral angle of zero between the low-resolution multispectral image and the optimal high-resolution multispectral image. This is accomplished by letting

$$E_s = \mu \sum_{i,j=1 \;\; i<j}^{N} \int_\Omega \left( u_i \cdot \uparrow M_j - u_j \cdot \uparrow M_i \right)^2 \; dx.$$

Combining all of these functionals together gives an energy that produces a reasonable

high-resolution multispectral image. For more details on each of the terms and the wavelet decomposition, see [77].

The functional $E_g$ will be used in one of the maximum penalized likelihood estimation methods proposed in the next chapter.

## 1.5 Part I Outline

The remainder of Part I is structured in the following way. In Sections 2.1 and 2.2, we present our set of methods, which we name the *Modified* Total Variation MPLE method and the *Weighted* $H_1$ MPLE method, respectively. Here, we also introduce the *Weighted* TV MPLE method to use for comparisons. In Section 2.3 we discuss the implementation and numerical schemes that we use to solve for the solutions of the models. We provide examples for validation of the methods and an example with actual residential burglary data in Section 2.4. In this section, we also compare our results to the Kernel Density Estimation method and the Total Variation MPLE methods. Finally, we give our conclusions in Section 2.5.

# CHAPTER 2

# Incorporating Spatial Information

Including additional information about a given region can improve density estimates. Using the framework of Maximum Penalized Likelihood Estimation (MPLE), we propose a set of models that incorporate a valid region determined from aerial images or other external spatial data.

## 2.1 The Modified Total Variation MPLE Model

The first model we propose is an extension of the Maximum Penalized Likelihood Estimation method given by Mohler, et al. [78],

$$u(\mathbf{x}) = \underset{\int_\Omega u d\mathbf{x}=1,\ \ 0\leq u}{\operatorname{argmin}} \left\{ \int_\Omega |\nabla u| \ d\mathbf{x} - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) \right\},$$

where $\Omega$ is the domain. Once we have determined a valid region, we wish to align the level curves of the density function $u$ with the boundary of the valid region. The Total Variation functional is well known to allow discontinuities in its minimizing solution [105]. By aligning the level curves of the density function with the boundary, we encourage a discontinuity to occur there to keep the density from smoothing into the invalid region.

Since $\frac{\nabla u}{|\nabla u|}$ gives the unit normal vectors to the level curves of $u$, we would like

$$\frac{\nabla(1_D)}{|\nabla(1_D)|} = \frac{\nabla u}{|\nabla u|},$$

where $(1_D)$ is the characteristic function of the valid region $D$. The region $D \subset \Omega$ is obtained from external spatial data, such as aerial images. To avoid division by zero, we use

$\theta := \frac{\nabla(1_D)}{|\nabla(1_D)|_\epsilon}$, where $|\nabla v|_\epsilon = \sqrt{v_x^2 + v_y^2 + \epsilon^2}$. To align the level curves of the density function and the boundary, one would want to minimize $|\nabla u| - \theta \cdot \nabla u$. Integrating this and applying integration by parts, we obtain the term $\int_\Omega |\nabla u| + u\nabla \cdot \theta \; d\mathbf{x}$. We propose the following *Modified* Total Variation penalty functional, where we adopt the more general form of the above functional

$$u(\mathbf{x}) = \underset{\int_\Omega u d\mathbf{x}=1, \quad 0\leq u}{\operatorname{argmin}} \left\{ \int_\Omega |\nabla u| \; d\mathbf{x} + \lambda \int_\Omega u\nabla \cdot \theta \; d\mathbf{x} - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) \right\}.$$

The parameter $\lambda$ allows us to vary the strength of the alignment term. Two pan-sharpening methods, P+XS and Variational Wavelet Pan-sharpening [77, 9], both include a similar term in their energy functional to align the level curves of the optimal image with the level curves of the high resolution pan-chromatic image.

## 2.2 The Weighted $H_1$ MPLE Model

A Maximum Penalized Likelihood Estimation method with penalty functional $\int_\Omega \frac{1}{2}|\nabla u|^2 \; d\mathbf{x}$, the $H_1$ Sobolev semi-norm, gives results equivalent to those obtained using Kernel Density Estimation [35]. We enforce the $H_1$ regularizer term away from the boundary of the invalid region. This results in the model

$$u(\mathbf{x}) = \underset{\int_\Omega u d\mathbf{x}=1, \quad 0\leq u}{\operatorname{argmin}} \left\{ \frac{1}{2} \int_{\Omega\setminus\partial D} |\nabla u|^2 \; d\mathbf{x} - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) \right\}.$$

This new term is essentially the smoothness term from the Mumford-Shah model [86]. We approximate the $H_1$ term by introducing the Ambrosio-Tortorelli approximating function $z_\epsilon(\mathbf{x})$ [4], where $z_\epsilon \to (1 - \delta(\partial D))$ in the sense of distributions. More precisely, we use a continuous function which has the property

$$z_\epsilon(\mathbf{x}) = \begin{cases} 1 & \text{if } d(\mathbf{x}, \partial D) > \epsilon, \\ 0 & \text{if } \mathbf{x} \in \partial D. \end{cases}$$

Thus, the minimization problem becomes

$$u(\mathbf{x}) = \underset{\int_\Omega u d\mathbf{x}=1,\ \ 0\leq u}{\operatorname{argmin}} \left\{ \frac{1}{2}\int_\Omega z_\epsilon^2 |\nabla u|^2 \, d\mathbf{x} - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) \right\}.$$

We will refer to this method as the *Weighted* $H_1$ MPLE model. The weighting away from the edges is used to control the diffusion into the invalid region. This method of weighting away from the edges can also be used with the Total Variation functional in our first model, and we will refer to this as our *Weighted* TV MPLE model.

## 2.3 Implementation

### 2.3.1 The Constraints

In the implementation for the Modified Total Variation MPLE method and Weighted $H_1$ MPLE method, we must enforce the constraints $0 \leq u(\mathbf{x})$ and $\int_\Omega u(\mathbf{x}) \, d\mathbf{x} = 1$ to ensure that $u(\mathbf{x})$ is a probability density estimate. The $u \geq 0$ constraint will be satisfied in our numerical solution by solving quadratic equations that have at least one nonnegative root.

We enforce the second constraint by first adding it to the energy functional as an $L_2$ penalty term. For the Weighted $H_1$ MPLE method, this change results in the new minimization problem

$$u(\mathbf{x}) = \underset{u}{\operatorname{argmin}} \left\{ \frac{1}{2}\int_\Omega z_\epsilon^2 |\nabla u|^2 \, d\mathbf{x} - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) + \frac{\gamma}{2}\left( \int_\Omega u(\mathbf{x}) \, d\mathbf{x} - 1 \right)^2 \right\}.$$

The constraint is then enforced by applying Bregman iteration [97]. Using this method, we formulate our problem as

$$(u,\ b) = \underset{u,\ b}{\operatorname{argmin}} \left\{ \frac{1}{2}\int_\Omega z_\epsilon^2 |\nabla u|^2 \, d\mathbf{x} - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) + \frac{\gamma}{2}\left( \int_\Omega u(\mathbf{x}) \, d\mathbf{x} + b - 1 \right)^2 \right\}$$

where $b$ is introduced as the Bregman variable of the sum to unity constraint. We solve this

problem using alternating minimization, updating the $u$ and the $b$ iterates as

$$(H1) \quad \begin{cases} u^{(k+1)} = & \text{argmin}_u \ \{\frac{1}{2}\int_\Omega z_\epsilon^2 |\nabla u|^2 \ d\mathbf{x} - \mu \sum_{i=1}^n \log(u(\mathbf{x}_i)) \\ & \quad + \frac{\gamma}{2}\left(\int_\Omega u(\mathbf{x}) \ d\mathbf{x} + b^{(k)} - 1\right)^2\} \\ b^{(k+1)} = & b^{(k)} + \int_\Omega u^{(k+1)} \ d\mathbf{x} - 1, \end{cases}$$

with $b^{(0)} = 0$. Similarly for the Modified TV MPLE method, we solve the alternating minimization problem

$$(TV) \quad \begin{cases} u^{(k+1)} & = \text{argmin}_u \left\{ \int_\Omega |\nabla u| \ d\mathbf{x} + \lambda \int_\Omega u \nabla \cdot \theta \ d\mathbf{x} - \mu \sum_{i=1}^n \log(u(\mathbf{x}_i)) \right. \\ & \qquad \left. + \frac{\gamma}{2}\left(\int_\Omega u(\mathbf{x}) \ d\mathbf{x} + b^{(k)} - 1\right)^2 \right\} \\ b^{(k+1)} & = b^{(k)} + \int_\Omega u^{(k+1)} \ d\mathbf{x} - 1 \end{cases}$$

with $b^{(0)} = 0$.

### 2.3.2 Weighted $H_1$ MPLE Implementation

For the Weighted $H_1$ MPLE model, the Euler-Lagrange equation for the $u$ minimization is given by

$$(H1) \qquad -\nabla \cdot \left(z_\epsilon^2 \nabla u\right) - \frac{\mu}{u(\mathbf{x})} \sum_{i=1}^n \delta(\mathbf{x} - \mathbf{x}_i) + \gamma \left(\int_\Omega u(\mathbf{x}) \ d\mathbf{x} + b^{(k)} - 1\right) = 0.$$

We solve this using a Gauss-Seidel method with central differences for the $\nabla z^2$ and $\nabla u$. Once we have discretized the partial differential equation, solving this equation simplifies to solving the quadratic

$$\left(4z^2 + \gamma\right) \ u_{i,j}^2 - \alpha_{i,j} u_{i,j} - \mu w_{i,j} = 0$$

for the nonnegative root, where

$$\begin{aligned}
\alpha_{i,j} = \quad & z_{i,j}^2 \left(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}\right) \\
+ \quad & \left(\frac{z_{i+1,j}^2 - z_{i-1,j}^2}{2}\right)\left(\frac{u_{i+1,j} - u_{i-1,j}}{2}\right) \\
+ \quad & \left(\frac{z_{i,j+1}^2 - z_{i,j-1}^2}{2}\right)\left(\frac{u_{i,j+1} - u_{i,j-1}}{2}\right) \\
+ \quad & \gamma\left(1 - b^{(k)} - \sum_{(i',j')\neq(i,j)} u_{i',j'}\right),
\end{aligned}$$

and where $w_{i,j}$ is the given number of sampled events that occurred at the location $(i, j)$. We chose our parameters $\mu$ and $\gamma$ so that the Gauss-Seidel solver will converge. In particular, we have $\mu = O\left((NM)^{-2}\right)$ and $\gamma = O\left(\mu\left(NM\right)\right)$, where the image is $N \times M$.

### 2.3.3 Modified TV MPLE Implementation

There are many approaches for handling the minimization of the Total Variation penalty functional. A fast and simple method for doing this is to use the Split Bregman technique (see [44],[78] for an in depth discussion, see also [135]). In this approach, we substitute the variable $\mathbf{d}$ for $\nabla u$ in the TV norm and then enforce the equality $\mathbf{d} = \nabla u$ using Bregman iteration. To apply Bregman iteration, we introduce the variable $\mathbf{g}$ as the Bregman vector of the $\mathbf{d} = \nabla u$ constraint. This results in a minimization problem in which we minimize with respect to both $\mathbf{d}$ and $u$.

Beginning the iteration with $\mathbf{g}^{(0)} = 0$, the minimization is written as

$$\begin{aligned}
\left(u^{(k+1)}, \mathbf{d}^{(k+1)}\right) = \quad & \underset{u,\ \mathbf{d}}{\operatorname{argmin}} \Big\{ ||\mathbf{d}||_1 + \lambda \int_\Omega u \nabla \cdot \theta d\mathbf{x} - \mu \sum_{i=1}^{n} \log(u(\mathbf{x}_i)) \\
& + \frac{\gamma}{2}\left(\int_\Omega u(\mathbf{x}) d\mathbf{x} + b^{(k)} - 1\right)^2 \\
& + \frac{\alpha}{2} ||\mathbf{d} - \nabla u - \mathbf{g}^{(k)}||_2^2 \Big\}
\end{aligned}$$

$$\mathbf{g}^{(k)} = \mathbf{g}^{(k-1)} + \nabla u^{(k)} - \mathbf{d}^{(k)}.$$

Alternating the minimization of $u^{(k+1)}$ and $\mathbf{d}^{(k+1)}$, we obtain our final formulation for the TV model as

$$(TV)\begin{cases} u^{(k+1)} & = \text{argmin}_u \left\{ \lambda \int_\Omega u \nabla \cdot \theta \; d\mathbf{x} - \mu \sum_{i=1}^n \log(u(\mathbf{x}_i)) \right. \\ & \qquad + \frac{\gamma}{2} \left( \int_\Omega u(\mathbf{x}) \; d\mathbf{x} + b^{(k)} - 1 \right)^2 \\ & \qquad \left. + \frac{\alpha}{2} ||\mathbf{d}^{(k)} - \nabla u - \mathbf{g}^{(k)}||_2^2 \right\} \\ d_j^{(k+1)} & = \text{shrink}\left( (\nabla u^{(k+1)})_j + g_j^{(k)}, \; \frac{1}{\alpha} \right) \\ \mathbf{g}^{(k+1)} & = \mathbf{g}^{(k)} + \nabla u^{(k+1)} - \mathbf{d}^{(k+1)} \\ b^{(k+1)} & = b^{(k)} + \int_\Omega u^{(k+1)} \; d\mathbf{x} - 1. \end{cases}$$

The shrink function is given by

$$\text{shrink}\,(z, \eta) = \max\{|z| - \eta, \; 0\} \left( \frac{z}{|z|} \right).$$

Solving for $\mathbf{d}^{(k+1)}$ and $\mathbf{g}^{(k+1)}$ we use forward difference discretizations, namely

$$\nabla u^{(k+1)} = (u_{i+1,j} - u_{i,j}, u_{i,j+1} - u_{i,j})^T.$$

The Euler-Lagrange equation for the variable $u^{(k+1)}$ is

$$\begin{aligned} 0 \;\; = \;\; & -\frac{\mu}{u(\mathbf{x})} \sum_{i=1}^n \delta(\mathbf{x} - \mathbf{x}_i) + \lambda \text{div}\theta - \alpha \left( \Delta u + \text{div}\, \mathbf{g}^k - \text{div}\, \mathbf{d}^k \right) \\ & + \gamma \left( \int_\Omega u \; d\mathbf{x} + b^k - 1 \right). \end{aligned}$$

Discretizing this simplifies to solving for the nonnegative root of

$$(4\alpha + \gamma) \; u_{i,j}^2 - \beta_{i,j} u_{i,j} - \mu w_{i,j} = 0$$

where

$$\begin{aligned}\beta_{i,j} = \alpha\left(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}\right) \quad & - \lambda \operatorname{div}\theta \\ - \quad \alpha\left(d^k_{x,i,j} - d^k_{x,i-1,j} + d^k_{y,i,j} - d^k_{y,i,j-1}\right) & \\ + \quad \alpha\left(g^k_{x,i,j} - g^k_{x,i-1,j} + g^k_{y,i,j} - g^k_{y,i,j-1}\right) & \\ + \quad \gamma\left(1 - b^{(k)} - \sum_{(i',j')\neq(i,j)} u_{i',j'}\right). &\end{aligned}$$

We solved for $u^{(k+1)}$ with a Gauss-Seidel solver. Heuristically, we found that using the relationships $\alpha = 2\mu N^2 M^2$ and $\gamma = 2\mu NM$ were sufficient for the solver to converge and provide good results. We also set $\lambda$ to have values between 1.0 and 1.2. The parameter $\mu$ is the last remaining free parameter. This parameter can be chosen using V-cross validation or other techniques, such as the sparsity $l_1$ information criterion [107].

## 2.4 Results

In this section, we demonstrate the strengths of our models by providing several examples. We first revisit the example from Chapter 1, providing an error comparison for both our methods and the existing methods. Next, we consider a dense data set, sampled from a piecewise-constant density. We then show our methods perform well for sparse data sets. Next, we explore an example with an aerial image and randomly selected events to show how these methods could be applied to geographic event data. Finally, we calculate probability density estimates for residential burglaries.

### 2.4.1 Continuation of Introductory Example

Chapter 1 provided a motivating example, showing existing methods in Figure 1.1 and our proposed methods in Figure 1.2. Visually, the methods we propose outperform the existing methods by preventing density from smoothing into the invalid region and by producing qualitatively similar densities in the valid region. To better understand how the methods

perform, we examine the $L_2$ error,

$$\left( \sum_{ij} |u_{ij} - u^*_{ij}|^2 \right)^{1/2},$$

and the relative $L_2$ error,

$$\frac{\left( \sum_{ij} |u_{ij} - u^*_{ij}|^2 \right)^{1/2}}{\left( \sum_{ij} |u^*_{ij}|^2 \right)^{1/2}},$$

where $u^*$ is the exact solution. The results for the introductory example are provided in Table 2.1.

| **4,000 Events** | $L_2$ Error | Relative $L_2$ Error |
|---|---|---|
| **Kernel Density Estimate** | 2.7191e-3 | 0.19393 |
| **TV MPLE** | 2.7862e-3 | 0.19871 |
| **Modified TV MPLE** | **7.6155e-4** | **0.05431** |
| **Weighted $H_1$ MPLE** | 1.4773e-3 | 0.10536 |
| **Weighted TV MPLE** | 1.9021e-3 | 0.13566 |

Table 2.1: These are the $L_2$ error and relative error comparisons of the five methods for the introductory example shown in Figure 1.1. Our proposed methods performed better than both the Kernel Density Estimation method and the TV MPLE method.

The error values demonstrate that all our proposed methods perform better than the existing Kernel Density Estimation and TV MPLE methods. Additionally, the Modified TV MPLE method performs the best for this example.

### 2.4.2 Dense Data Set Example

To further validate the use of our methods, we took a predefined probability map with sharp gradients, see Figure 2.1(a). The chosen valid region and the 8,000 selected events are displayed in Figures 2.1(b) and 2.1(c), respectively. We performed density estimates with the Gaussian Kernel Density Estimation and the Total Variation MPLE methods. The variance used for the Kernel Density Estimation is $\sigma = 2$. The results are provided in Figures 2.1(d) and 2.1(e). The density estimates obtained from our Modified TV MPLE method and

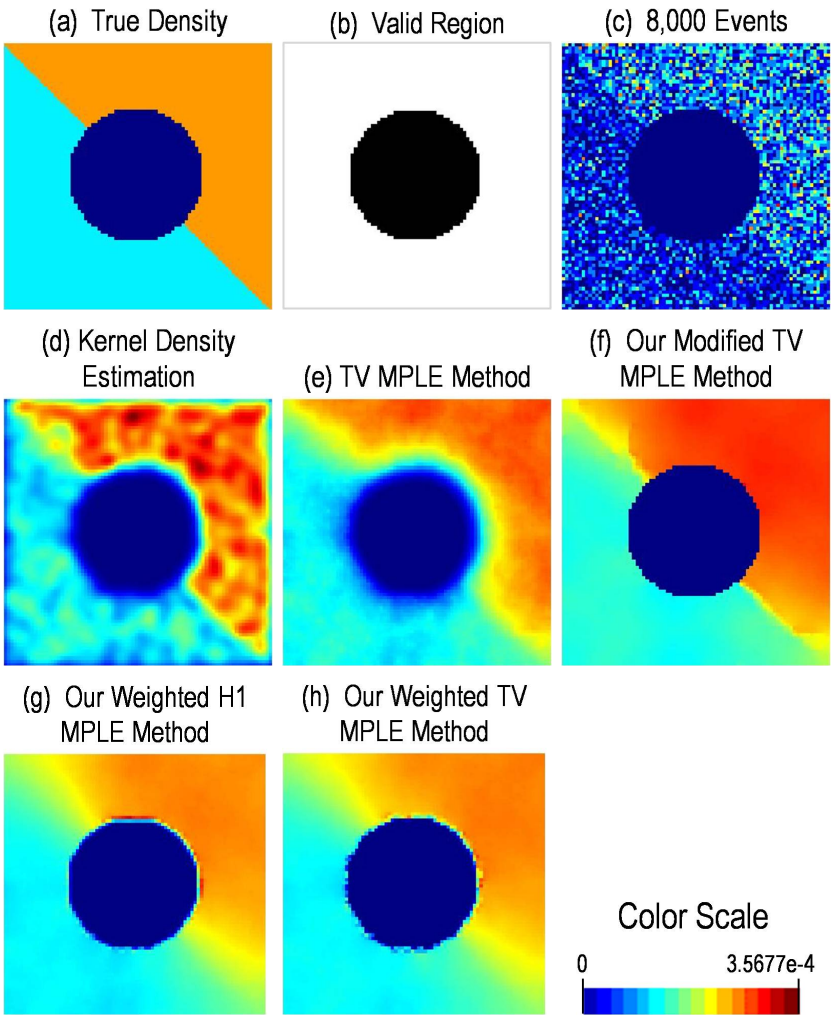


Figure 2.1: This is a model validating example with dense data set of 8000 events. The piecewise-constant true density is given in (a), and the valid region is provided in (b). The sampled events are shown in (c). Figures (d) and (e) show the two current density estimation methods, Kernel Density Estimation and TV MPLE. Figures (f), (g), and (h) show the density estimates from our methods. The color scale represents the relative probability of an event occurring in a given pixel. The images are 80 pixels by 80 pixels.

Weighted $H_1$ MPLE method are shown in Figures 2.1(f) and 2.1(g), respectively. We also included our Weighted TV MPLE in Figure 2.1(h).

Our methods maintain the boundary of the invalid region and appear close to the true solution. In addition, they keep the sharp gradient in the density estimate. The $L_2$ errors and relative $L_2$ errors for these methods are located in Table 2.2.

| **8,000 Events** | $L_2$ Error | Relative $L_2$ Error |
|---|---|---|
| **Kernel Density Estimate** | 2.8474e-3 | 0.1940 |
| **TV MPLE** | 2.5721e-3 | 0.1753 |
| **Modified TV MPLE** | 2.0301e-3 | 0.1383 |
| **Weighted $H_1$ MPLE** | **1.9691e-3** | **0.1342** |
| **Weighted TV MPLE** | 2.0783e-3 | 0.1416 |

Table 2.2: These are the $L_2$ error and relative error comparisons of the five methods shown in Figure 2.1. Our proposed methods performed better than both the Kernel Density Estimation method and the TV MPLE method.

### 2.4.3 Sparse Data Example

Crimes and other types of events may be quite sparse in a given geographical region. Consequently, it becomes difficult to determine the probability that an event will occur in the area. It is challenging for density estimation methods that do not incorporate the spatial information to distinguish between invalid regions and areas that have not had any crimes but are still likely to have events. Using the same predefined probability density from the introduction section in Figure 1.1(b), we demonstrate how our methods maintain these invalid regions for sparse data. The 40 events selected are shown in Figure 2.2(b). The density estimates for current methods and our methods are given in Figure 2.2. We used a variance $\sigma = 15$ for the Gaussian Kernel Density Estimate.

| **40 Events** | $L_2$ Error | Relative $L_2$ Error |
|---|---|---|
| **Kernel Density Estimate** | 4.8021e-3 | 0.3425 |
| **TV MPLE** | 5.0346e-3 | 0.3591 |
| **Modified TV MPLE** | 3.7875e-3 | 0.2701 |
| **Weighted $H_1$ MPLE** | **1.9608e-3** | **0.1398** |
| **Weighted TV MPLE** | 3.9983e-3 | 0.2852 |

Table 2.3: These are the $L_2$ error and relative error comparisons of the five methods for the sparse example shown in Figure 2.2. Our proposed methods performed better than both the Kernel Density Estimation method and the TV MPLE method.

For this sparse problem, our Weighted $H_1$ MPLE and Modified TV MPLE methods maintain the boundary of the invalid region and appear close to the true solution. Table 2.3 contains the $L_2$ errors and relative $L_2$ errors for this example. Notice that the Weighted $H_1$ MPLE was exceptionally better for the sparse data set. The Weighted TV MPLE method

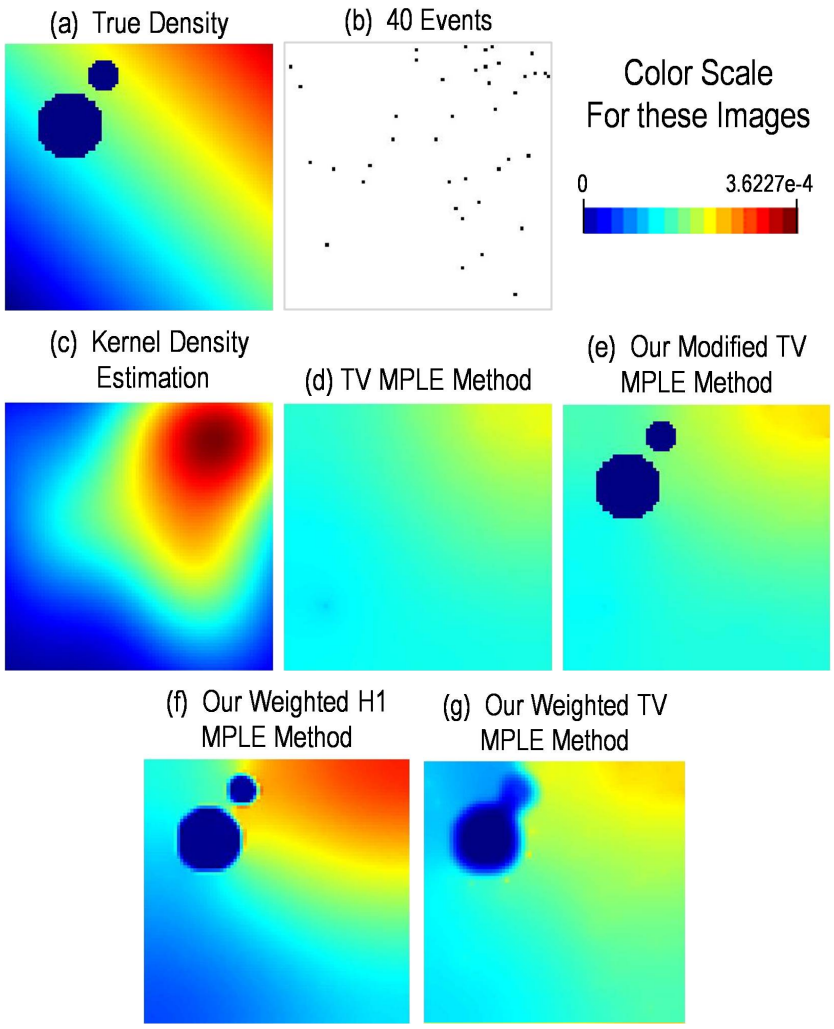


Figure 2.2: This is a sparse example with 40 events. The true density is given in (a), and it is the same density from the example in the introduction. The sampled events are shown in (b). Figures (c) and (d) show the two current density estimation methods, Kernel Density Estimation and TV MPLE. Figures (e), (f), and (g) show the density estimates from our methods. The color scale represents the relative probability of an event occurring in a given pixel. The images are 80 pixels by 80 pixels.

does not perform as well for sparse data sets and fails to keep the boundary of the valid region. Since the rest of the examples contain sparse data sets, we will omit the Weighted TV MPLE method from the remaining sections.

### 2.4.4 Orange County Coastline Example

To test the models with external spatial data, we obtained from Google Earth $^{TM}$ a region of the Orange County coastline with clear invalid regions (see Figure 2.3(a)). For the purposes of this example, it was determined to be impossible for events to occur in the ocean, rivers, or large parks located in the middle of the region. One may use various segmentation methods for selecting the valid region. For this example, we only have data from the true color aerial image, not multispectral data. To obtain the valid and invalid regions, we removed the "texture" (i.e. fine detailed features) using a Total Variation based denoising algorithm [105]. The resulting image, shown in Figure 2.3(b), still contains detailed regions obtained from large features, such as large buildings. We wish to remove these and maintain prominent regional boundaries. Therefore, we smooth away from regions of large discontinuities. This is shown in Figure 2.3(c). Since oceans, rivers, parks, and other such areas have generally lower intensity values than other regions, we threshold to find the boundary between the valid and invalid regions. The final valid region is displayed in Figure 2.4(a).

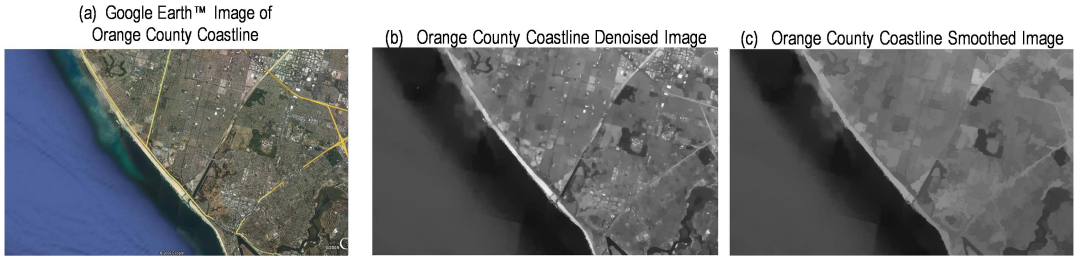


Figure 2.3: This shows how we obtained our valid region for the Orange County Coastline example. Figure (a) is the initial aerial image of the region to be considered. The region of interest is about 15.2 km by 10 km. Figure (b) is the denoised version of the initial image. We took this denoised image and smoothed away from regions of large discontinuities to obtain Figure (c).

From the valid region, we constructed a toy density map to represent the probability density for the example and to generate data. It is depicted in Figure 2.4(b). Regions with colors farther to the right on the color scale are more likely to have events. Sampling from this constructed density, we took distinct data sets of 200, 2,000 and 20,000 selected events given in Figure 2.5. For each set of events, we included three probability density estimations for comparison. We first give the Gaussian Kernel Density Estimate followed by our Modified

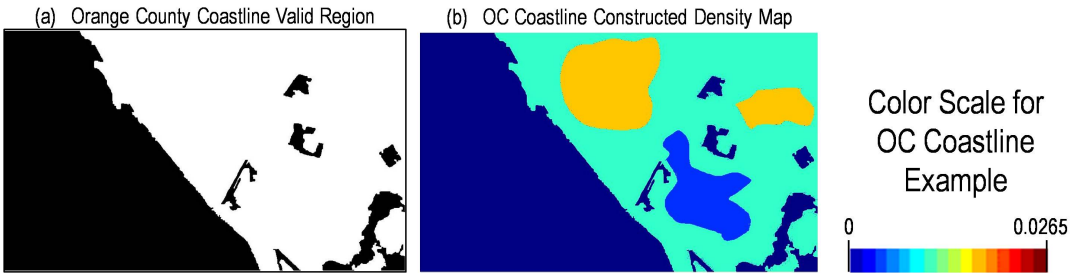


Figure 2.4: After thresholding the intensity values of Figure 2.3(c), we obtain the valid region for the Orange County Coastline. This valid region is shown in (a). We then constructed a probability density shown in Figure (b). The color scale represents the relative probability of an event occurring per square kilometer.

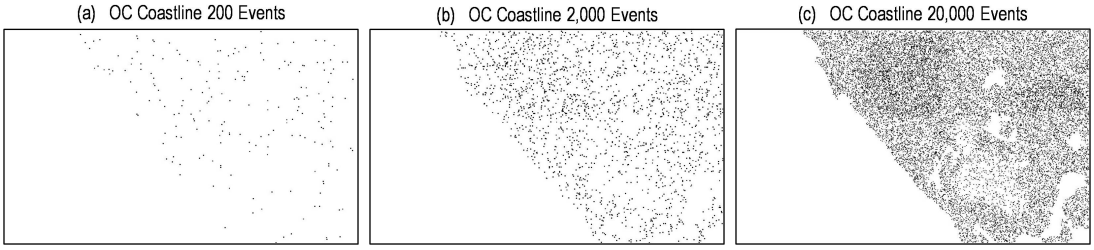


Figure 2.5: From the probability density in Figure 2.4, we sampled 200, 2,000, and 20,000 events. These events are given in (a), (b), and (c), respectively.

Total Variation MPLE model and our Weighted $H_1$ MPLE model. We provide all images together to allow for visual comparisons of the methods.

Summing up Gaussian distributions gives a smooth density estimate. Figure 2.6 contains the density estimates obtained using the Kernel Density Estimation model. The standard deviation $\sigma$ of the Gaussians are given with each image. In all of these images, a nonzero density is estimated in the invalid region.

Taking the same set of events as the Kernel Density Estimation, the images in Figure 2.7 were obtained using our first model, the Modified Total Variation MPLE method with the boundary edge aligning term. The parameter for $\lambda$ must be sufficiently large in the TV method in order to prevent the diffusion of the density into the invalid region. In doing so, the boundary of the valid region may attain density values too large in comparison to the rest of the image when the size of the image is very large. To remedy this, we may take the resulting image from the algorithm and set the boundary of the valid region to zero and

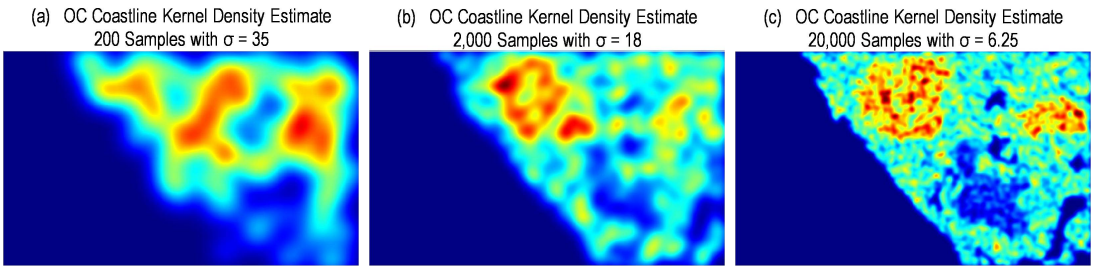


Figure 2.6: These images are the Gaussian Kernel Density estimates for 200, 2,000, and 20,000 sampled events of the Orange County Coastline example. The color scale for these images are located in Figure 2.4.

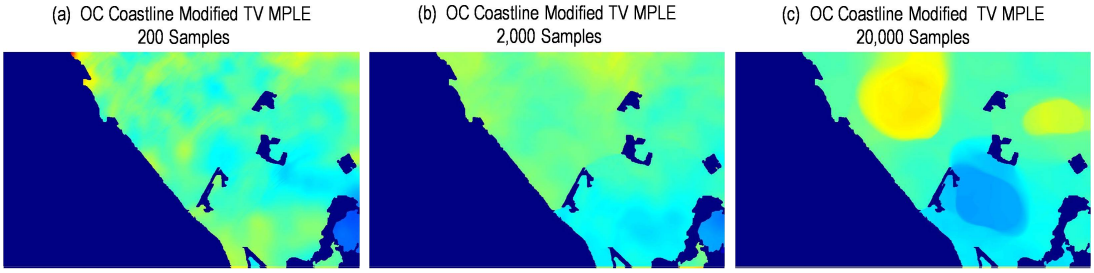


Figure 2.7: These images are the Modified TV MPLE estimates for 200, 2,000, and 20,000 sampled events of the Orange County Coastline example. The color scale for these images are located in Figure 2.4.

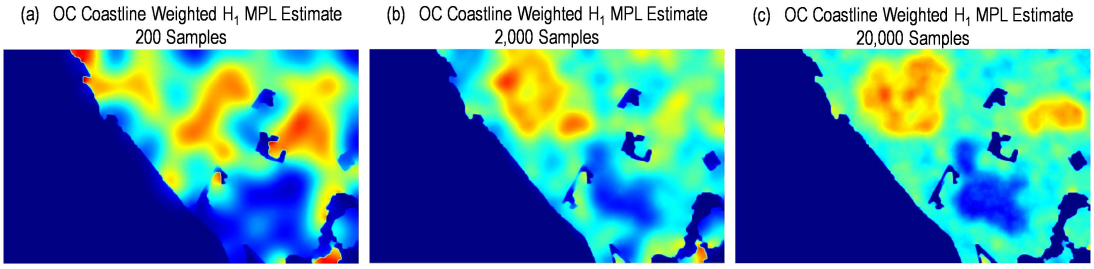


Figure 2.8: These images are the Weighted $H_1$ MPLE estimates for 200, 2,000, and 20,000 sampled events of the Orange County Coastline example. The color scale for these images are located in Figure 2.4.

rescale the image to have a sum of one. The invalid region in this case sometimes has a very small non-zero estimate. For visualization purposes we have set this to zero. However, we note that the method has the strength that density does not diffuse through small sections of the invalid region back into the valid region on the opposite side. Events on one side of an object, such as a lake or river, should not necessarily predict events on the other side.

The next set of images in Figure 2.8 estimate the density using the same sets of event data but with our Weighted $H_1$ MPLE model. Notice the difference for the invalid regions with our models and the Kernel Density Estimation model. This method does very well for

the sparse data sets of 200 and 2,000 events.

#### 2.4.4.1 Model Comparisons

The density estimates obtained from using our methods have a clear improvement in maintaining the boundary of the valid region. To determine how our models did in comparison to one another and to the Kernel Density Estimate, we calculated the $L_2$ errors and relative errors located in Tables 2.4, 2.5, and 2.6.

| **200 Events** | $L_2$ Error | Relative $L_2$ Error |
|---|---|---|
| **Kernel Density Estimate** | 8.3868e-4 | 0.4646 |
| **Modified TV MPLE** | **5.5494e-4** | **0.3074** |
| **Weighted $H_1$ MPLE** | 7.3931e-4 | 0.4095 |

Table 2.4: These are the $L_2$ error and relative error comparisons of the three methods for the Orange County Coastline example with 200 data points shown in Figures 2.6, 2.7, and 2.8. Our proposed methods performed better than the Kernel Density Estimation method.

| **2,000 Events** | $L_2$ Error | Relative $L_2$ Error |
|---|---|---|
| **Kernel Density Estimate** | 5.3709e-4 | 0.2975 |
| **Modified TV MPLE** | 5.1569e-4 | 0.2857 |
| **Weighted $H_1$ MPLE** | **3.9985e-4** | **0.2215** |

Table 2.5: These are the $L_2$ error and relative error comparisons of the three methods for the Orange County Coastline example with 2,000 data points shown in Figures 2.6, 2.7, and 2.8. Our proposed methods performed better than the Kernel Density Estimation method.

| **20,000 Events** | $L_2$ Error | Relative $L_2$ Error |
|---|---|---|
| **Kernel Density Estimate** | 3.9781e-4 | 0.2204 |
| **Modified TV MPLE** | 2.9892e-4 | 0.1656 |
| **Weighted $H_1$ MPLE** | **2.4091e-4** | **0.1335** |

Table 2.6: These are the $L_2$ error and relative error comparisons of the three methods for the Orange County Coastline example with 20,000 data points shown in Figures 2.6, 2.7, and 2.8. Our proposed methods performed better than the Kernel Density Estimation method.

Our models consistently outperform the Kernel Density Estimation model. The Weighted $H_1$ MPLE method performs the best for the 2,000 and 20,000 events and visually appears closer to the true solution for the 200 events than the other methods. Qualitatively, we have noticed that with sparse data, the TV penalty functional gives results which are near

constant. Thus, it gives a good $L_2$ error for the Orange County Coastline example, which has piecewise-constant true density, but gives a worse result for the sparse data example of Figure 2.2, where the true density has a nonzero gradient. Even though the Modified TV MPLE method has a lower $L_2$ error in the Orange County Coastline example, the density estimation fails to give a good indication of regions of high and low likelihood.

### 2.4.5 Residential Burglary Example

The following example uses actual residential burglary information from the San Fernando Valley in Los Angeles. In Figure 2.9 is the area of interest and the locations of 4,487 burglaries that occurred in the region during 2004 and 2005. The aerial image was obtained using Google Earth$^{TM}$. We assume that residential burglaries cannot occur in large parks, lakes, mountainous areas without houses, at airports, and industrial areas. Using census or other types of data, housing density information for a given region can be calculated. Figure 2.9(c) is the housing density for our region of interest. The housing density provides us with the exact locations of where residential burglaries may occur. However, our methods prohibit the density estimates from spreading through the boundaries of the valid region. If we were to use this image directly as the valid region, then crimes on one side of a street will not have an effect on the opposite side of the road. Therefore, we fill in small holes and streets in the housing density image and use the image located in Figure 2.9(d) as our valid region.

Using our Weighted $H_1$ MPLE and Modified TV MPLE models, the Gaussian Kernel Density Estimate with variance $\sigma = 21$, and the TV MPLE method, we obtained the density estimations shown in Figure 2.10.

## 2.5 Conclusions

In this paper we have studied the problem of determining a more geographically accurate probability density estimate. We demonstrate the importance of this problem by showing how common density estimation techniques, such as Kernel Density Estimation, fail to

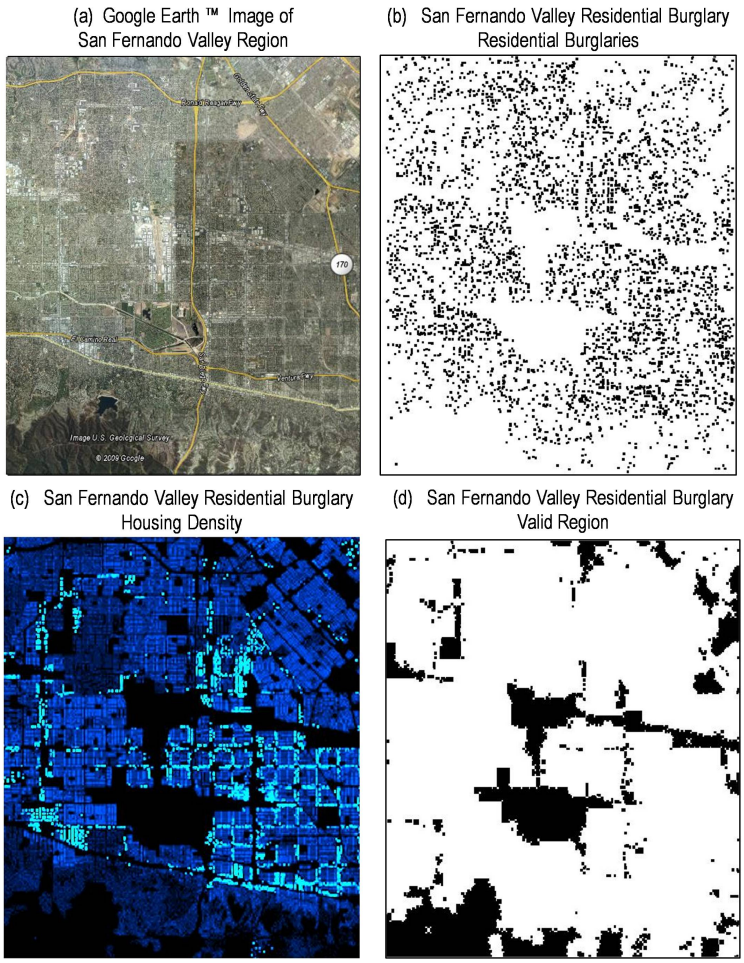


Figure 2.9: These Figures are for the San Fernando Valley residential burglary data. In (a), we have the aerial image of the region we are considering, which is about 16 km by 18 km. Figure (b) shows the residential burglaries of the region. Figure (c) gives the housing density for the San Fernando Valley. We show the valid region we obtained from the housing density in Figure (d).

restrict the support of the density in a set of realistic examples.

To handle this problem, we proposed a set of methods, based on Total Variation and $H_1$-regularized MPLE models, that demonstrates great improvements in accurately enforcing the support of the density estimate when the valid region has been provided *a priori*. Unlike the TV-regularized methods, our $H_1$ model has the advantage that it performs well for very sparse data sets.

The effectiveness of the methods is shown in a set of examples in which burglary probability densities are approximated from a set of crime events. Regions in which burglaries

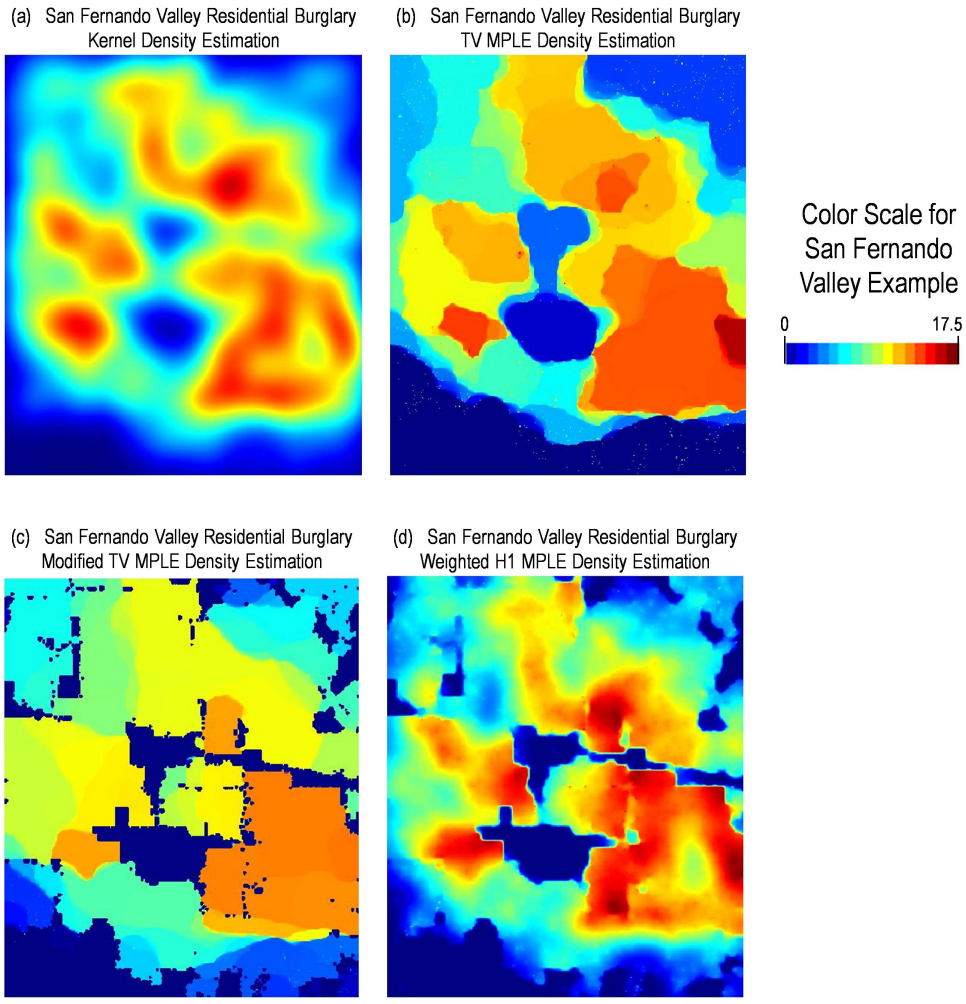


Figure 2.10: These images are the density estimates for the San Fernando Valley residential burglary data. Figures (a) and (b) show the results of the current methods Kernel Density Estimation and TV MPLE, respectively. The results from our Modified TV MPLE method and our Weighted $H_1$ MPLE method are shown in Figures (c) and (d), respectively. The color scale represents the number of residential burglaries per year per square kilometer.

are impossible, such as oceans, mountains, and parks, are determined using aerial images or other external spatial data. These regions are then used to define an invalid region in which the density should be zero. Therefore, our methods are used to build geographically accurate probability maps.

Part II

# Modeling Gang Rivalries and Spatial Behaviors

# CHAPTER 3

# Introduction to Modeling Street Gangs

Street gangs have a long history in Los Angeles [83, 128] and are major contributors to violent acts in the region [41, 82, 85, 102, 134]. Gangs compete to secure both instrumental (e.g., money, drugs, space) and expressive (e.g., reputation, power) resources, often causing a partitioning of the city into gang territories [123]. Disputes over spatial resources may lead to violence and acts of retribution, augmenting or creating a rivalry between gangs [29, 125].

With the prevalence of street gangs in some regions, police departments have developed strategies for mitigating gang violence. These tactics may include gang injunctions, enforcing gang territory boundaries, or interventions [128]. While these methods are important, they may be costly or difficult to implement. Before making large scale changes in policing strategies, it is beneficial to test theories through inexpensive models.

## 3.1 Existing Gang Models

Several models have recently been proposed for analyzing various gang behaviors. Retaliatory behavior among gangs has been simulated via self-exciting point processes and an agent-based method [34, 116]. These methods focus on the temporal aspect of gang violence. Using this gang characteristic of retaliation, attempts have been made to fill in the missing data of which gang is responsible for a crime [51, 122]. From a different perspective, the authors of [7] use an approach from epidemiology to simulate membership in gangs.

The Lotka-Volterra competition model has been used to analyze the creation of territorial boundaries between two gangs in [20]. Through this approach, the authors examined the locations of violent interactions with respect to the theoretical boundary between the two

gangs and concluded that violence typically clusters about this dividing curve. Another model uses graffiti as the mechanism for interactions between gang members and analyzes the circumstances under which territories are formed [11]. Both of these methods focus only on situations with two gangs. However, many areas with gang violence contain more than two gangs. Thus, we will need to consider another approach to fully examine a larger system of active gangs.

## 3.2 Motivation for Proposed Models

The goal of the proposed models is to simulate the spatial behaviors of gangs to assist in the understanding of violent interactions and the formation of rivalries and territories. If achieved, these models could provide information of how the system may change given different scenarios and intervention strategies. We aim to model gangs' spatial dynamics through one agent-based model and one approach with a system of partial differential equations. Each method outputs different types of information for analysis.

### 3.2.1 The Agent-Based Approach

One approach to modeling complex systems is through the class of agent-based models (ABM). These types of models create a number of individuals or groups, labeled *agents*, and often have them interact, learn from the environment, and make decisions through computational simulations. These methods use a "bottom-up approach," attempting to simulate behavioral dynamics for the system from the actions of the individual agents. This allows for examination of the emergent patterns based on simple rules. These models are flexible and allow complex dynamics, rules, and strategies to be incorporated into the system.

With a stochastic approach, a model can capture some randomness that is found in human activity for which the underlying reasoning is unknown. We examine the formation of rivalries through a stochastic agent-based model that is coupled to a rivalry network in Chapter 4. A node of the network represents a geographic position that is the central location to a gang's activities known as a *set space* [102, 126]. An edge in the network is present

if there is a rivalry between the two gangs. In this model, agents move about the city and interact with members of other gangs. Additionally, geographic boundaries in the region are incorporated into this model to restrict movement across rivers and large freeways.

### 3.2.2 The PDE Approach

Using a deterministic approach with a derivation based on a stochastic process, we can examine other information such as gang densities and territories. Gang territories are dynamic and the boundaries change over time with the emergence of new gangs and collapsing of other gangs. By creating a flexible model that can incorporate these fluctuations, theories can be tested to investigate the system dynamics given certain changes [32, 47].

Chapter 5 proposes a deterministic model of a system of partial differential equations to describe the equilibrium densities of gangs and gang graffiti. Gangs identify their territories through the use of graffiti [3, 25, 69]. Gang graffiti marks the boundaries of contested regions claimed by multiple gangs, which may make these areas more prone to violence. Our goal is to model the spatial behaviors for gangs and their taggings.

One approach for modeling human behavior is through the framework of ecological models. The ecological models for optimal foraging have been applied to hunters and gatherers [50] and to people seeking information [106]. College drinking patterns have also been modeled through an ecological approach [1]. Ecology has been essential to epidemiological modeling of wildlife populations and diseases [61]. These models have in turn been applied to humans for not only infectious diseases, but also to the social sciences through the spread of ideas [15] and political affiliations [104]. Adapting ecological models created for territorial animals, we model the spatial behaviors of street gangs, incorporating the geographic features of the region that influence movement.

#### 3.2.2.1 Existing Ecological Territorial Models

Many species exhibit territorial behavior, where individuals and/or groups are willing to defend space they currently hold from incursions by members of the same species (intraspecific

competition) or other species (interspecific competition). These territorial behaviors are also exemplified by street gangs which use violence to control local drug markets or simply defend their neighborhood from outsiders. We consider existing ecological territorial methods to model these gang territorial behaviors. Several models have recently been proposed to describe the territorial behavior of species as diverse as wolves [140, 23, 67, 68, 139], coyotes [80, 81], killer wasps [31], chimpanzees [137], foxes [43], and birds [58].

The paper [55] provides a review of partial differential equation models proposed for studying the spatial distributions of species. Some approaches include pure diffusion models with Brownian random motion [94, 33, 87],

$$\frac{\partial u(x,y,t)}{\partial t} = D\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right),$$

diffusion with drift and convection models [53, 10],

$$\frac{\partial u}{\partial t} = D\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) - w_x\,\frac{\partial u}{\partial x} - w_y\,\frac{\partial u}{\partial y},$$

biased random motion models [49],

$$\frac{\partial u}{\partial t} = D\,\frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x}\left(ku\frac{\partial u}{\partial x}\right),$$

models with aggregation or repulsion that depends on the current organism density [6, 133],

$$\frac{\partial u}{\partial t} = D\,\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 \psi(u)}{\partial x^2}$$

and models that depend on environmental features [112, 111, 95, 22],

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[u\frac{d}{dx}\phi(E)\right].$$

Additionally, reaction-diffusion,

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[D(u)\frac{\partial u}{\partial x}\right] + \frac{\partial}{\partial y}\left[D(u)\frac{\partial u}{\partial y}\right] + f(u),$$

and predator-prey models incorporate competition into the system.

Two competing groups might never physically meet, yet still have defined territories. Coyotes and wolves are animals that exhibit territorial behavior by claiming spatial regions via scent marking from raised leg urination or other means [99]. When an individual comes in contact with a marking from a different pack, the animal increases its own urination and moves towards its home den located on the interior of its territory [67]. This avoidance of markings has been incorporated into wolf [140, 23, 67, 68, 139] and coyote models [80, 81] through a coupled system of partial differential equations for both the pack densities and the marking densities. These mechanistic home range models have recently been connected to models of resource selection analysis (RSA) [79].

One of the models of [81] includes a mechanism to avoid areas where the terrain is steep. It is this model that we will modify to examine the spatial behaviors of street gangs, using gang graffiti as a proxy for scent marking. Figure 3.1 displays the results from this paper, indicating the ability of this model to approximate the coyote densities in Yellowstone. Here, the colored curves give the contours of the coyote densities produced by the "steep terrain avoidance plus conspecific avoidance (STA+CA)" model. The colored dots depict the locations where coyotes have been found and documented. The different colors represent different packs of coyotes. The black curves show the elevation contours of the region, demonstrating how the steep changes in the terrain influence the locations where coyotes are found. The STA+CA model will be presented in Chapter 5, along with its derivation.

## 3.3 Case Study: Street Gangs in Hollenbeck

The city of Los Angeles is plagued by gang violence. Hollenbeck, the eastern most division of the Los Angeles Police Department, is home to twenty-nine active Latino street gangs in

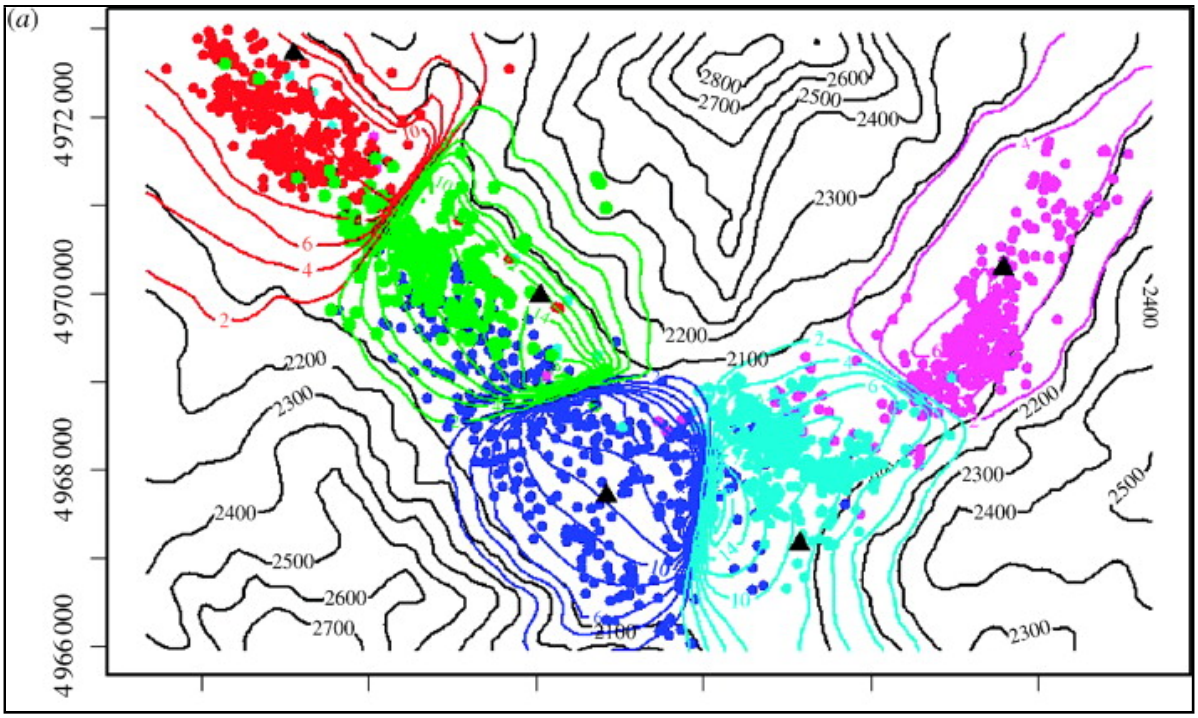


Figure 3.1: This image was taken from P. R. Moorcroft, M. A. Lewis, and R. L. Crabtree, "Mechanistic home range models capture spatial patterns and dynamics of coyote territories in yellowstone," Proceedings: Biological Sciences, 2006, 273(1594):1651–1659, by permission of the Royal Society [81]. Here, the colored curves give the contours of the coyote densities produced by the "steep terrain avoidance plus conspecific avoidance (STA+CA)" model. The colored dots depict the locations where coyotes have been found and documented. The different colors represent different packs of coyotes. The black curves show the elevation contours of the region.

only 15.2 square miles [102, 128]. Due to the geography of the area, the rivalries among the gangs in this region are generally restricted to Hollenbeck. Most of the rivalries and disputes among these gangs is linked to neighborhood-based territoriality rather than drugs or other conflicts.

To examine the structure of the rivalries in Hollenbeck, we view a graph that is embedded in space with nodes representing the set spaces of the gangs. If a rivalry exists between two gangs, then an edge is present between the two respective nodes. Among the twenty-nine gangs in Hollenbeck, there are sixty-nine observed rivalries. This observed network has been examined in [102, 128]. The authors of these papers argue that the structure of this network is highly dependent on the geography of the region and the highways and rivers that pass through the area. Five major freeways, I-5, I-10, I-710, CA-60, and US-101, divide

Hollenbeck into many sections. Figure 3.2 shows the spatially embedded rivalry network and the features that partition the region. It is clear that these physical boundaries impact the rivalry structure by limiting the number of rivalries between gangs in separate sections. Thus, it is desirable to include these features into our models.

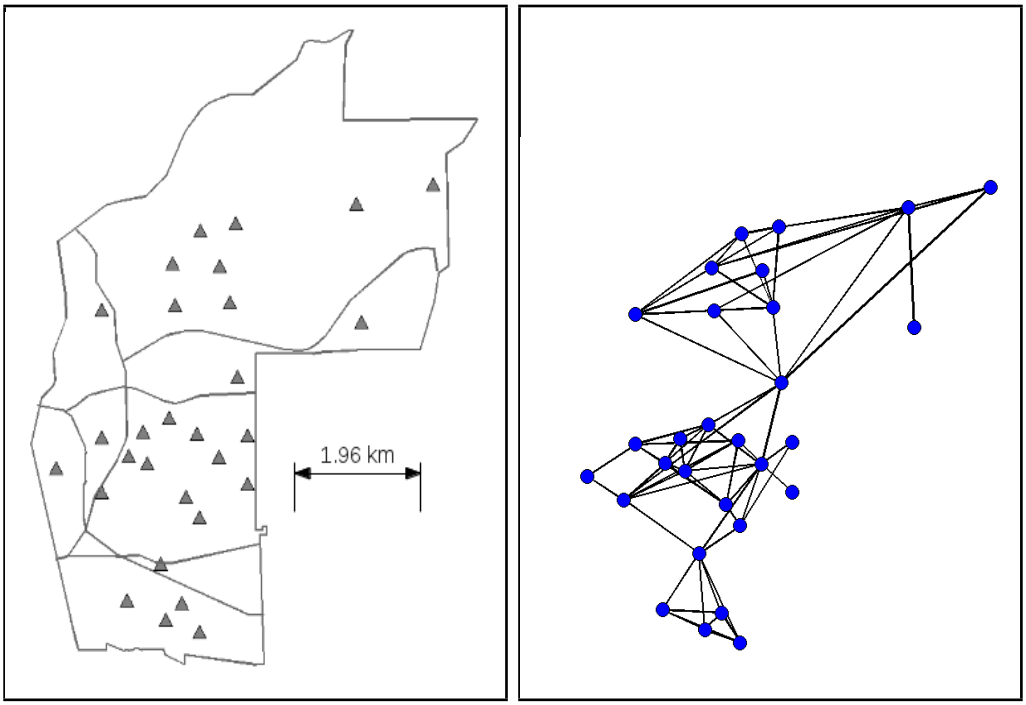


Figure 3.2: Set spaces embedded in Hollenbeck (Left) and the corresponding observed rivalry network (Right). Each solid triangle on the left and solid circle on the right gives the location of a gang's set space. The faint curves (Left) show the major roads and rivers that pass through the region, impacting the rivalry structure. The policing division boundary is also included as the bottom and right boundaries. The observed network from [102] is displayed on the right, where an edge between two gangs indicates an existing rivalry between the gangs.

## 3.4 Overview

The remainder of Part II contains the details of two approaches to modeling gangs' spatial behaviors and rivalries. Both methods use the Hollenbeck policing division as a case study for evaluation. Chapter 4 is the published version of the article on the stochastic agent-based approach [52]. Chapter 5 describes the proposed deterministic model based on an ecological territorial model [120].

# CHAPTER 4

# Geographical Influences of an Emerging Network of Gang Rivalries

This chapter is reprinted from Physica A: Statistical Mechanics and its Applications, volume 390, Rachel Hegemann, Laura Smith, Alethea Barbaro, Andrea Bertozzi, Shannon Reid, and George Tita, "Geographical influences of an emerging network of gang rivalries," pages 3894-3914, 2011, with permission from Elsevier [52].

## Abstract

We propose an agent-based model to simulate the creation of street gang rivalries. The movement dynamics of agents are coupled to an evolving network of gang rivalries, which is determined by previous interactions among agents in the system. Basic gang data, geographic information, and behavioral dynamics suggested by the criminology literature are integrated into the model. The major highways, rivers, and the locations of gangs' centers of activity influence the agents' motion. We use a policing division of the Los Angeles Police Department as a case study to test our model. We apply common metrics from graph theory to analyze our model, comparing networks produced by our simulations and an instance of a Geographical Threshold Graph to the existing network from the criminology literature.

## 4.1 Introduction

Street gangs are a growing problem around the world [27, 64, 62]. In fact, recent statistics from The National Gang Intelligence Center estimate there are 1 million active gang members

in the United States alone [101]. Violence is intrinsic to street gangs, and rival gangs battle to gain respect and street reputation [125, 29]. Criminal activities perpetrated by gang members, including armed robbery, homicide, drug dealing, and auto theft, drain cities and governments of tight resources and also pose safety threats to community members. Much of the research on street gangs has been conducted within the United States, though there have been some efforts to understand the phenomenon in Europe and other parts of the world [27, 64, 62].

Violence perpetrated by gang members is frequently against members of a different gang. In areas with numerous gangs, it is common for gangs to have multiple violent interactions with many of the other gangs. Further, street gang members typically have locations, known as *set spaces*, where they spend large quantities of time [126, 98]. It is therefore reasonable to think of each gang as a node embedded in Euclidean space [102, 128]. Within this framework, the existence of persistent violence between two gangs becomes an edge connecting two nodes. From this construction, one can view a collection of gangs as a spatially embedded network [127]. The Hollenbeck policing division of eastern Los Angeles is marked by a particularly high degree of violent crimes involving gang members, including homicides and aggravated assaults [102, 56]. It is for this reason and others listed in Section 4.1.4 that we consider Hollenbeck as a case study for our model.

Statistical approaches are often used to analyze gang activities. For example, in their recent paper, the authors of [102] used a network statistical approach called CONCOR to partition the region of Hollenbeck into areas of similar violence activity and geographic proximity. From this, they were able to conclude that a relationship exists between the location of a gang and its rivals. Other authors have utilized various statistical approaches to examine gang related data [128, 126]. The methodology of determining factors that correlate with gang activity has been able to describe certain features of the observed system. This approach has made important contributions in the field, but it cannot make solid causal arguments or test theories [72, 131]. Though one could argue that creating an experiment would help validate causal arguments, many of these experiments are not feasible due to monetary and resource constraints and could be unethical and infringe on basic human

rights. Even in circumstances where the resources are available, interventions may fail to be implemented as planned. This is evident in the work of [128].

Obtaining complete and valid data sets are common issues in the field of criminology. Data sets are many times unreliable through inaccuracies, underreporting, and potential bias [32]. This provides a fundamental problem in the conclusions made from faulty data. In [47], a strong case is made to move beyond statistical modeling and instead model social phenomena using a mathematical approach. We use such an approach with the aim to understand the plausible mechanisms for the formation of gang rivalries and to provide social scientists a means by which to test social theories. For comparison, we propose two different mathematical techniques: a pure network and an agent-based approach.

Our goal in this work is to understand how long-term rivalries among gangs develop by using a model including geography, social dynamics, and human mobility patterns. Because we have only one observed rivalry network, this is a very difficult inverse problem. Furthermore, statistical approaches are not able to determine causal effects. As a solution, we propose an agent-based model that is coupled to a dynamically evolving network. This bottom-up approach simulates the mobility of gang members by using the conclusions from current literature in human mobility patterns, see Section 4.2.1 for more details. These agents interact to form a rivalry network. We compare the resulting simulated network to the gang rivalry network observed in the eastern Los Angeles division of Hollenbeck [128, 102].

The outline of our paper is as follows: network and agent-based modeling approaches are outlined in Sections 4.1.1 and 4.1.2, and previous work on crime models is discussed in Section 4.1.3. Detailed information concerning the particulars of Hollenbeck are found in Section 4.1.4. In Section 4.2, we outline the proposed model. In Section 4.3, we describe two baseline models, one instance of a Geographical Threshold Graph and a network derived from Brownian Motion. We contrast these simpler models with our model to demonstrate the need for a more complex approach. In Section 4.4, we describe a series of metrics from network theory, examine long term behavior of the model, and compare the networks against the metrics. Section 4.5 provides a sensitivity analysis of our model. We conclude and give future directions in Section 4.6.

### 4.1.1 Network Models

General network models and the corresponding analysis are useful for describing the behavior of complex systems and have played an increasingly active role [88, 89, 91]. One way networks are treated in the literature is by analyzing the statistical properties of a given network. Another approach is to consider the construction of a network. There are many instances where the network of interest is not known, but there is some knowledge of the processes by which the network is formed. One popular method to construct a network is to view it as a random graph. Each edge is added with a predetermined probability, often dependent on the weight of the nodes [2, 92, 93].

In some applications, including gang rivalry networks, the geographic location of the nodes influences the structure of the network. In such cases, geographic features should be considered as part of the random network model. For example, interstate highways have been shown to be structurally different from scale-free networks such as Internet and airline flight networks [42]. The importance of geography is also seen in friendship networks in [138]. This work was descriptive by nature, and was therefore limited in the conclusions that could be drawn. In their paper, Liben-Nowell et al. use the publicly accessible location of 495,836 bloggers in the Live Journal to investigate effects of spatial proximity on friendships [70]. The study found that an estimated 69% of a person's friends can be described by geography. This paper constructed a simulation and was able to create plausible scenarios and test theories of information spreading on a social network, again highlighting a strength of a mathematical modeling approach.

In the scenario of gang rivalries, the geographic location of gangs play a role in the observed gang rivalry structure [128]. Because of this, one of our approaches is to construct a simple network model that incorporates the proximities between gang set spaces. One method for incorporating geographical information into the random graph construction is by using a Geographical Threshold Graph [74, 17, 18]. This is a random graph on a set of randomly weighted nodes, where the nodes are located in a metric space and the connections are determined by thresholding a function of the distance and the weights. This provides an

efficient way to construct a rivalry network while incorporating some geographic information. We use an instance of a Geographical Threshold Graph as a baseline against which we compare our model.

### 4.1.2 Agent-Based Models

Network models can provide a computationally inexpensive means to reproduce an observed network, but these models do not lend themselves to describing phenomena of interest beyond the structure of the network. An alternative modeling approach is to use agent-based models. These models are generally used for complex systems, since they are able to capture details at the level of an individual, or *agent*. In this class of models, agents often move through phase space and interact amongst themselves, producing complex dynamics and patterns from simple behavioral rules. Agent-based models can answer different questions from a network approach since they focus on the way individuals actions can determine the behavior at the system level.

This type of modeling approach is widely used in economics [132, 124, 141], epidemiology [37], sociology [119], biology [54], and other situations in which complex systems are encountered. The strength of agent-based modeling is that it allows for responses at the individual level to be directly incorporated into the model. For example, agent-based models have been proposed in the context of searching and swarming [71], and much interesting mathematical analysis has been done on the behavior of these systems [30, 26]. In this way, modelers can include information and behavioral dynamics from scientists who study the complex system of interest. Agent-based methods have been proposed to describe social systems and economics since they can provide a means by which to test theories about individual dynamics in cases where the dynamics are not precisely known [38, 72].

In our model, we are interested in the coupling between the network and the underlying system. There has been some exploration of this in the literature. For example, Schweitzer and Tilch provide one example of a model that uses an agent-based approach to form an emerging network [108, 109]. They model the chemical trail formed by ants searching for

food at an unknown location. As the ants search their environment, networks of chemical trails form with which the ants interact. Another example is that of the EpiSims model [130, 76]. Here, the contact networks of the populations are evolving over time and depend on the internal attributes of the people in the population. In turn, as a disease is spread through the contact network, the movements of the people change in response to the disease, producing a non-trivial interaction between the system and the network.

One of the major strengths of this approach is the flexible framework available for these models. For example, this method can easily incorporate environmental and spatial information inherent to the system. One example of this is shown in [12], where the authors were able to incorporate environmental information in the form of temperature and current, as well as geography of landmasses, to accurately model and predict the migration patterns of a species of pelagic fish. In our agent-based model, we use information about the environment in the form of freeways, rivers, and road density. In a network context, coupling an agent-based model to a network is a novel way to explore how changing dynamics of individual agents can affect the evolution of the network, providing control parameters which would be inaccessible in a graph-based model.

### 4.1.3 Previous Work on Crime Modeling

Many mathematical models have been created to study the patterns, mechanisms, and potential interventions associated with crime. Research has been conducted to address various aspects of burglaries such as hot spot formation [113, 115, 19], policing strategies [60, 100], criminal cooperation [114], and geographical profiling [96]. Some attempts have also been made to model gang behaviors. The study in [116] examines the short term retaliatory behavior of the rivalries based on between gang violence data from the LAPD. In this work, each violent event between two gangs is considered an instance of a point process associated with that pair of gangs. The intensity of the rivalry depends directly on the network of unidirectional violent interactions. This provides a top-down approach to understanding immediate consequences of violence among gangs within a system.

In the model proposed by [34], an agent-based approach was used to simulate the location of violent interactions and gang retaliations in Hollenbeck. Embedded in the model was a rivalry network. Though the model recreated similar features to the violence data, the model made unrealistic assumptions about the mobility of the agents. For instance, the agents' movements were influenced by the location of every rival gang member, an implausible assumption. Further, agents only move towards rival gangs members, ignoring their own set space, or center of activity. However, according to the criminology literature, this retaliatory behavior is only seen on short time scales [116]. Furthermore, gang members tend to avoid the territory of rival gangs and spend large quantities of time at their set spaces [5, 69]. Another concern is that the the agents ignore geographic features, such as highways and rivers, known to correlate with the rivalry structure [102]. It has been shown in other models simulating human movement that highways restrict movements [48].

We propose a bottom-up approach with an agent-based model that incorporates movement rules from literature on human mobility in order to capture the long-term gang rivalry structure. These rules consider geographical features known to relate to movement dynamics. We also consider current literature on known gang behavior as a basis for directional decisions. More details are seen in Subsection 4.2.1. We use Hollenbeck as a case study for our model.

### 4.1.4 Hollenbeck

Hollenbeck is a policing division located in eastern Los Angeles, surrounded by downtown Los Angeles to the west, Pasadena to the northeast, Vernon to the south, and to the east the unincorporated area of East Los Angeles, see Figure 4.1. Hollenbeck provides a diverse geography with many highways cutting through the region and is bounded by the Los Angeles River. It encompasses an area of roughly 39.4 $km^2$. Hollenbeck is home to approximately twenty-nine active gangs with sixty-nine rivalries among them [102, 128]. The set spaces for the gangs and the corresponding observed rivalry network are displayed in Figure 4.1, as given in [102].

Certain properties of Hollenbeck make it accessible to modeling the gang rivalry networks outlined in [128, 102]. First, it is a closed system in that the gang activity within Hollenbeck is generally isolated from gang activity outside of Hollenbeck. Further, the motivation for violence between gangs is largely characterized by disputes over geographical gang territories, as opposed to drug and racially motivated violence. Data on the geography of Hollenbeck is easily accessible, and there has been explicit documentation of the observed rivalry network.

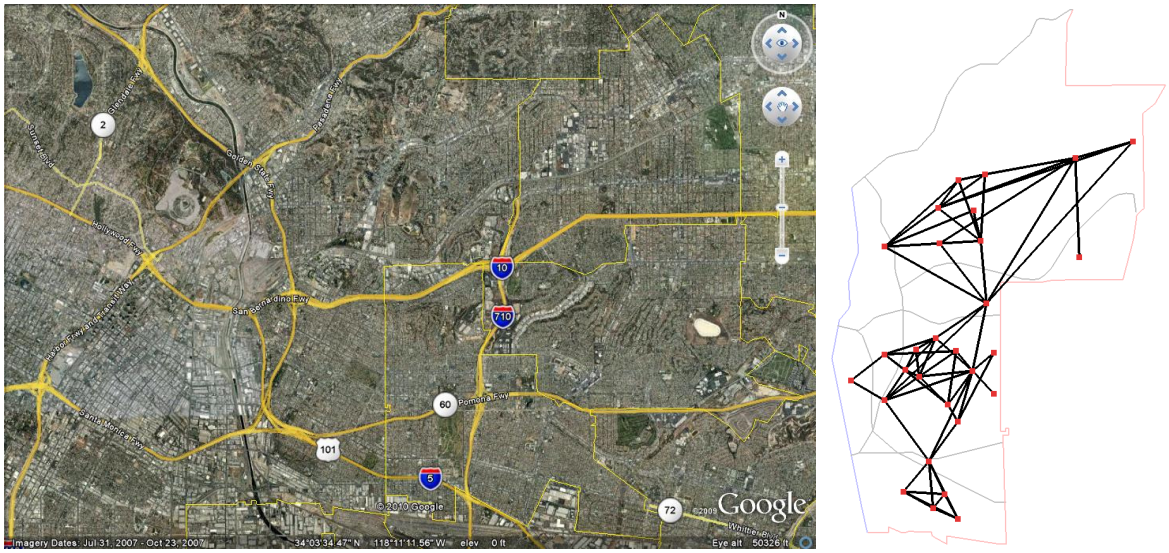


Figure 4.1: Google Earth$^{TM}$ Image of the Hollenbeck area (left). Map of the Hollenbeck area with the location of the gang set spaces, or centers of a gang's activity, and the corresponding rivalry network approximated by [102], where a node of the network represents a set space, and an edge represents a rivalry between two gangs (right). Major roads, highways, the Los Angeles river, and division lines are also seen in both images.

## 4.2 Our Model

Our objective is to model the long-term gang rivalry structure and gang member mobility by incorporating simple behavioral rules and geographical factors, such as road density, highways, and locations of gangs' set spaces (centers of activity). The movement of each agent is not intended to model each detail involved in an individual's mobility, but rather capture the statistical behavior of people's movements observed from the literature. Agents in the model move based on their location with respect to their and other gangs' set spaces and in-

teract with agents of different gang affiliations. We count the number of interactions between gangs, and when agents of different gangs move within a certain distance of each other, the number of interactions between those gangs increases by one. As the simulation progresses, a network structure emerges. The weighted network of interactions in turn influences the directional decisions of the agents.

### 4.2.1 Motivation for Model Construction

The intent of this model is to capture the broad statistical features of human mobility with an emphasis on gang members' movements. Empirical data on the location and individual movements of each gang member is inaccessible, so we characterize the movements of the individual gang members in a statistical sense based on the literature on human mobility. Several studies give compelling evidence that when people move unconstrained environment, the jump lengths between movements is distributed like a power law [24, 103]. Further, in the presence of obstacles such as roads and buildings, the jump lengths more accurately follow a bounded power law distribution [45].

Determining the statistical properties of the jump length is only one aspect of movement dynamics. In their paper, Rhee et al. discuss the need to incorporate geographical features and the tendency for people to go home [103]. Gonzalez et al. confirmed in their data that humans do tend to frequent a small number of locations often [45]. For these reasons, the agents in our model pick their jump length from a Bounded Pareto distribution and have a directional choice in movement.

In the case of gangs in Hollenbeck, it is reasonable to assume that the gang members have a clear sense of the location of their home territory, or set space, as well as the location of their rival gangs' set spaces [126]. Literature on gang activity suggests that, in general, gang members tend to stay away from their rival gangs' set spaces [69]. Unlike other criminal groups, such as organized crime syndicates and insurgency groups that strive for secrecy, street gangs are social organizations that proudly demarcate their territory and announce their enemies through the use of graffiti. Gangs create social boundaries and therefore

areas of avoidance [5]. Our model incorporates this social geography into agents' movement dynamics.

One aspect of modeling human mobility that was touched on, but not fully explored, by the previous literature is the role of physical features specific to urban areas that may constrain agents' movement. The first consideration is the ease with which an agent can move through a city. We posit that in areas where there is a dense street network, the likelihood of an agent to move long distances is small due to such obstacles as the high density of people and cars, as well as traffic lights. On the other hand, areas where the road density is lower, agents should be able to move longer distances. A second physical consideration that affects human mobility in a city are the highway systems and rivers that can cut across the region. These features are not impassible, in that there are underpasses and bridges. Simulations have shown that they provide an obstacle to mobility [48]. It has been posited that these features play a role in gang rivalry networks [102]. Therefore in our model, we view major roads, highways, and rivers as semi-permeable boundaries that affect the agents' movements.

### 4.2.2 Model Details

#### 4.2.2.1 Agents

The agents of this model are gang members in a city. Each agent is associated with exactly one gang. For simplicity we assume agents' directional choice is dictated only by the location of the gang set spaces, or centers of activity. All agents know the location of their home and rivals' set spaces. We divide the city into regions based on geographical boundaries, such as rivers and highways. An agent knows which region it is currently in as well as the region of any prospective new locations. Embedded in the city is a rivalry network among the gangs in the city. This network is encoded in the rivalry matrix $R$. Each element $(i, j)$ of $R$ corresponds to the number of interactions between gang $i$ and gang $j$. When two agents are within interaction range, we consider them to have interacted and the corresponding element of the rivalry matrix, $R$, is updated. Refer to Section 4.2.2.3 for details on $R$.

#### 4.2.2.2 Environment

The environment of interest is on the scale of a small city. Agents and gang set spaces, or centers of gang activity, have a coordinate location in free space. The rest of the physical features of the city are approximated by an NxM grid. Each point in free space is identified with the nearest grid element. The size and number of grid elements are constant throughout the simulation and are limited by the available data and the memory of the computer.

Two specific features encoded by this grid are the road density and semi-permeable boundaries represented by a region map. Each grid element of the road density contains a number between 0 and 1. A value of 0 implies a low road density whereas a value of 1 implies high road density. The semi-permeable boundaries, corresponding to such objects as highways and rivers, are assumed to split the environment into distinct regions. We pair this region grid with a transition matrix that stores the associated probability of an agent to cross from one region to another. These probabilities are determined at the start and remain constant throughout the simulation. This is implemented to discourage agents from crossing freeway boundaries.

#### 4.2.2.3 Rivalries

The network structure of the rivalries is encoded in a weighted adjacency matrix, $R$. Each element $R_{ij}$ contains the current history of interactions between gang $i$ and gang $j$. At the end of a simulation, we construct a thresholded rivalry graph where an edge between gang $i$ and $j$ exists if either $\rho_i(j)$ or $\rho_j(i)$ is larger than a given threshold $T$, where

$$\rho_i(j) = \frac{R_{ij}}{\sum_{k=1}^{N} R_{ik}} \qquad \text{and} \qquad \rho_j(i) = \frac{R_{ji}}{\sum_{k=1}^{N} R_{jk}}. \tag{4.1}$$

The quantity $\rho_i(j)$ represents the proportion of gang $i$'s interactions which have occurred with gang $j$. Note that $\rho_i(j)$ is not necessarily equal to $\rho_j(i)$; however, this thresholding yields a bidirectional network or, equivalently, a symmetric adjacency matrix. Although the final rivalry network is symmetric, the influence of the rivalry matrix $R_{ij}$ on the agents'

movements during the course of the simulations are not symmetric. The data available to us was in the form of a symmetric network, and so we chose the thresholding rule to be comparable with the data. Other thresholding rules that result in asymmetric networks could be implemented. For example, if the thresholding rule for gang $i$ solely depended on the quantity $\rho_i(j) > T$, then the resulting network could be asymmetric.

### 4.2.3 Process Overview and Scheduling

At each iteration an agent is chosen from the set of all agents with equal probability. The selected agent performs one step of a random walk by choosing a jump length and direction from probability distributions for its new prospective location. Depending on the distributions for the jump length and direction, different random walks will occur. The literature on human mobility suggests humans move according to a truncated Lévy walk, motivating our model selection. Further, it is unreasonable to assume that an person's direction of movement is solely determined by set space locations. Therefore, we use a statistical distribution to simulate an agent's directional choice. See Section 4.2.1 for more motivational details.

The jump length, $x$, is chosen from the Bounded Pareto probability density function,

$$P(x; k, x_m, x_M) = \frac{k x_m^k x^{-k-1}}{1 - \left(\frac{x_m}{x_M}\right)^k} \quad k > 0\ , \quad x_M \geq x \geq x_m > 0. \tag{4.2}$$

The minimum and maximum jump lengths, $x_m$ and $x_M$ respectively, provide the bounds for the jump length $x$. The scaling parameter, $k$, determines how quickly the probability density function decays from the minimum jump length to the maximum jump length. For all agents the minimum jump length, $x_m$, and scale, $k$, are fixed parameters. To determine the maximum jump length, $x_M$, the agent uses the approximated road density of the agent's corresponding location from the environment grid, with $x_M$ ranging from $a$ to $A$. These values remain constant throughout the simulation. Here, $A$ is called the largest maximum jump length and $a$ is the smallest maximum jump length. Then $x_M$ is calculated via

$$x_M = (1 - \delta) \cdot A + a, \tag{4.3}$$

where the road density at the agent's location, $\delta$, is between 0 and 1.

The second quantity needed to move the agent to the next location is the new direction, $\theta$. This is obtained by constructing a deterministic direction of bias. This bias incorporates the agent's location with respect to its home set space and the location of its rival gangs' set spaces. Agents have a higher probability of moving in this direction. However, with lower probability, they have the ability to move any direction. To account for this mobility pattern, the von Mises distribution is used to simulate the direction the agent will move, $\theta$.

More specifically, given an agent in gang $i$, the bias direction, $\mu_i$, is defined as

$$\langle y, z \rangle = H_i(||\overrightarrow{G_i}||_2) \ \frac{\overrightarrow{G_i}}{||\overrightarrow{G_i}||_2} + \sum_{j \neq i} D_{ij}(||\overrightarrow{G_j}||_2) \ \frac{\overrightarrow{G_j}}{||\overrightarrow{G_j}||_2} \tag{4.4}$$

$$\mu_i = \tan^{-1}\left(\frac{z}{y}\right).$$

Here, $\overrightarrow{G_l}$ is the vector that points to the set space of gang $l$ from the location of the agent. When $l = i$, this vector points towards the agent's home set space, and when $l \neq i$, it points towards a different gang's set space. This concept is shown in the cartoon example in Figure 4.2.

In Equation 4.4, $H_i$ gives the rules for weighting towards a gang member's own home set space. The weightings toward or away from different gangs' set spaces are determined by $D_{ij}$. Our $H_i$ and $D_{ij}$ take the following form:

$$H_i(||\overrightarrow{G_i}||_2) = h_i \ ||\overrightarrow{G_i}||_2, \tag{4.5}$$

$$D_{ij}(||\overrightarrow{G_j}||_2) = w_{ij}(R) \ \frac{1}{||\overrightarrow{G_j}||_2}. \tag{4.6}$$

One notable feature about these equations is that $H_i(\cdot)$ is large when an agent in gang $i$ is far from his or her gang's set space, but the $D_{ij}(\cdot)$ function is large when the agent is close to a rival gang $j$'s set space. The factors $h_i$ and $w_{ij}(R)$ of the weighting functions are

chosen according to the rules for agent movement. In our implementation, the factor $w_{ij}(R)$ depends on the current state of the rivalry network. Negative values of these functions result in repulsion and positive values result in attraction.

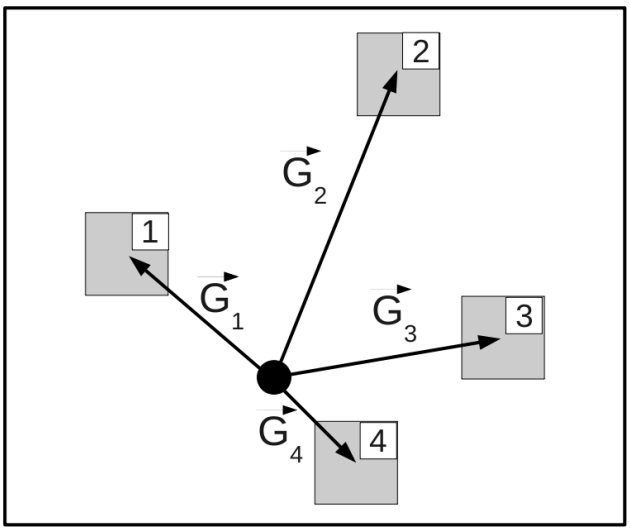


Figure 4.2: Cartoon example of the direction vectors incorporated in the direction of bias formula, Equation 4.4. The agent in this example is located at the dot. Here $G_1$, $G_2$, $G_3$, and $G_4$ show the vectors pointing toward the set spaces of gangs 1 through 4, respectively. Depending on the choices of $H_i$ and $D_{ij}$, different movement dynamics are possible.

After determining the direction of bias from Equation 4.4, we must choose in which direction the agent will move. The direction, $\theta$, is drawn from a von Mises distribution (also known as the Circular Normal distribution) [73, 57, 14]. For $\theta \in [-\pi,\ \pi]$, the von Mises distribution is given by

$$f(\theta|\mu,\kappa) = \frac{\exp\left(\kappa\cos(\theta-\mu)\right)}{2\pi I_0(\kappa)}.$$

Here $I_0$ is a modified Bessel function of order zero. The von Mises distribution requires two parameters, one for the angle of bias, $\mu$, and one for the strength of the bias, $\kappa$. We can think of $\mu$ as being the mean of the distribution, and $\frac{1}{\kappa}$ as being comparable to the variance. The larger $\kappa$ is, the stronger the bias is for the direction $\mu$. If $\kappa = 0$, this is a uniform distribution on a circle.

From the direction, $\theta$, and jump length, $x$, a prospective location is calculated. The new location is then checked to see if the result would move the agent into a different region.

Within the same region movement is unrestricted. However, if its next move would result in a region change, i.e. it is crossing a semi-permeable boundary, it has a given probability of crossing into that region. If the agent moves, it searches the other agents to see if it is close enough to interact with agents of other gangs. When an interaction does occur, the rivalry matrix, $R$, is updated. The location of interactions is also recorded and could be of interest to other applications, see discussion in Section 4.6 and Figure 4.14.

The model is run until limiting behavior is observed in all of the metrics. In the absence of an observed network, it could be helpful to run simultaneous simulations with different random seeds and calculate the variance of each metric over the course of the simulation run. When the variances of each metric levels off, terminate the simulation. In the case of Hollenbeck, the final network is taken after 20,000,000 iterations and then thresholded to ignore infrequent interactions. For more details on long term behavior of the model, see Section 4.4.2.

### 4.2.4 Initialization and Input Data

Before the simulation begins, the region map and an estimated density of the road networks must be provided in matrix form on the same grid. The probability of crossing each boundary must also be provided. Additionally, parameter values must be specified. Table 4.1 describes the full list of parameters needed for implementation. At the start of the simulation all of the agents are located at their gang's set space. The size of each gang must also be specified.

### 4.2.5 Hollenbeck Parameters

The grid of environment features of Hollenbeck was approximated from the Google Earth$^{TM}$ image in Figure 4.1. Hollenbeck is about 39.4 km$^2$ [102, 128]. In our implementation, one Hollenbeck city block corresponds to approximately six grid elements. The interaction radius between agents is 3 units, or roughly half a city block. The approximated road density and region grids are show in Figure 4.3. The boundaries of the Hollenbeck region were approximated using points from the geographic features visible from Google Earth$^{TM}$. These

| Parameters | Acceptable Values | Hollenbeck Values | Tested Range |
|---|---|---|---|
| Agent | | | |
| $x_m$ | $0 < x_m < a$ | 0.1 | |
| $k$ | $0 < k$ | 1.1 | [1, 1.9] |
| $\kappa$ | $0 \leq \kappa$ | 3.5 | [1.5, 5] |
| $h_i$ | $-\infty < h_i < \infty$ | 1 | |
| $w_{ij}(R)$ | $-\infty < w_{ij}(R) < \infty$ | $-\rho_i(j)$ | |
| Environment | | | |
| $N_i$ | $N_i \in \mathbb{Z}^+$ | $14 \leq N_i \leq 598$ | |
| $S_i$ | $S_i \in \mathbb{R}^2$ | see Figure 4.1 | |
| $A$ | $a < A$ | 200 | [100, 400] |
| $a$ | $x_m < a < A$ | 100 | [100, 200] |
| $B$ | $0 \leq B \leq 1$ | 0.2 | [0, .5] |
| Network | | | |
| $T$ | $0 \leq T$ | 0.04 | [0, 0.06] |

| Parameters | Description |
|---|---|
| Agent | |
| $x_m$ | Minimum jump length |
| $k$ | Bounded Pareto scaling parameter |
| $\kappa$ | Von Mises scaling parameter |
| $h_i$ | Home weighting |
| $w_{ij}(R)$ | Rival gang weighting |
| Environment | |
| $N_i$ | Number of gang members in gang $i$ |
| $S_i$ | Location of gang $i$ set space |
| $A$ | Largest maximum jump length |
| $a$ | Smallest maximum jump length |
| $B$ | Permeability of boundaries |
| Network | |
| $T$ | Threshold for existence of an edge |

Table 4.1: Parameters needed for model implementation are listed in the first column. The second column lists theoretically acceptable parameter values. The values corresponding to the SBLN are displayed in the Hollenbeck Values Column. The Tested Range column provides the range for each variable for simulations run. The second table provides a description of each of the parameter values.

boundaries were used to construct the region grid. To approximate the road density, the Weighted $H^1$ Maximum Penalized Likelihood Estimation method was used [121]. Although

the area of Hollenbeck does not have large invalid regions of agent movement, alternative cities could have regions where human mobility is not expected to occur, such as in lakes, mountains, and oceans. Other methods for density estimation, such as kernel density estimation or other Maximum Penalized Likelihood Estimation (MPLE) methods, could also be used to construct the road density [118, 35, 78, 46]. To extend the approximated road density to the same sized grid as the region grid, the average value of the density over Hollenbeck was computed and used for the extended regions. The number of agents in each gang reflects historical information obtained from the LAPD.


Figure 4.3: The image on the left shows the location of Hollenbeck in the N $\times$ M environment grid. The semi-permeable boundaries encoded in the model are displayed in the center image. The shades of gray of this image are used to distinguish among regions. On the right, we used a Weighted $H^1$ Maximum Penalized Likelihood Estimation method with a road map as the initial data to approximate the road density of Hollenbeck [121]. The scale, seen on the far right, gives the approximated road density intensity. Light shades of gray correspond to high density values near one and dark shades correspond to low densities near zero.

The boundary crossing probability between the regions was calculated by the minimum number of boundaries one must cross to get from one region to the next. For instance, if region 1 and region 2 were separated by one boundary, the agent would have a probability, $B$, of accepting a move from region 1 to region 2. If region 1 and 2 were separated by $\alpha$ boundaries, then the agent would have a $B^\alpha$ probability accepting the move.

## 4.3 Baseline Comparison Models

### 4.3.1 Geographical Threshold Graphs (GTG)

For comparison to the networks produced by our simulations, we constructed an instance of a Geographical Threshold Graph (GTG). Geographical Threshold Graphs are random graphs that use spatial proximity to assist in determining whether or not two nodes are connected with an edge [74, 17, 18]. Geographical Threshold Graphs randomly assign weights $\eta_i$ to the $N$ nodes. Then, using an interaction function $F(\eta_i, \eta_j)$, an edge between nodes $n_i$ and $n_j$ exists only if

$$\frac{F(\eta_i, \eta_j)}{d(n_i, n_j)^\beta} \geq \text{Threshold},$$

where $d(n_i, n_j)$ is the distance between nodes $n_i$ and $n_j$. Constructing an instance of this graph is fast and computationally inexpensive.

Geographical Threshold Graphs are not deterministic in general. However, we are using this framework to construct one network in order obtain a reasonable comparison to our proposed model. To this end, we provide deterministic weights, $\eta_i$, taken from data to be the size of each gang. In our case, we take the multiplicative weight function $F(\eta_i, \eta_j) = \eta_i \cdot \eta_j$, since this is the number of possible pairings between members of gang $i$ and gang $j$. We use Euclidean distance for the $d(n_i, n_j)$ function and choose $\beta = 2$. The threshold was chosen to give the same number of rivalries as the observed rivalry network.

### 4.3.2 Brownian Motion Network (BMN)

Another model we use to compare with the simulated network is a simplified version of the proposed model using Brownian Motion and unbiased movement rules. Experiments without encoding the semi-permeable boundaries were conducted with unsatisfactory results. Therefore, the semi-permeable boundaries of the model are incorporated. In this model, each agent chooses the next prospective location from a standard normal distribution, ignoring any directional decisions. These simplifications reduce the number of variables to the threshold, $T$, and the permeability, $B$, while still incorporating the geographic boundaries. The parameter

space around the Hollenbeck values was explored and run for $2 \cdot 10^7$ iterations. *A priori*, it was unclear how many iterations to run the simulation. We observed that the accuracy of the Brownian Motion networks peaked around $1.2 \cdot 10^7$ iterations and then decreased as the simulations progressed. The parameters and number of iterations that produced the highest accuracy were used for analysis. We will refer to the resulting network as the Brownian Motion Network (BMN).

Inherent in the BMN is a level of stochasticity. To understand how this stochasticity influences the final rivalry network and the resulting metrics, the BMN simulation was run for 100 different seed values. The resulting collection of final networks will be called the Ensemble BMN.

### 4.3.3 Baseline Network Graphs

Figure 4.4 displays the resulting GTG and BMN as compared to the observed rivalry network. The lower portion of the GTG graph has similar shape to the observed network, but contains more connections. The GTG does not make long connections. This is particularly evident in the upper half of Hollenbeck. The BMN picks up many of the longer connections, but includes far too many connections.

## 4.4 Results

The proposed model produces strong results when compared to the baseline models. For analysis of the model, this section is divided into two parts, the internal properties of the model and the comparison among models. Due to the stochastic nature of the movement rules, the final network is not deterministic. Despite fluctuations among simulation runs, within a single run the model exhibits long term stable behavior in the metrics used for analysis. The stochasticity and long term behavior are the internal properties we examine in detail in Subsection 4.4.1 and Subsection 4.4.2, respectively. The stochastic nature of the model allows for a more realistic scenario, in the case where the observed network is just one instance of a random process. Further, the existence of stable long term behavior in

Figure 4.4: A visual comparison of the observed rivalry network (left), GTG (center), and BMN (right). Here, a node of the network represents a set space, and an edge represents a rivalry between two gangs.

our model is important to replicating the observed system, since research has demonstrated that the rivalry networks among gangs tend to be stable over time [127, 102, 98]. A more detailed discussion of this can be seen at the end of Subsection 4.4.2.

After an examination of the internal properties of the model, we subject the models to a number of metrics in order to capture the features of the observed network as well as the accuracy of model networks. Keeping in mind the potential for variants of an observed network, the measures we chose to evaluate the performance of our model are fairly robust to small perturbations to the observed network. The metrics used to assess the shape of the network and the accuracy are defined in Subsection 4.4.3, and their corresponding results are displayed in Subsection 4.4.4.

For analysis and comparison, we took one simulation run as a showcase of the model. This network was obtained by searching the parameter space within the ranges specified in the fourth column of Table 4.1, allowing for dependencies between parameters. The 34,128 simulated networks were then sorted according to accuracy, defined in Equation 4.7. Because each of the gangs in Hollenbeck are active, the graph with the highest accuracy with all non-

zero degree nodes was chosen. The parameter values for the optimal run are found in the third column of Table 4.1. We will to refer to this as Simulated Biased Lévy walk Network (SBLN). Figure 4.5 displays the network with our optimal parameters. The SBLN has a shape and structure similar to the observed network, but does not capture all of the longer edges. We also verified that all of the metrics we use to evaluate our model have reached a statistical equilibrium for the SBLN.

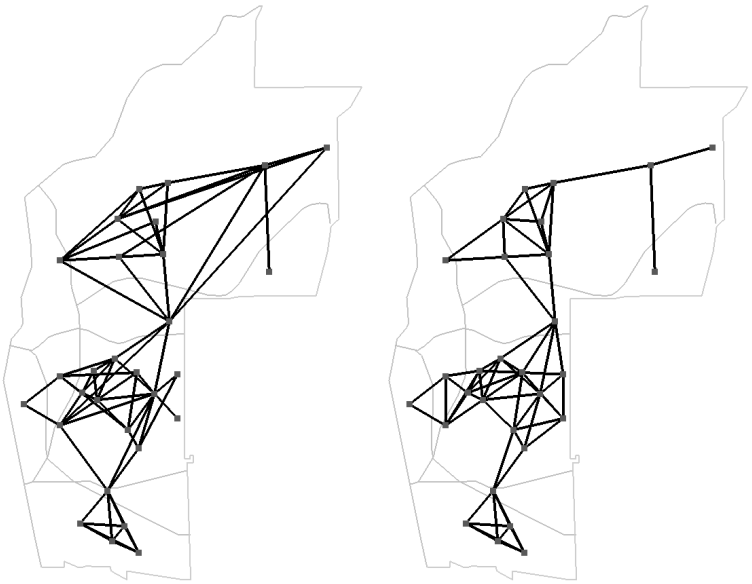

Figure 4.5: Comparison of the observed rivalry network (left) and the SBLN (right). Here, a node of the network represents a set space, and an edge represents a rivalry between two gangs. The SBLN has a shape and structure similar to the observed network, but does not capture many of the longer edges.

### 4.4.1 Stochastic Effects Observed in the Simulated Biased Lévy walk Network (SBLN)

Implicit in the model is a degree of stochasticity intended to capture the gross features of human movement. In particular, the jump length and direction choice are sampled from probability distributions, and the directional bias is determined by the (inherently stochastic) current rivalry structure. These elements affect the inclusion and exclusion of rivalry network edges. To understand the effect of stochasticity on the network produced by the model,

each simulation was run 100 times with different random seed values with the same SBLN parameter values. We refer to the collection of runs as the Ensemble SBLN. Each simulation was run independently and evaluated with several metrics. The resulting metrics were then averaged for analysis.

We also recorded the persistence of each edge in the ensemble of networks, and this is denoted as the *percent edge agreement*. For example, an ensemble network with 10% edge agreement refers to a network consisting of all edges that appear in at least 10% of the runs. Figure 4.6 displays the Ensemble SBLN with 100%, 50%, and 1% edge agreement next to the observed rivalry network. As expected, increasing the percent edge agreement decreases the number of edges present in the network. The network constructed with 100% edge agreement does not give a close representation of the observed network, because there are too few edges. However, allowing for 50% edge agreement produces a similar shape to the observed network. The Ensemble SBLN 1% edge agreement network shows all possible edges observed in the ensemble of simulation runs. Taken together, these images demonstrate the stochastic effects inherent in the model.

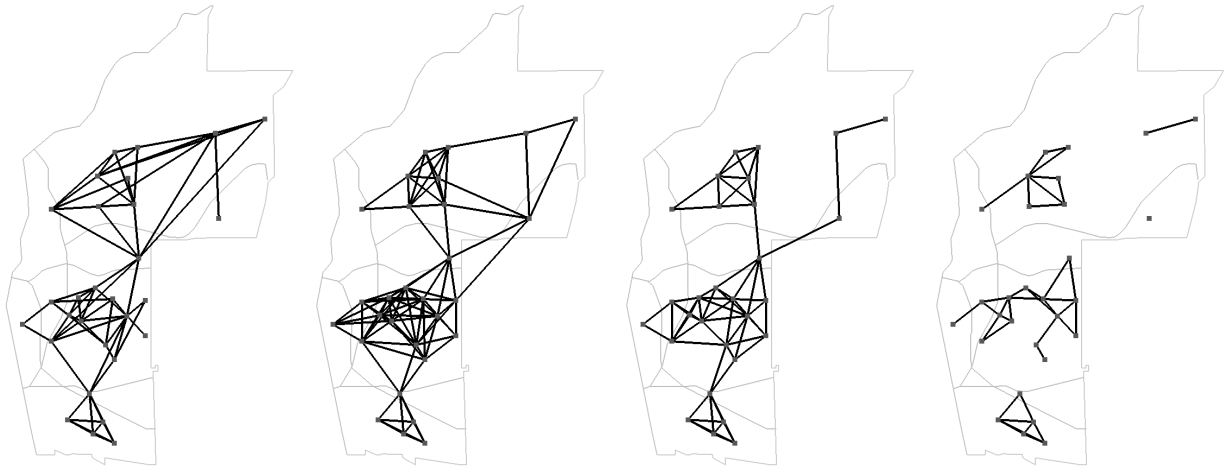

Figure 4.6: Percent edge agreement for the ensemble of runs for the SBLN parameter values. These four images give a comparison of, from left to right, the observed rivalry network, the Ensemble SBLN 1% edge agreement, the Ensemble SBLN 50% edge agreement, and the Ensemble SBLN 100% edge agreement. Here, a node of the network represents a set space, and an edge represents a rivalry between two gangs.

For comparison, we simulated a random model that incorporates only the distance be-

tween nodes. In particular, we constructed a collection of randomly weighted Geographical Threshold Graphs by fixing the locations of the nodes and sampling the weights, $\eta_i$, independently from a uniform distribution. We selected a threshold to yield a median of 69 edges. Figure 4.7 displays the percent agreement of each possible edge for the Ensemble SBLN, a collection of randomly weighted Geographical Threshold Graphs, and the Ensemble BMN. For visualization, the edges for each ensemble were sorted separately in descending order based on percent edge agreement. In the Ensemble SBLN, there is 100% edge agreement for the existence of 39 of the edges (corresponding to the first 39 edges of the Ensemble SBLN along the horizontal axis in Figure 4.7). The 100% edge agreement network in Figure 4.6 shows these edges. All runs in the Ensemble SBLN consistently agree on the nonexistence of 309 edges (corresponding to the last 309 edges of the Ensemble SBLN in Figure 4.7). These are the edges not appearing in the 1% edge agreement network in Figure 4.6.

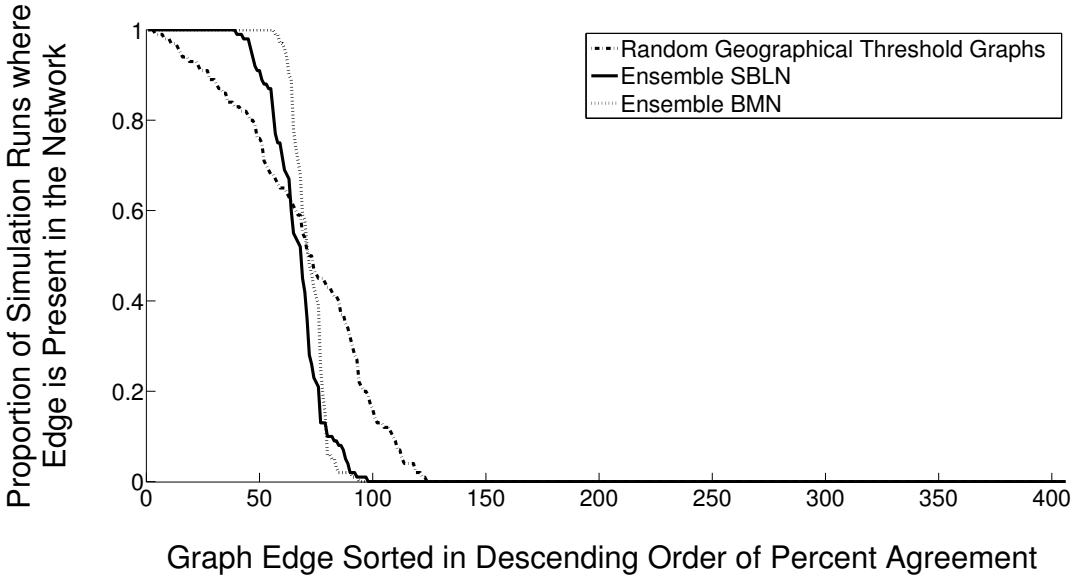


Figure 4.7: Plot of the edge persistence for the Ensemble SBLN (solid), Ensemble BMN (thin-dash), and an ensemble of random Geographical Threshold Graphs (thick-dash). The randomly weighted Geographical Threshold Graphs were constructed with random weights and have a median of 69 edges present. The edges were sorted in descending order according to the proportion of simulation runs where the edge is present in the network. Each ensemble of runs were sorted separately, yielding different edge numbers among ensembles.

The transition between edge existence and nonexistence in the Ensemble SBLN is marked

by a steep drop over 58 edges. The collection of randomly weighted Geographical Threshold Graphs displays a large degree of stochasticity indicated by fewer edges with 100% edge agreement and the more gradual decline of edge agreement. The Ensemble BMN appears to have a smaller degree of stochasticity with more edges with 100% edge agreement and a steeper decline than the Ensemble SBLN and the collection of randomly weighted Geographical Threshold Graphs. Despite the stochasticity observed in these models, there is agreement among the edges of the Ensemble BMN and Ensemble SBLN, maintaining some structure within the simulated networks.

Further analysis was conducted on the stochastic nature of the proposed SBLN model by looking at the metrics calculated on the network produced with varying levels of percent edge agreement. Plots of the accuracy, density, and nodal degree variance for the collection of randomly weighted Geographical Threshold Graphs, the Ensemble SBLN, and the Ensemble BMN are displayed in Figure 4.8. A definition of these metrics can be found in Subsection 4.4.3. As the percent edge agreement increases, the accuracy metric generally increases for the randomly weighted Geographical Threshold Graphs. However, the Ensemble SBLN and Ensemble BMN do not show much variation, though it is notable that the Ensemble SBLN consistently has higher accuracy across percent edge agreements. For the density metric, as the percent edge agreement increases, the density decreases as expected. The randomly weighted Geographical Threshold Graphs have the greatest variation in the density metric and nodal degree variance metric. This implies that there is a greater degree of stochasticity for this model. Alternatively, more structure is seen in the Ensemble SBLN and Ensemble BMN. The two ensembles show small variations in the density and nodal degree variance metrics for percent edge agreements between 15 and 75. Note that the Ensemble BMN has little change in nodal degree variance across all percent edge agreements. This indicates a lower degree of stochasticity in the model, which agrees with the steep decline in Figure 4.7.

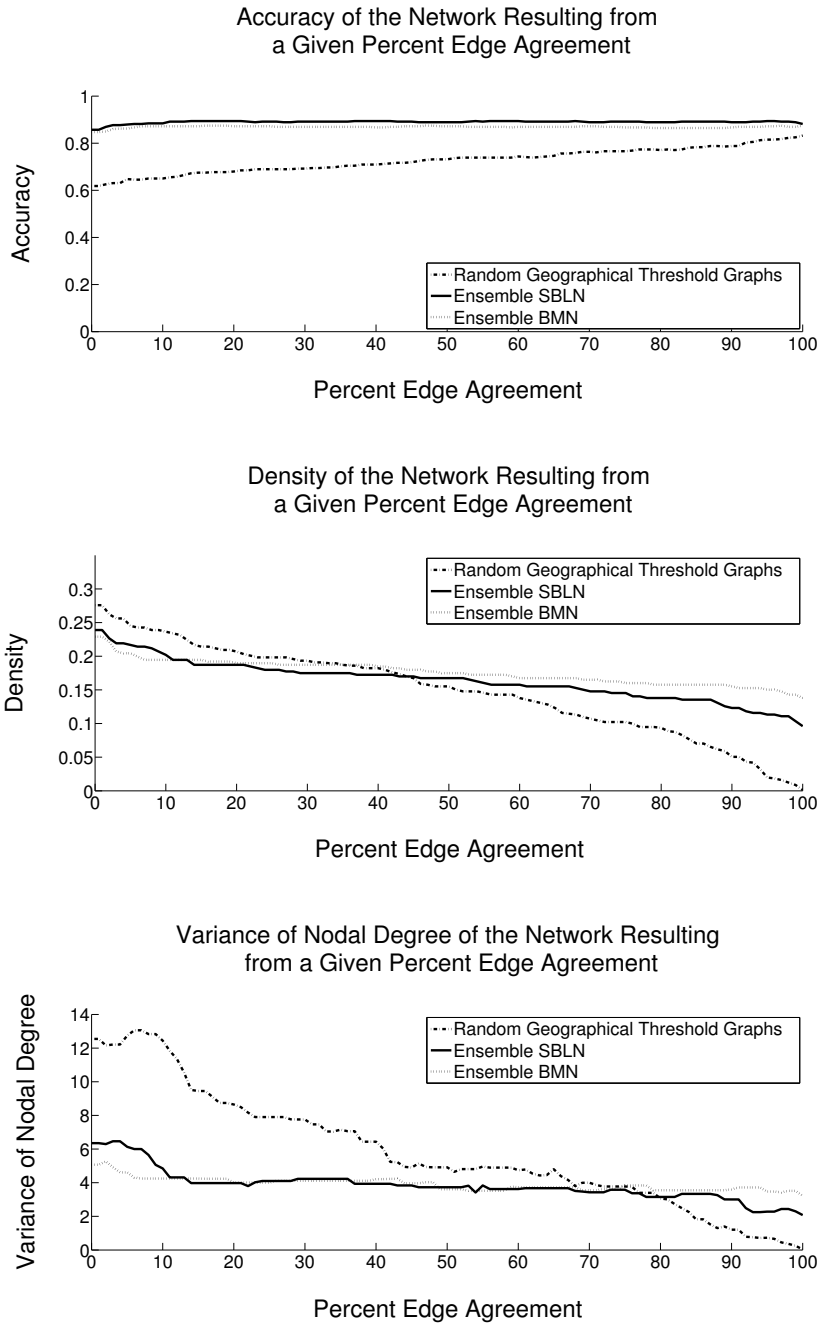


Figure 4.8: Plots of the measured Accuracy (Top), Density (Middle), and Variance (Bottom) of the networks constructed from a given percent edge agreement. The metric values corresponding to the Ensemble SBLN (solid), Ensemble BMN (thin-dash), and an ensemble of random Geographical Threshold Graphs (thick-dash) are plotted together. The randomly weighted Geographical Threshold Graphs were constructed with random weights and have a median of 69 edges present.

### 4.4.2 Long Term Behavior of the SBLN

The simulated network, through the movements of each of the agents, evolves as the simulation progresses. Because of this evolution, it is natural to ask if any sort of steady state is achieved. Keeping in mind the stochasticity of the model and the interaction between the network and the agents' movements, an equilibrium in the strictest sense cannot be obtained. Despite this, the results indicate there is limiting behavior of the observed metrics as the simulation progresses. Figure 4.9 displays the density and accuracy over the progression of the simulations for the Ensemble SBLN. In general, the accuracy metric we used is a measure of how close the simulated network is to the observed network. The density of a graph is proportional to the average number of rivals of all gangs; for further definitions of these metrics, refer to Subsection 4.4.3. Each run is observed every 1,000 iterations, and the results of each simulation are shown as a thin line. The average metric value at each iteration is calculated and plotted as the thick line. For visual investigation, the vertical axis on the accuracy plot has been refined to include only the area of interest. Accuracy values can range from 0 to 1. Both of these plots suggest that after a short phase of initialization, the metrics of each run seem to stabilize. For the average values of the density and accuracy of the last iteration, refer to Tables 4.2 and 4.3. Further, we tracked the variance of the metrics over the course of the simulation; plots of the variance for density and accuracy are shown in Figure 4.10. We observe that the variance of the network metrics across simulations levels out, indicating an appropriate time to terminate the simulation.

The Ensemble SBLN exhibits stable long term behavior the simulated rivalry network, with some variation due to stochasticity. Despite this variation, the network emerging from the model results in metrics with a small deviation from the average. Further, the stochasticity observed may provide a more realistic model of the true rivalry structure. Research has demonstrated that the rivalry networks that link gangs tend to be stable over time [127, 102, 98], and that the activity spaces of gangs are anchored to specific places [126, 84]. However, over longer periods of times, the membership ranks of gangs may ebb and flow due to incarceration, individuals "aging out" of active status, or other forms of incapaci-

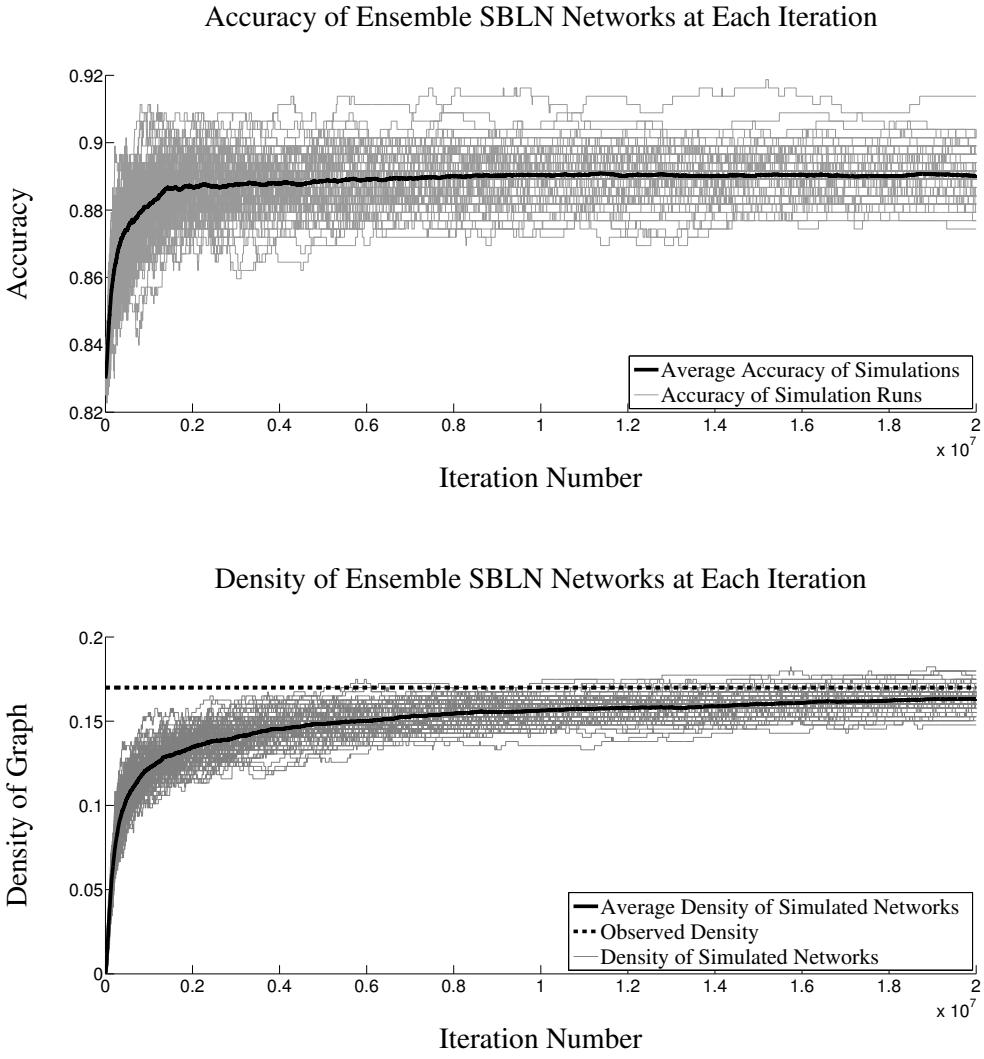


Figure 4.9: Plots of the accuracy (top) and the density (bottom) of the SBLN over the $2 \cdot 10^7$ iterations. Each of the 100 Ensemble SBLN runs are plotted by thin lines. The average over all the runs at each sampled iteration is shown with the solid, thick line. The density of the observed network is shown in the thick, dashed line. For visual investigation the vertical axis on the accuracy plot has been refined to include the area of interest. Accuracy values can range from 0 to 1.

tation [128]. Thus, gangs may lay dormant and, though identified in the rivalry network, not actually participate in violence. In extreme cases, either through high levels of victimization at the hands of rival gangs or through the focused enforcement of law enforcement agencies, a gang may simply disappear altogether. As more data become available, inherent

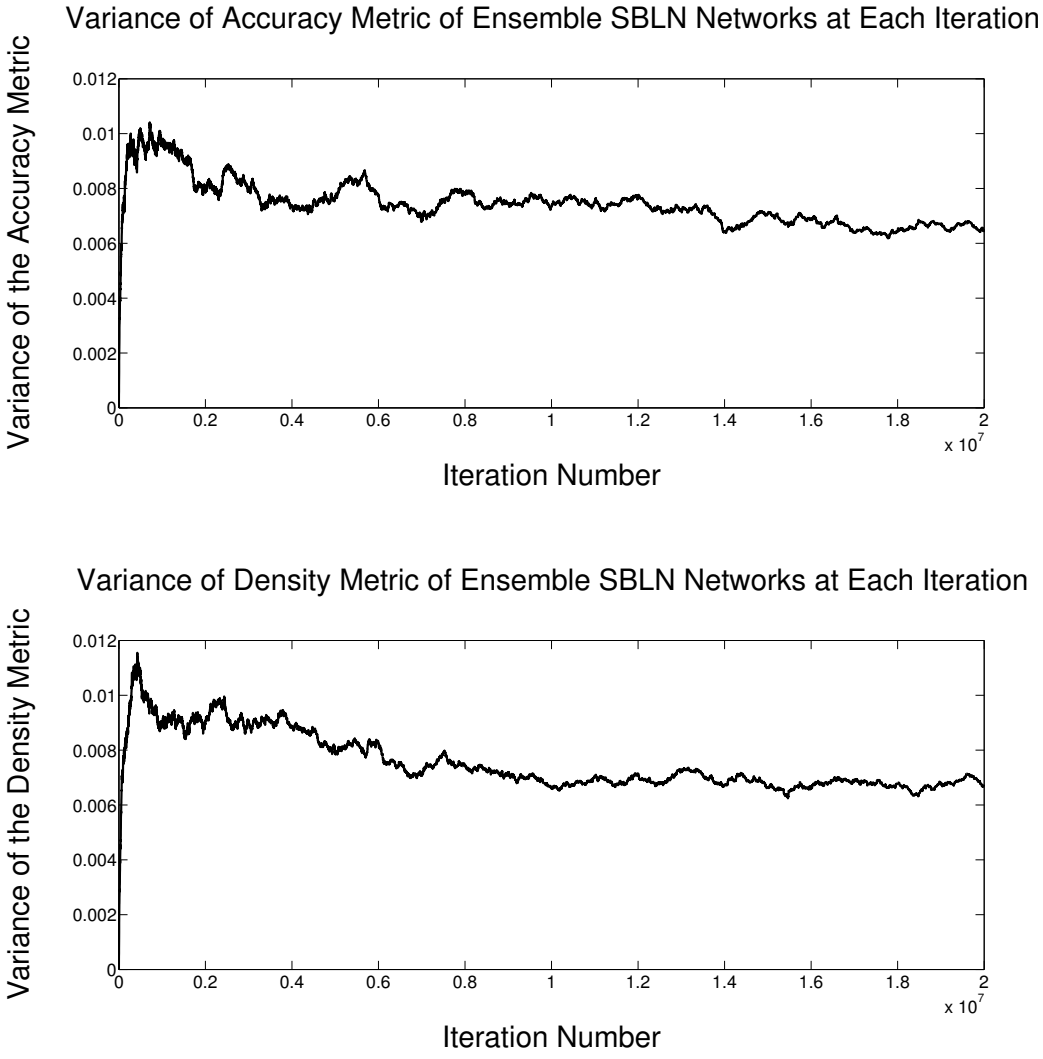


Figure 4.10: Plots of the variance of the accuracy (top) and the density (bottom) measures of the Ensemble SBLN over the $2 \cdot 10^7$ iterations. The variances of the 100 Ensemble SBLN runs were calculated at 1,000 iteration increments.

stochasticity in the model may allow for further understanding of the rivalry structure.

### 4.4.3 Metrics Used for Analysis

We analyze our model according to several common metrics of accuracy and shape. Since we are examining a rivalry network, there are certain popular measures that are not applicable here. For example, the clustering coefficient is often used in social and friendship networks,

describing the proportion of a node's neighbors who are also neighbors [2, 90, 136]. In the analysis of gangs, the rival of a rival gang is not necessarily a rival. In addition, there may be errors in an observed rivalry network. With these in mind, we carefully choose measures that are applicable to gang rivalry networks and robust to small perturbations in the observed data. One measure we consider, the degree distribution, is commonly used in social network analysis [142, 129, 90, 92, 91, 2]. In the context of gangs, the degree distribution of the network describes the number of rivalries of each gang. Knowing the distribution of rivalries at the city level may be helpful to policy makers for determining how to allocate resources. The measures used are motivated and defined in Sections 4.4.3.1 and 4.4.3.2.

#### 4.4.3.1 Accuracy Metrics

The first measures of interest are the raw values for the number of correct and incorrect edges. These values provide a means for evaluating the performance of the model. However, when comparing the observed network with the constructed network, each edge can be correct in two ways and incorrect in two ways. First, the constructed network can correctly identify an edge, *true positive* (TP), and correctly identify the lack of an edge, *true negative* (TN). The constructed network can also be wrong in two different ways. It can place an edge where there is none, *false positive* (FP), and also fail to place an edge where there is one, *false negative* (FN).

There are three quantities that are of particular interest that summarize the TP, TN, FP, and FN values. First is the *accuracy* of the model. The accuracy in the context of edges on a graph is defined by

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN}. \tag{4.7}$$

The ACC ranges between 0 and 1, with 1 being a perfect reproduction of the observed network. This measure is proportional to the $Q_\alpha$ measure discussed in [8]. The *F1 score* provides another measure to analyze the accuracy of the predicted network, [110, 143], and

is defined as

$$\text{F1} = \frac{2TP}{2TP + FP + FN}. \tag{4.8}$$

An exact replication of the network would have an F1 score of 1. The other summary statistic for the raw closeness to the network is the *Matthews Correlation Coefficient (MCC)* [75, 8]. This measurement varies between $-1$ and 1, where a value of 1 is a perfect prediction. The MCC is defined as follows:

$$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \tag{4.9}$$

The ACC, F1 score, and MCC provide a summary of the TP, TN, FP, and FN. These measures are fairly robust to changes in the observed data. For example if one edge were added or removed in the observed network, at most the the ACC would change by $\frac{2}{N(N-1)}$, where $N$ is the number of nodes.

The measurements of the TP, TN, FP, and FN provide one means by which to determine the success of the model. However, they do not describe how these correct or incorrect measurements affect the overall network structure. A strong model would create a network that is not only accurate but also, maintains the same network structure, even in the event that the individual connections are the not same.

#### 4.4.3.2 Shape Metrics

We would like to verify that the simulated network has a similar shape to that of the observed network. To do this, we calculate the graph *density*, standard *variance* of nodal degree and Freeman's *centrality measure* of the graph [136, 40]. Given $N$ nodes, the density of a graph is defined by

$$\sum_{i=1}^{N} \frac{\text{degree}(i)}{N(N-1)}. \tag{4.10}$$

In the context of gangs, the degree of a gang is equivalent to the number of rivals of the gang. The density of the rivalry network is the average number of rivalries scaled by a normalization, $\frac{1}{N-1}$. Networks with the same number of edges and nodes have the same density measure. Further, this metric is fairly robust to perturbations of the graph. For instance, if one edge were added or removed, the density metric would only change by $\frac{2}{N(N-1)}$.

Two other metrics, the variance of nodal degree and centrality measure, give an indication of the spread of degrees among the gangs. The variance of nodal degree for a graph is

$$\sum_{i=1}^{N} \frac{(\text{degree}(i) - \text{aveDegree})^2}{N}. \tag{4.11}$$

The centrality measure of the graph is defined to be

$$\sum_{i=1}^{N} \frac{\text{maxDegree} \ - \ \text{degree}(i)}{(N-1)(N-2)}. \tag{4.12}$$

These measures provide summary statistics with which to describe the general shape of the graph. This becomes useful when comparing the output to multiple models. For instance, if the observed network has a few gangs with high degrees and the rest of the gangs with low degrees, then this will be reflected high values in the spread statistics. A good model should capture this feature.

The *degree distribution* has been widely used to understand the overall network structure [90, 92, 91, 2]. We compare the nodal degree cumulative distribution function (CDF) of our simulations with the observed network.

### 4.4.4 Evaluating Models Using Graph Metrics

#### 4.4.4.1 Accuracy Metric Results

Table 4.2 provides the accuracy measures for the GTG, BMN, Ensemble BMN, SBLN, and Ensemble SBLN. The SBLN outperforms all of the other networks on all of the accuracy metrics. Observe that the GTG also performs well on these metrics. The Ensemble SBLN

metrics are comparable to the GTG and BMN metrics. In particular the average number of true negatives (TN) and false positives (FP) perform slightly better for the Ensemble SBLN than for the GTG, BMN, and Ensemble BMN. The Ensemble SBLN average of the true positives (TP) and false negatives (FN) performs slightly worse than the GTG and BMN. Only the GTG and SBLN have higher accuracy, F1 Score, and MCC values than the Ensemble SBLN average.

| | TP | TN | FP | FN |
|---|---|---|---|---|
| SBLN | 50 | 320 | 17 | 19 |
| Ensemble Average SBLN $\pm \sigma$ | 45.50 $\pm$ 1.269 | 316.1 $\pm$ 2.424 | 20.90 $\pm$ 2.424 | 23.50 $\pm$ 1.269 |
| GTG | 48 | 316 | 21 | 21 |
| BMN | 47 | 313 | 24 | 22 |
| Ensemble Average BMN $\pm \sigma$ | 43.61 $\pm$ 1.380 | 309.2 $\pm$ 1.390 | 27.76 $\pm$ 1.390 | 25.39 $\pm$ 1.380 |

| | ACC | F1 Score | MCC |
|---|---|---|---|
| SBLN | 0.9113 | 0.7353 | 0.6822 |
| Ensemble Average SBLN $\pm \sigma$ | 0.8906 $\pm$ 0.0077 | 0.6722 $\pm$ 0.020 | 0.6069 $\pm$ 0.025 |
| GTG | 0.8966 | 0.6957 | 0.6333 |
| BMN | 0.8867 | 0.6714 | 0.6031 |
| Ensemble Average BMN $\pm \sigma$ | 0.8691 $\pm$ 0.0051 | 0.6213 $\pm$ 0.016 | 0.5424 $\pm$ 0.019 |

Table 4.2: Accuracy measures for the SBLN, Ensemble SBLN, GTG, BMN, and Ensemble BMN. The $\sigma$ denotes the standard deviation of the ensemble metric values.

#### 4.4.4.2 Shape Metric Results

Table 4.3 provides the shape measures for the observed network, GTG, BMN, Ensemble BMN, SBLN, and Ensemble SBLN. Note that the density of the GTG is exactly the same as the observed rivalry network by construction, but it does not perform well for the nodal degree variance. The density for the BMN, Ensemble BMN, SBLN, and Ensemble SBLN are all close to the observed network. The BMN and the Ensemble BMN average have the closest nodal degree variance to the observed network's nodal degree variance. The centrality

measure for the SBLN is the closest to that of the observed network.

The cumulative distribution function (CDF) of nodal degree for the observed network, GTG, BMN, normalized Ensemble BMN, SBLN, and the normalized Ensemble SBLN are shown in Figure 4.11. A normalized ensemble CDF shows the CDF of the degree distribution of all runs divided by the number of runs. The SBLN and the normalized Ensemble BMN have the most similar distributions as the observed network. The normalized Ensemble SBLN performs better than the GTG and the BMN. In the same figure, the normalized Ensemble BMN and SBLN are plotted with two standard deviations above and below together with the observed network distribution. Here we see that there is a smaller standard deviation for the normalized BMN than the normalized SBLN. Even with the standard deviations, the degree distributions of both classes of networks are close to that of the observed degree distribution.

| | Density | Variance of Nodal Degree | Centrality |
|---|---|---|---|
| Observed | 0.16995 | 4.32105 | 0.20106 |
| SBLN | 0.16503 | 3.54578 | 0.16799 |
| Ensemble Average SBLN $\pm \sigma$ | 0.16355 $\pm$ 0.005593 | 3.66423 $\pm$ 0.48395 | 0.15040 $\pm$ 0.01883 |
| GTG | 0.16995 | 9.97622 | 0.27778 |
| BMN | 0.17488 | 3.88585 | 0.15741 |
| Ensemble Average BMN $\pm \sigma$ | 0.17579 $\pm$ 0.004546 | 3.93926 $\pm$ 0.41351 | 0.16065 $\pm$ 0.02635 |

Table 4.3: This table provides the shape measures for the observed network, SBLN, Ensemble SBLN, GTG, BMN, and Ensemble BMN. The $\sigma$ denotes the standard deviation of the ensemble metric values. Note that the density of the GTG is exactly the same as the observed rivalry network by construction.

### 4.4.5 Summary of Results

In all metrics except the density, the SBLN performs better than the GTG (note that the density measure of the GTG is exactly the same as the observed network by construction).

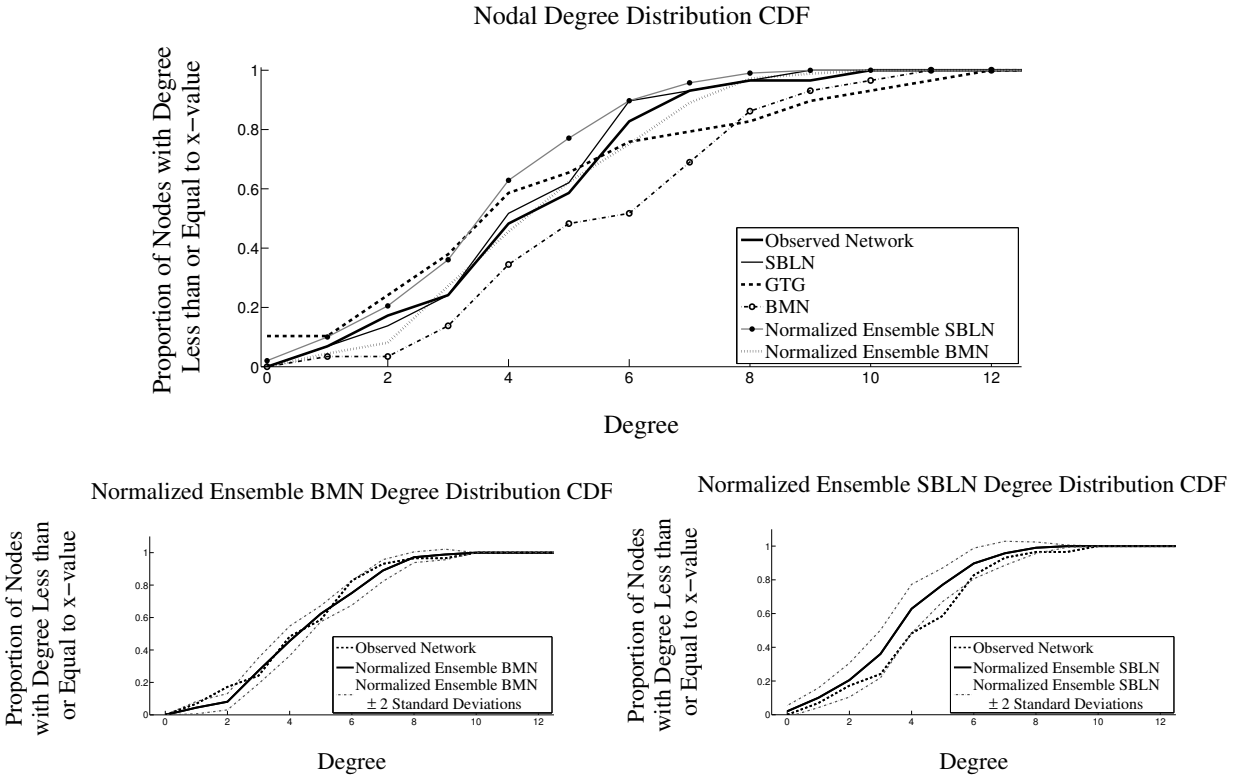


Figure 4.11: The top figure plots together the cumulative distribution functions of the degree distribution for the observed network (thick-solid), GTG (thick-dashed), BMN (dot-dash), normalized Ensemble BMN (thin-dash), SBLN (thin-solid), and normalized Ensemble SBLN (dot-solid). A normalized ensemble CDF shows the CDF of the degree distribution of all runs divided by the number of runs. The normalized Ensemble BMN (bottom left) and SBLN (bottom right) are plotted with two standard deviation above and below (thin-dash) with the observed network distribution (thick-dash).

Although the GTG is unable to closely replicate the standard shape measures, it has fairly high accuracy values. The Ensemble SBLN average performs similarly to the GTG in the accuracy, but performs better with shape measures, even with the stochastic considerations. On average, the Ensemble SBLN produces a slightly more accurate degree distribution than the GTG. The BMN is able to reproduce the degree distribution fairly well, however, the BMN and Ensemble BMN average have lower values for the accuracy (ACC), Matthews Correlation Coefficient (MCC), and F1 Score when compared to the other models. Our analysis demonstrates that the SBLN is the strongest model in reproducing the observed rivalry network.

## 4.5 Sensitivity Analysis

Our objective in this section is to understand the effects of the input parameters on the system by comparing the different metrics of the resulting networks as the parameters change. Due to computational constraints, we perform a local analysis of the parameter space around the SBLN parameters specified in column 3 of Table 4.1.

In particular, we perturb one parameter at a time by 30% from the SBLN parameter values in 10% increments. To account for the stochasticity inherent in the model, each perturbation was run using the same 25 seed values for the random number generator. The range of each parameter examined is listed in Table 4.4.

| | | | |
|---|---|---|---|
| Bounded Pareto Scaling Parameter | $k$ | $\in$ | [0.77 , 1.43] |
| Von Mises Scaling Parameter | $\kappa$ | $\in$ | [2.45 , 4.55] |
| Largest Maximum Jump Length | $A$ | $\in$ | [140 , 260] |
| Smallest Maximum Jump Length | $a$ | $\in$ | [70 , 130] |
| Boundary Permeability | $B$ | $\in$ | [0.14 , 0.26] |
| Network Threshold | $T$ | $\in$ | [0.028 , 0.052] |

Table 4.4: Ranges of the parameters used in the sensitivity analysis. Each parameter was changed 30% from the SBLN parameters in 10% increments. For SBLN parameter values refer to the Hollenbeck column of Table 4.1

For each simulation run, we compute the accuracy, Matthews Correlation Coefficient, F1 score, centrality measure, variance of nodal degree, and density for the resulting network. Plots of each combination of metric versus parameter values were created for the general analysis. Three examples of parameter and metric combinations with more dramatic results are plotted in Figure 4.12. In this figure, we display the variance of nodal degree versus the smallest maximum jump length, $a$, and the network threshold, $T$. We also display the density versus the Bounded Pareto scaling parameter, $k$, where the vertical axis has been rescaled for visualization. The dots represent the metric values of the simulation run at the specified parameter. The solid curve indicates the average metric value over all runs at each parameter value.

As seen in Figure 4.12, the plots varying the network threshold and Bounded Pareto

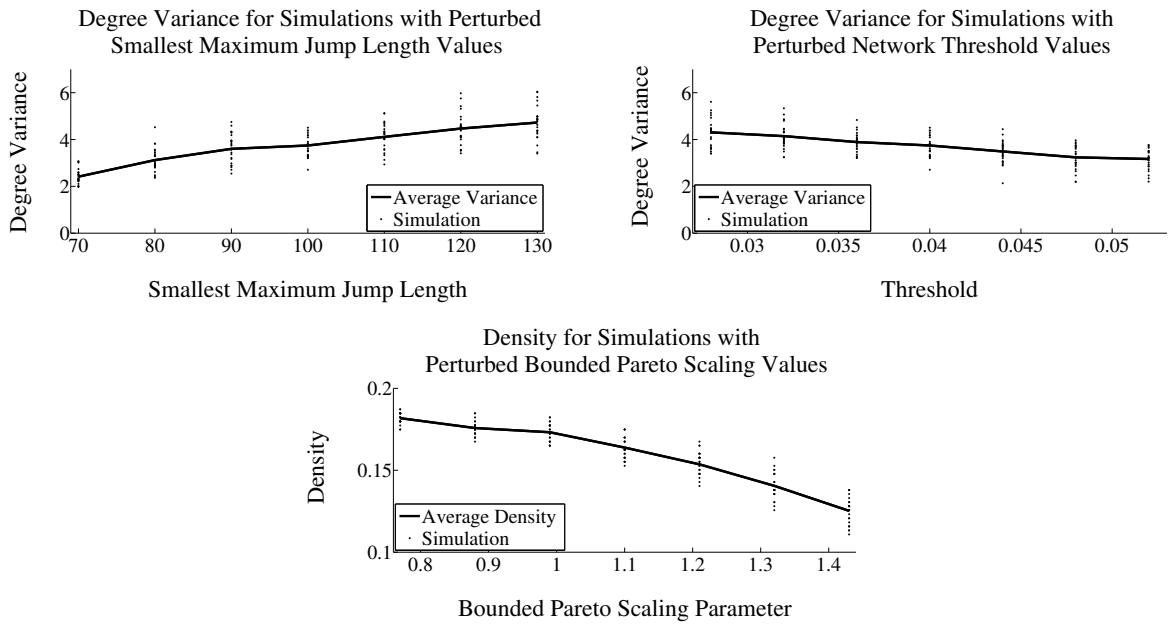


Figure 4.12: Plots of the nodal degree variance versus the smallest maximum jump length (top left), and the network threshold (top right). We also display the density versus the Bounded Pareto scaling parameter (bottom), where the vertical axis has been rescaled for visualization. The solid curve indicates the average metric value over all runs at each parameter value. The dots represent the metric values of the simulation run at the specified parameter.

scaling parameters have a negative trend on average. The smallest maximum jump length, however, shows a positive trend. The stochastic effects can also be observed by the range of metric values associated with each parameter input, as illustrated by the dots in Figure 4.12. These plots suggests that stochasticity may influence the metric values for a particular run, and on average the resulting metric output is sensitive with respect to these parameters.

These plots give a view of how the particular metric and parameter value interact. We changed all of the parameter values by the same 30% from the SBLN parameters, and so we can compare plots with the same metric. For example in Figure 4.12, we can see that in general nodal degree variance for the smallest maximum jump length has a steeper trend than the nodal degree variance for the threshold, but we can not compare the trend of the nodal degree variance plots directly to that of the density plot.

To compare the effects of all the parameters on all metrics, we rescale the data points

to percent deviation from the SBLN parameter values. For example, when considering the affects of the Bounded Pareto scaling parameter, $k$, on the density metric, we rescaled the observed data points

$$(k_i, \text{density}_i) \mapsto \left( \frac{k_i - k_{\text{ SBLN}}}{k_{\text{ SBLN}}} \ , \ \frac{\text{density}_i - \overline{\text{density}}_{\text{ SBLN}}}{\overline{\text{density}}_{\text{ SBLN}}} \right),$$

where $k_{\text{ SBLN}}$ is the SBLN Bounded Pareto scaling parameter. Here, $\overline{\text{density}}_{\text{ SBLN}}$ is the average density at the $k_{\text{ SBLN}}$ value for all 25 runs. A line was fitted to the rescaled data points, and the slope of this line was recorded. This process was repeated for each parameter and metric value combination.

The results are recorded in Table 4.5 and visualized in Figure 4.13. In Table 4.5, negative values indicate a negative slope of the best fit line to the scaled data, and positive values indicate a positive slope. Slopes with a greater magnitude indicate a stronger correlation between the metric and parameter. To get a clearer impression of overall sensitivity of the system, this information is displayed in Figure 4.13. The dark, and light, intensities of the color map represent large positive, and negative, values of the best fit line slope.

| | k | $\kappa$ | A | a | B | T |
|---|---|---|---|---|---|---|
| Accuracy | -0.0120 | -0.0031 | 0.0011 | 0.0001 | 0.0031 | 0.0000 |
| MCC | -0.2066 | -0.0161 | 0.0023 | 0.1458 | 0.0293 | -0.0000 |
| F1 Score | -0.2149 | -0.0146 | 0.0018 | 0.1562 | 0.0278 | -0.0000 |
| Centrality | -0.1705 | -0.0100 | -0.0131 | 0.7119 | 0.0195 | -0.1751 |
| Nodal Degree Variance | -0.4385 | -0.0489 | -0.0146 | **0**.9456 | 0.0154 | **-**0.5412 |
| Density | **-**0.7410 | -0.0080 | -0.0114 | 0.6131 | 0.0640 | -0.2460 |

Table 4.5: Slope of the best fit to the rescaled data for each metric and parameter combination. For reference, coefficients that correspond to the images in Figure 4.12 are highlighted in bold font. Figure 4.13 displays this information in a color map.

In general, the metrics are not very sensitive to the von Mises parameter, $\kappa$, the largest maximum jump length, $A$, and the boundary permeability, $B$, within the parameter space investigated. On the other hand, the Bounded Pareto scaling parameter, $k$, the smallest

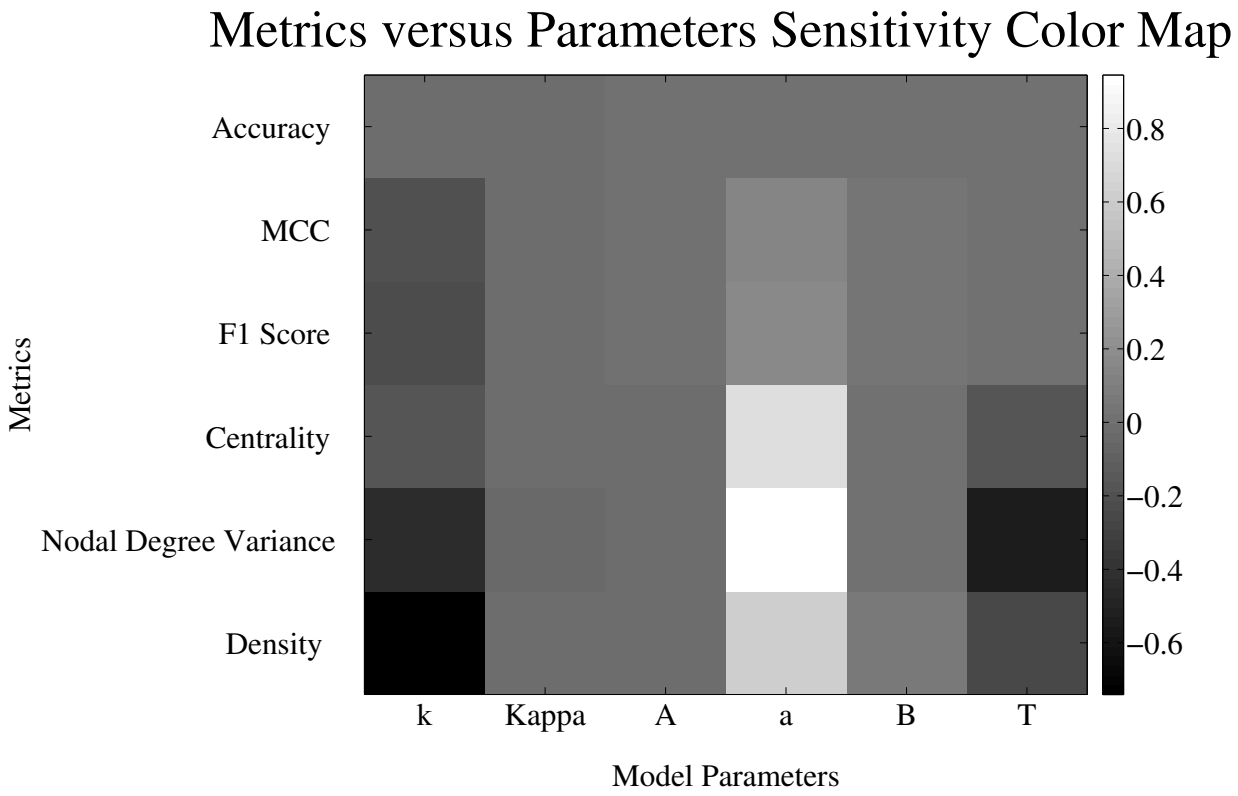


Figure 4.13: Slopes of the best fit line to the rescaled data for each of the parameter and metric combinations depicted in a color map. The parameters varied include the Bounded Pareto scaling parameter, $k$, the von Mises scaling parameter, Kappa, the largest maximum jump length, $A$, the smallest maximum jump length, $a$, the boundary permeability, $B$, and the network threshold, $T$. The scale to the right of the image gives the slope values. Tones close to the center of the scale represent combinations where the metrics are not very sensitive to the respective parameter. Combinations with tones at the ends of the spectrum (black and white) represent metrics that are sensitive to the respective parameter. The numerical values are also stored in Table 4.5.

maximum jump length, $a$, and the network threshold, $T$, have the most influence on the metrics. As seen in the table and figure, the accuracy measures are fairly robust to changes in all parameter values. Further, note that nodal degree variance and density measures appear to be the most affected by the changes in these parameters.

The Bounded Pareto scaling parameter values result in negative slopes for all metrics. This is to be expected because an increase in the Bounded Pareto scaling parameter will decrease the likelihood of larger jumps and result in fewer edges. This phenomena is particularly evident in the density metric. Also this parameter appears to have the most effect on

the accuracy measures, in particular the MCC and F1 score.

Increasing the network threshold parameter also has a negative effect on the shape metrics. By increasing the network threshold, the number of connections decreases. This in turn decreases the density, nodal degree variance, and centrality. On the other hand, increasing the smallest maximum jump length increases the connectivity of the network by allowing for larger jumps in areas of high road density. The effect of changing this parameter is more significant than changing the largest maximum jump length. Interestingly, as the largest maximum jump length increases, the connectivity decreases. This could be where attempts to cross boundaries are more likely to occur. The lower portion of Hollenbeck is approximately 300 units wide and has many boundaries. When varying the largest maximum jump length between 140 to 260, it becomes very probable that at least one boundary cross would be attempted. At this point, the boundary permeability is expected to play a stronger role in the simulation.

Depending on the network, changes in the number of connections could be more or less beneficial in terms of accuracy. Further, small changes in the connectivity, i.e. the existence or non-existence of an edge, could have small effects on the accuracy measures and large effects on the shape measures, as seen for our simulations in the case of the network threshold parameter.

## 4.6 Discussion

Using biased truncated Lévy walks with semi-permeable boundaries, we have designed an agent-based model for gang members in Hollenbeck that incorporates quasi-realistic movement rules as well as physical geographic features existing in Hollenbeck. We have shown that it is able to simulate a gang rivalry network similar to the one observed in [128, 102]. The Simulated Biased Lévy walk Network (SBLN), the Brownian Motion Network (BMN), and an instance of a Geographical Threshold Graph (GTG) were compared to the observed rivalry network using measures of accuracy and shape.

In choosing the metrics for the analysis of the models, we took into account that the

observed network may not represent the actual set of rivalries in a given region. If the true rivalry network was only a small perturbation from the observed network, then ideally, we would like the metrics used for comparison to only vary slightly. For this reason, we evaluated these models with the accuracy and shape metrics presented. The accuracy metrics provide a means to determine whether the simulations are recreating the same rivalries that are observed. The variance of nodal degree and centrality measures indicate the spread in degrees for each gang. These measures could be used to help policy makers know how to distribute resources. For example, if one gang has several rivals and the rest of the gangs only have one, then the resources should be mainly directed towards the one gang that is highly centralized within the network. Alternatively, if all gangs have the same number of rivals, then resources should be spread more equally through the area. The density metric gives an indication as to the connectivity of the network. In terms of gang rivalries, a higher density indicates a larger number of rivalries are present in the system. Creating a model that produces a network with similar shape metrics is desirable, since the connectivity and degree distributions could be useful for determining intervention strategies and methods for implementation.

The baseline models did not perform as well as the SBLN, but they did provide some insight into the modeling of gang rivalry networks. The GTG is a simple model designed to compare with the agent-based models. This method performs well on the accuracy metrics and provides an alternative, computationally inexpensive method to construct the rivalry network. One could extend this model to incorporate boundary information by increasing the distance function $d(n_i, n_j)$ if $n_i$ and $n_j$ are in distinct regions, see Section 4.3.1. The GTG model indicates that geography and size of gangs matter for the rivalry structure. However, the GTG is limited to reproducing only the rivalry network and does not lend itself to understanding other phenomena, such as the gang member mobility and the locations of interactions between gang members. It is not obvious how to extend the modeling framework of a GTG to include policing strategies, the location of violence, retaliatory behavior, and effects of injunctions, unlike agent-based models.

The BMN is a simplified version of the SBLN model. Although the BMN accuracy results

were not as strong as the GTG and SBLN results, this method was able to reproduce a similar shape as the observed rivalry network. This model was able to incorporate geographical features, but ignored directional decisions of the agents. The presence of semi-permeable boundaries in this model reduced the number of connections between regions, giving the simulated network a more similar shape to the observed. This suggests that the incorporation of boundaries plays an important role in the replication of the observed network. This is corroborated in [102]. In the absence of the directional decisions of the agents, we see a fundamental difference between the SBLN and the BMN. Without directional decisions the BMN has too many high degree nodes. One major problem with the BMN model is that the stopping criterion for the model was artificial, in that we chose to stop it at the observed peak in accuracy. In general, there may not be an observed network, and so it would be difficult to determine stopping criterion. Unlike the BMN, our proposed SBLN model exhibits long term stabilization of the accuracy and density metrics. This is a direct consequence of the directional decisions of the agents. The need for incorporating directional decisions into a mobility model is consistent with the literature [45].

The SBLN is the best model in replicating the observed network. Further, it allows for easy incorporation of geographic features and alternate movement dynamics, while maintaining a high level of accuracy and allowing for evolution in the observed system. This model is admittedly more complex than the other models, however, it provides better results in terms of accuracy and shape of the networks. The stochasticity in the model is beneficial not only to estimating the expected resulting network, but also potential networks that could arise from the seemingly random movements of individuals. One important feature is that this model produces stable long term behavior, meaning there is a reasonable point when to stop the number of iterations. This has practical consequences when the true rivalry network is not known. In the absence of an observed network, with which to calibrate the parameters of the model, various parameters could be initialized using data when available. There has been some work done on approximating the probability density function for jump lengths in a given region. Using similar methodology as [59] one could approximate the parameters associated with this region. In addition, an analysis of traffic flow under highways and within

regions could be used to approximate the boundary cross probability. This would greatly reduce the number of variables of the model.

This modeling framework allows for a variety of data output, depending on the questions of interest. For example, this model is able track the location of the agents' interactions during the simulation. This can be compared to violence data for the Hollenbeck area, and preliminary work has been done in this direction. Using this violence data would provide an additional data set to use for testing the external validity of our model. The issues of internal and external credibility are discussed in [13]. Figure 4.14 shows the locations of the interactions among agents for one of the Ensemble SBLN simulation runs and the density estimation of gang-related violent crimes in Hollenbeck from 1998 through 2000. The juxtaposition of these two plots emphasizes the similarities between the two and illustrates the potential predictive capabilities of this kind of approach. Though movement and interaction rules may need to be slightly altered to provide a closer match to the data, the current model provides a baseline model for further analysis and investigation of the gang rivalry violence in Hollenbeck.

This highly flexible model provides a framework with which to test sociological theories related to gang activities. One question of great interest to social scientists is the role of territories in the motivation for gang violence [128]. One could encode territories into the model by having each agent place a marker, associated with its gang, on the locations where it has been. By leaving these markings, one can see that territories would begin to form. Once these territories are established, the behavioral rules could be changed to avoid or attack the territories of gangs, instead of one point in space. From this, social scientists would be able to play out various scenarios and test hypotheses. Another interesting phenomenon observed is the presence of alliances or truce between gangs, as has been observed in Chicago [16] and Los Angeles [128, 85]. The current model does not account for the difference between positive, negative, and neutral interactions. Instead the SBLN records interactions between agents with the implicit assumption that these are negative interactions. A known truce between two gangs could be incorporated into the model by assigning negative values to the corresponding element of the gang rivalry matrix $R$. Then, all future interactions between

gangs with an alliance will not influence the movement decisions for agents of these gangs. Our model could also be used to examine this question.

Pursuing a model that accurately describes the violent behavior in Hollenbeck is of great value, since Hollenbeck is one of the most violent areas in Los Angeles [128, 56]. There are several advantages of approaching this serious problem using a mathematical model. One major strength over a network or statistical approach is that once the model has been sufficiently calibrated, it provides a powerful tool for social scientists to test theories and hypotheses. The paper [47] highlights that correlation does not imply causality, and though statistical analysis provides important contributions to social science as a field, it falls short in terms of identifying underlying mechanisms and testing causation hypotheses. Though one could argue that creating an experiment would help validate or reject causal arguments, such experiments are expensive and may bring up ethical concerns. Finally, if the simulation can accurately model the social phenomena of interest, then we might gain some insight into how intervention strategies could alter the existing gang rivalry system. The costs of implementing these changes in the simulation are relatively small compared to those costs of public funds needed to implement experimental interventions. If the Hollenbeck area can be well understood by this approach, there may be hope in understanding, and potentially mitigating, other areas of intense violent behavior.

## 4.7 Acknowledgments

We would like to thank Allon Percus for his useful feedback and Peterson Trethewey for his generous programming advice. We also would like to thank the LAPD, in particular Sean Malinowski, Gabriel Barboza, Arnold Suzukamo, Javier Macias, and the Hollenbeck Gang Division for their time and the enlightening conversations. In our work we used Google Earth$^{TM}$ images. This work was sponsored by NSF grant DMS-0907931, NSF grant DMS-0968309, AFOSR MURI grant FA9550-10-1-0569,ARO grant 58344-MA, ONR grant N000141010221, and ARO MURI grant 50363-MA-MUR.

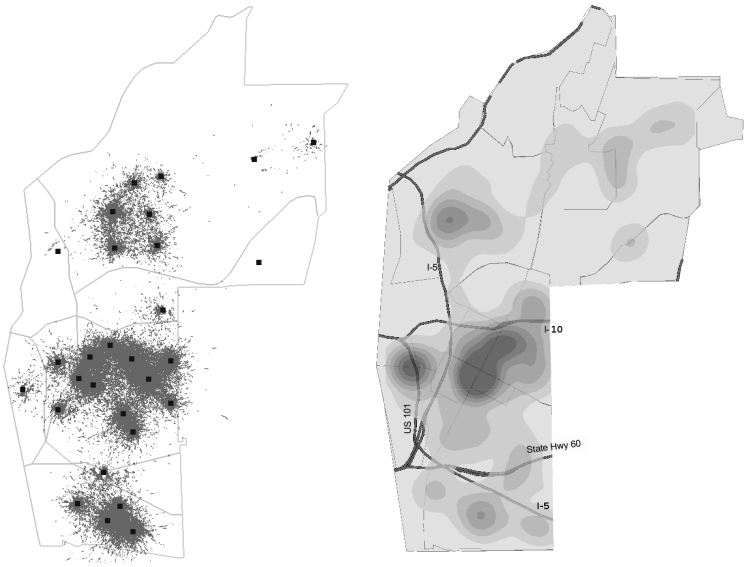


Figure 4.14: Locations of all the interactions between agents for one of the Ensemble SBLN runs (left). Density map of gang-related violent crimes in Hollenbeck between 1998 and 2000 (right).

# CHAPTER 5

# Adaptation of an Ecological Territorial Model to Street Gang Spatial Patterns in Los Angeles

Traditional gangs and territorial animals have distinctly claimed areas that are well-established [63]. They use markings to inform members of both their own group as well as other groups of territorial boundaries. Individuals may respond to markings of other groups by increasing their own marking density and by biasing their movement towards their respective den site for animals or set space for gangs. A gang's *set space* is the central location to the gang's activities [102, 126].

This chapter examines the formation of gang territories and territorial marking modifying the approach given in [81] that is based on a system of partial differential equations. The derivation of the PDEs is given in Section 5.1, and the adaptation for street gangs follows in Section 5.2. The implementation details are given in Section 5.3, and a description of the data sets are provided in Section 5.4. The results follow in Section 5.5. Section 5.6 concludes the paper with a discussion of the model implications.

## 5.1 Derivation of the Ecological Territorial Model

The particular model of interest in [81] is called, "steep terrain avoidance plus conspecific avoidance (STA+CA)." The supplement to [81] contains a derivation for the equations used, which will be roughly followed in this section with additional details.

We first consider the probability density function $u(\mathbf{x}, t)$ for an individual member's expected position at time $t$. Here, we have $\mathbf{x} = (x, y) \in \mathbb{R}^2$. Let $k_{\Delta t}(\mathbf{w}, \mathbf{x'})$ be the redistribution

kernel that gives the probability of moving from location **x'** with displacement **w** in a time step $\Delta t$. Thus,

$$\int_{\mathbb{R}^2} k_{\Delta t}(\mathbf{w}, \mathbf{x'})\ d\mathbf{w} = 1 \qquad \text{for all } \mathbf{x'}. \tag{5.1}$$

Then the density function $u$ at time $t + \Delta t$ can be described by the equation

$$u(\mathbf{x}, t + \Delta t) = \int_{\mathbb{R}^2} u(\mathbf{x'}, t)\, k_{\Delta t}(\mathbf{x} - \mathbf{x'}, \mathbf{x'})\ d\mathbf{x'}. \tag{5.2}$$

Using a change of variables with $\mathbf{q} = \mathbf{x} - \mathbf{x'} = (q_1, q_2)$, we have

$$u(\mathbf{x}, t + \Delta t) = \int_{\mathbb{R}^2} u(\mathbf{x} - \mathbf{q}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x} - \mathbf{q})\ d\mathbf{q}. \tag{5.3}$$

Using the Taylor expansion for the integrand about $\mathbf{x}$, we obtain

$$\begin{aligned} u(\mathbf{x}, t + \Delta t) = \int_{\mathbb{R}^2} [ & u(\mathbf{x}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x}) \\ & - (q_1) \frac{\partial}{\partial x} (u(\mathbf{x}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x})) \\ & - (q_2) \frac{\partial}{\partial y} (u(\mathbf{x}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x})) \\ & + \left(\frac{q_1^2}{2!}\right) \frac{\partial^2}{\partial x^2} (u(\mathbf{x}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x})) \\ & + \left(\frac{q_1 q_2}{2!}\right) \frac{\partial^2}{\partial x \partial y} (u(\mathbf{x}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x})) \\ & + \left(\frac{q_1 q_2}{2!}\right) \frac{\partial^2}{\partial y \partial x} (u(\mathbf{x}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x})) \\ & + \left(\frac{q_2^2}{2!}\right) \frac{\partial^2}{\partial y^2} (u(\mathbf{x}, t)\, k_{\Delta t}(\mathbf{q}, \mathbf{x})) \\ & + \ldots ]\ d\mathbf{q}. \end{aligned} \tag{5.4}$$

Rearraning the equation and changing the order of integration and differentiation, we find

$$\begin{aligned} u(\mathbf{x}, t+\Delta t) &= u(\mathbf{x}, t) \int_{\mathbb{R}^2} k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q} \\ &\quad -\frac{\partial}{\partial x}\left(u(\mathbf{x}, t)\ \int_{\mathbb{R}^2} q_1\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad -\frac{\partial}{\partial y}\left(u(\mathbf{x}, t)\ \int_{\mathbb{R}^2} q_2\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\left(\frac{1}{2!}\right)\frac{\partial^2}{\partial x^2}\left(u(\mathbf{x}, t)\ \int_{\mathbb{R}^2} \left(q_1^2\right)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\left(\frac{1}{2!}\right)\frac{\partial^2}{\partial x \partial y}\left(u(\mathbf{x}, t)\ \int_{\mathbb{R}^2} (q_1 q_2)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\left(\frac{1}{2!}\right)\frac{\partial^2}{\partial y \partial x}\left(u(\mathbf{x}, t)\ \int_{\mathbb{R}^2} (q_1 q_2)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\left(\frac{1}{2!}\right)\frac{\partial^2}{\partial y^2}\left(u(\mathbf{x}, t)\ \int_{\mathbb{R}^2} (q_2)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\ldots. \end{aligned} \tag{5.5}$$

To further analyze Equation (5.5), we examine each of the integrals and notice that they are the moments of $k_{\Delta t}(\mathbf{q}, \mathbf{x})$. Equation (5.1) determines that the first integral, $\int_{\mathbb{R}^2} k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}$, equals one. Since the probability of $\mathbf{q}$ being large is very small in a time step $\Delta t$, we assume that the first and second moments of $k_{\Delta t}$ dominate all of the higher moments as $\Delta t \to 0$. Thus, we do not consider the higher order terms. Rearranging the equation and dividing by $\Delta t$, we obtain

$$\begin{aligned} \left(\frac{u(\mathbf{x}, t+\Delta t) - u(\mathbf{x}, t)}{\Delta t}\right) &= -\frac{\partial}{\partial x}\left(u(\mathbf{x}, t)\ \frac{1}{\Delta t}\int_{\mathbb{R}^2} q_1\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad -\frac{\partial}{\partial y}\left(u(\mathbf{x}, t)\ \frac{1}{\Delta t}\int_{\mathbb{R}^2} q_2\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\frac{\partial^2}{\partial x^2}\left(u(\mathbf{x}, t)\ \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)\left(q_1^2\right)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\frac{\partial^2}{\partial x \partial y}\left(u(\mathbf{x}, t)\ \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)(q_1 q_2)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\frac{\partial^2}{\partial y \partial x}\left(u(\mathbf{x}, t)\ \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)(q_1 q_2)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right) \\ &\quad +\frac{\partial^2}{\partial y^2}\left(u(\mathbf{x}, t)\ \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)(q_2)\ k_{\Delta t}(\mathbf{q}, \mathbf{x})\ d\mathbf{q}\right). \end{aligned} \tag{5.6}$$

We now take the limit as $\Delta t \to 0$, and use the following definitions:

$$\begin{aligned}
\mathbf{c}(\mathbf{x},t) &= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\int_{\mathbb{R}^2} \mathbf{q}\; k_{\Delta t}(\mathbf{q},\mathbf{x})\; d\mathbf{q} \qquad (5.7)\\
d_{xx}(\mathbf{x},t) &= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)\left(q_1^2\right)\quad k_{\Delta t}(\mathbf{q},\mathbf{x})\; d\mathbf{q}\\
d_{xy}(\mathbf{x},t) &= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)(q_1 q_2)\quad k_{\Delta t}(\mathbf{q},\mathbf{x})\; d\mathbf{q}\\
d_{yx}(\mathbf{x},t) &= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)(q_1 q_2)\quad k_{\Delta t}(\mathbf{q},\mathbf{x})\; d\mathbf{q}\\
d_{yy}(\mathbf{x},t) &= \lim_{\Delta t\to 0} \frac{1}{\Delta t}\int_{\mathbb{R}^2} \left(\frac{1}{2!}\right)\left(q_2^2\right)\quad k_{\Delta t}(\mathbf{q},\mathbf{x})\; d\mathbf{q}.
\end{aligned}$$

Thus, Equation (5.8) in the limit becomes

$$\begin{aligned}
u_t(\mathbf{x},t) &= -\nabla\cdot(\mathbf{c}(\mathbf{x},t)\,u(\mathbf{x},t)) \qquad (5.8)\\
&\quad + \frac{\partial^2}{\partial x^2}\left(u(\mathbf{x},t)\,d_{xx}(\mathbf{x},t)\right) + \frac{\partial^2}{\partial x\partial y}\left(u(\mathbf{x},t)\,d_{xy}(\mathbf{x},t)\right)\\
&\quad + \frac{\partial^2}{\partial y\partial x}\left(u(\mathbf{x},t)\,d_{yx}(\mathbf{x},t)\right) + \frac{\partial^2}{\partial y^2}\left(u(\mathbf{x},t)\,d_{yy}(\mathbf{x},t)\right).
\end{aligned}$$

Assuming isotropic diffusion, this indicates that the mixed partial terms are equal to zero, and the terms $d_{xx}$ and $d_{yy}$ are equal. Letting $d := d_{xx} = d_{yy}$ and $d_{xy} = d_{yx} = 0$, we obtain the following differential equation,

$$u_t(\mathbf{x},t) = -\nabla\cdot(\mathbf{c}(\mathbf{x},t)\,u(\mathbf{x},t)) + \Delta\left(u(\mathbf{x},t)\,d(\mathbf{x},t)\right). \qquad (5.9)$$

### 5.1.1 Evaluating the Coefficients from the Redistribution Kernel

In Equation (5.9), the functions $\mathbf{c}(\mathbf{x},t)$ and $d(\mathbf{x},t)$ depend on the redistribution kernel, $k_{\Delta t}(\mathbf{q},\mathbf{x})$. The authors of [81] assume that this kernel depends on the jump length and the direction of the coyote's home den site and the change in elevation. They make the definitions listed in Table 5.1. To help visualize these definitions, Figure 5.1 illustrates these quanitities.

| Variable | Description |
|---|---|
| $\mathbf{x}$' | Starting Location |
| $\mathbf{x}$ | Ending Location |
| $\mathbf{x}_H$ | Home Den Site Location |
| $\phi$ | Angle of the Jump (angle of vector $\mathbf{x} - \mathbf{x}$') |
| $\phi_H$ | Angle of the Home Den Site from Ending Location (for vector $\mathbf{x}_H - \mathbf{x}$) |
| $\theta$ | $\phi - \phi_H$ |
| $\rho = \lvert\mathbf{x} - \mathbf{x}'\rvert$ | Jump Length = Distance from Starting Location to Ending Location |

Table 5.1: Variable definitions for starting location, ending location, jump length, and angles with respect to the jump and the home den site.

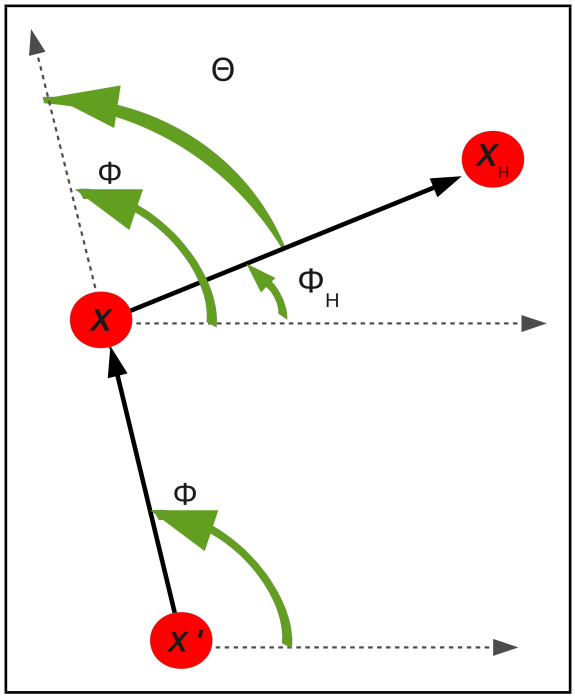


Figure 5.1: This illustrates the definitions of Table 5.1.

The authors of [81] assume the redistribution kernel takes the form

$$k_{\Delta t}\left(\mathbf{x} - \mathbf{x}', \mathbf{x}'\right) = \frac{1}{\rho}\, f_{\Delta t}\left(\rho\right) \cdot K_{\Delta T}\left(\phi - \phi_H\right), \tag{5.10}$$

giving the probability of starting at the location $\mathbf{x}$' and ending at $\mathbf{x}$. The term $f_{\Delta t}(\rho)$ is the probability density function for the jump length. The term $K_{\Delta t}(\phi - \phi_H) = K_{\Delta t}(\theta)$ gives

the probability density for the direction of movement. This was chosen to be the von Mises distribution,

$$K_{\Delta t}(\theta, \mathbf{x}, t) = \frac{1}{2\pi I_0(\kappa_{\Delta t})} \cdot \exp\left(\kappa_{\Delta t} \cos\theta\right). \tag{5.11}$$

In this equation, the $I_0$ is a modified Bessel function, and $\kappa_{\Delta t}$ gives the concentration of the bias about $\theta$. If $\kappa_{\Delta t} = 0$, then this distribution is a uniform distribution on a circle.

We now rewrite Equation (5.7) in terms of the redistribution kernel of Equation (5.10) using simplified notation (excluding $t$ and $\mathbf{x}$ from the function arguments), giving

$$\mathbf{c} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_0^{2\pi} \int_0^{\infty} (\mathbf{x} - \mathbf{x'}) \; f_{\Delta t}(\rho) \; K_{\Delta t}(\theta) \; d\rho \; d\theta \tag{5.12}$$

$$d = \lim_{\Delta t \to 0} \frac{1}{2\Delta t} \int_0^{2\pi} \int_0^{\infty} (x - x')^2 \; f_{\Delta t}(\rho) \; K_{\Delta t}(\theta) \; d\rho \; d\theta. \tag{5.13}$$

Consider the vector

$$\overrightarrow{\mathbf{x}} = \frac{\mathbf{x}_H - \mathbf{x}}{|\mathbf{x}_H - \mathbf{x}|} = (\cos\phi_H, \sin\phi_H)^T,$$

the vector that points in the direction of the den site from the ending location. We take the inner product of $\mathbf{c}$ and $\overrightarrow{\mathbf{x}}$ to give

$$\begin{aligned}
\mathbf{c} \cdot \overrightarrow{\mathbf{x}} &= \left( \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_0^{2\pi} \int_0^{\infty} (\mathbf{x} - \mathbf{x'}) \; f_{\Delta t}(\rho) \; K_{\Delta t}(\theta) \; d\rho \; d\theta \right) \cdot (\cos\phi_H, \sin\phi_H) \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_0^{2\pi} \int_0^{\infty} \rho \left(\cos\phi, \sin\phi\right) \cdot (\cos\phi_H, \sin\phi_H) \, f_{\Delta t}(\rho) \, K_{\Delta t}(\theta) \; d\rho \; d\theta \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_0^{2\pi} \int_0^{\infty} \rho \left(\cos\phi \cdot \cos\phi_H + \sin\phi \cdot \sin\phi_H\right) f_{\Delta t}(\rho) \, K_{\Delta t}(\theta) \; d\rho \; d\theta \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_0^{2\pi} \int_0^{\infty} \rho \cos\left(\phi_H - \phi\right) f_{\Delta t}(\rho) \, K_{\Delta t}(\theta) \; d\rho \; d\theta \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_0^{2\pi} \int_0^{\infty} \rho \cos(\theta) \, f_{\Delta t}(\rho) \, K_{\Delta t}(\theta) \; d\rho \; d\theta \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_0^{2\pi} \cos(\theta) \, K_{\Delta t}(\theta) \left[ \int_0^{\infty} \rho f_{\Delta t}(\rho) \; d\rho \right] d\theta \\
\mathbf{c} \cdot \overrightarrow{\mathbf{x}} &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[ \int_0^{2\pi} \cos(\theta) \, K_{\Delta t}(\theta) \, d\theta \right] \left[ \int_0^{\infty} \rho f_{\Delta t}(\rho) \; d\rho \right].
\end{aligned} \tag{5.14}$$

We now consider a vector orthogonal to $\overrightarrow{\mathbf{x}}$,which will be denoted as

$$\mathbf{x}^{\perp} = \left(\cos\left(\phi_H - \frac{\pi}{2}\right), \sin\left(\phi_H - \frac{\pi}{2}\right)\right)^T.$$

Taking the inner product of $\mathbf{c}$ and $\mathbf{x}^{\perp}$ and using a similar argument gives

$$\mathbf{c} \cdot \mathbf{x}^{\perp} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left[\int_0^{2\pi} \sin(\theta)\, K_{\Delta t}(\theta)\, d\theta\right] \left[\int_0^{\infty} \rho f_{\Delta t}(\rho)\ d\rho\right]. \tag{5.15}$$

Since $K_{\Delta t}(\theta)$ is even and $\sin(\theta)$ is odd, both with period $2\pi$, we have that

$$\mathbf{c} \cdot \mathbf{x}^{\perp} = 0.$$

We note that $\overrightarrow{\mathbf{x}}$ and $\mathbf{x}^{\perp}$ form an orthogonal basis for $\mathbb{R}^2$. Then using Equations (5.14) and (5.15), we have

$$\mathbf{c} = |\mathbf{c}|\overrightarrow{\mathbf{x}}, \qquad\qquad |\mathbf{c}| = \mathbf{c} \cdot \overrightarrow{\mathbf{x}}.$$

We now do a similar analysis for $d$, noting that $(x - x')^2 = \rho^2 \cos^2\phi$. We make this substitution into Equation (5.13) and use several trigonometric identities to yield

$$\begin{aligned}
d &= \lim_{\Delta t \to 0} \frac{1}{2\Delta t} \int_0^{2\pi} \int_0^{\infty} \rho^2 \cos^2\phi\ f_{\Delta t}(\rho)\ K_{\Delta t}(\theta)\ d\rho\, d\theta \\
&= \lim_{\Delta t \to 0} \frac{1}{2\Delta t} \int_0^{2\pi} \int_0^{\infty} \frac{\rho^2}{2}(1 + \cos(2\phi))\ f_{\Delta t}(\rho)\ K_{\Delta t}(\theta)\ d\rho\, d\theta \\
&= \lim_{\Delta t \to 0} \frac{1}{4\Delta t} \int_0^{2\pi} (1 + \cos(2\phi))\ K_{\Delta t}(\theta) \left[\int_0^{\infty} \rho^2\ f_{\Delta t}(\rho)\ d\rho\right] d\theta \\
&= \lim_{\Delta t \to 0} \frac{1}{4\Delta t} \left[\int_0^{2\pi} K_{\Delta t}(\theta)\, d\theta + \int_0^{2\pi} \cos(2\phi)\ K_{\Delta t}(\theta)\, d\theta\right] \left[\int_0^{\infty} \rho^2\ f_{\Delta t}(\rho)\ d\rho\right] \\
&= \lim_{\Delta t \to 0} \frac{1}{4\Delta t} \left[1\ + \int_0^{2\pi} \cos(2\theta + 2\phi_H)\ K_{\Delta t}(\theta)\, d\theta\right] \left[\int_0^{\infty} \rho^2\ f_{\Delta t}(\rho)\ d\rho\right]
\end{aligned} \tag{5.16}$$

Looking only at the first integral, we use multiple identites for sine and cosine to simplify

the equations, giving

$$\begin{aligned}
\int_0^{2\pi} \cos(2\theta \;+\; 2\phi_H)\; K_{\Delta t}(\theta)\, d\theta & \\
&= \int_0^{2\pi} \left[\cos(2\theta)\cos(2\phi_H) - \sin(2\theta)\sin(2\phi_H)\right] K_{\Delta t}(\theta)\, d\theta \\
&= \cos(2\phi_H)\int_0^{2\pi} \cos(2\theta)\;\; K_{\Delta t}(\theta)\, d\theta - \sin(2\phi_H)\int_0^{2\pi} \sin(2\theta)\;\; K_{\Delta t}(\theta)\, d\theta \\
&= \cos(2\phi_H)\int_0^{2\pi} \cos(2\theta)\;\; K_{\Delta t}(\theta)\, d\theta - \sin(2\phi_H)\int_0^{2\pi} 2\sin\theta\cos\theta\;\; K_{\Delta t}(\theta)\, d\theta \\
&= \cos(2\phi_H)\int_0^{2\pi} \cos(2\theta)\;\; K_{\Delta t}(\theta)\, d\theta + 0.
\end{aligned}$$

The last equality comes from the fact that $\sin(\theta)$ is odd, $K_{\Delta t}$ and $\cos(\theta)$ are even, and all have a period of $2\pi$. Putting this together with Equation (5.16), we obtain

$$d \;=\; \lim_{\Delta t \to 0}\; \frac{1}{4\Delta t}\left[1 \;+ \cos(2\phi_H)\int_0^{2\pi} \cos(2\theta)\;\; K_{\Delta t}(\theta)\, d\theta\right]\left[\int_0^{\infty} \rho^2\;\, f_{\Delta t}(\rho)\;\; d\rho\right]. \quad (5.17)$$

The authors of [81] additionally assume that $f_{\Delta t}(\rho)$ and $\kappa_{\Delta t}(\mathbf{x}, t)$ go to zero as $\Delta t \to 0$. We denote the first moment of $f_{\Delta t}$ as

$$\bar{\rho_{\Delta t}} = \int_0^{\infty} \rho f_{\Delta t}(\rho)\; d\rho.$$

The $\bar{\rho_{\Delta t}}$ represents the average jump length in the time interval $\Delta t$. We also define

$$\bar{\rho^2_{\Delta t}} = \int_0^{\infty} \rho^2 f_{\Delta t}(\rho)\; d\rho.$$

Using an identity for the von Mises distribution and an expansion for Bessel functions, we will obtain $\mathbf{c}$ and $d$ [57]. More specifically, we have the identity

$$\int_0^{2\pi} \cos(n\theta) K_{\Delta t}(\theta)\; d\theta = \frac{I_n(\kappa_{\Delta t})}{I_0(\kappa_{\Delta t})}. \quad (5.18)$$

Using the Bessel function series expansion for $I_0(\kappa)$ and $I_1(\kappa)$, we obtain

$$\begin{aligned}
|\mathbf{c}| &= \mathbf{c}\cdot\overrightarrow{\mathbf{x}} \\
&= \lim_{\Delta t\to 0}\ \frac{1}{\Delta t}\left[\int_0^{2\pi}\cos(\theta)\,K_{\Delta t}(\theta)\,d\theta\right]\left[\int_0^{\infty}\rho f_{\Delta t}(\rho)\ \ d\rho\right] \\
&= \lim_{\Delta t\to 0}\ \frac{1}{\Delta t}\left[\frac{I_1(\kappa_{\Delta t})}{I_0(\kappa_{\Delta t})}\right]\left[\bar{\rho_{\Delta t}}\right] \\
&= \lim_{\Delta t\to 0}\ \frac{1}{\Delta t}\left[\frac{\frac{1}{2}\kappa_{\Delta t}+O\left(\kappa_{\Delta t}^2\right)}{1+\frac{1}{4}\kappa_{\Delta t}^2+O\left(\kappa_{\Delta t}^3\right)}\right]\left[\bar{\rho_{\Delta t}}\right].
\end{aligned}$$

Keeping only the $\kappa_{\Delta t}$ of leading order, we have

$$|\mathbf{c}| = \lim_{\Delta t\to 0}\ \frac{1}{\Delta t}\left[\frac{1}{2}\kappa_{\Delta t}\right]\left[\bar{\rho_{\Delta t}}\right] = \lim_{\Delta t\to 0}\left(\frac{\kappa_{\Delta t}\ \bar{\rho_{\Delta t}}}{2\Delta t}\right) \tag{5.19}$$

$$\mathbf{c} = |\mathbf{c}|\overrightarrow{\mathbf{x}} = \lim_{\Delta t\to 0}\left(\frac{\bar{\rho_{\Delta t}}}{2\Delta t}\right)(\kappa_{\Delta t})\,\overrightarrow{\mathbf{x}} \tag{5.20}$$

For $d$, a similar derivation is used with the Bessel function series expansions for both $I_2$ and $I_0$.

$$\begin{aligned}
d &= \lim_{\Delta t\to 0}\ \frac{1}{4\Delta t}\left[1\ +\cos(2\phi_H)\left(\frac{I_2(\kappa_{\Delta t})}{I_0(\kappa_{\Delta t})}\right)\right]\left[\bar{\rho^2_{\Delta t}}\right] \\
&= \lim_{\Delta t\to 0}\ \frac{\bar{\rho^2_{\Delta t}}}{4\Delta t}\left[1\ +\cos(2\phi_H)\left(\frac{\frac{1}{8}\kappa_{\Delta t}^2+O\left(\kappa_{\Delta t}^3\right)}{1+\frac{1}{4}\kappa_{\Delta t}^2+O\left(\kappa_{\Delta t}^3\right)}\right)\right] \\
&= \lim_{\Delta t\to 0}\ \frac{\bar{\rho^2_{\Delta t}}}{4\Delta t}\left[1+0\right] \\
d &= \lim_{\Delta t\to 0}\left(\frac{\bar{\rho^2_{\Delta t}}}{4\Delta t}\right).
\end{aligned} \tag{5.21}$$

These formulas are used in the following two subsections to derive the $\mathbf{c}$ and $d$ for the Conspecific Avoidance (CA) model and the Steep Terrain Avoidance plus Conspecific Avoidance (STA+CA) Model [81]. The latter of these is the one that will be used for the street gang model.

### 5.1.2 Conspecific Avoidance (CA) Model

For specific models, the angle of bias $\kappa_{\Delta t}$ needs to be determined. For the Conspecific Avoidance (CA) approach of [81], the choice for $\kappa_{\Delta t}$ was chosen to depend on the density of the scent marks for different packs. If the individual is from pack $i$, then we take

$$\kappa_{\Delta t}(\mathbf{x}, t) = b\bar{\rho}_{\Delta t} \sum_{i \neq j}^{n} p^{(j)}(\mathbf{x}, t). \tag{5.22}$$

The $p^{(j)}(\mathbf{x}, t)$ is the marking density for pack $j$, and $n$ is the number of packs of coyotes. Additionally, $b$ gives the strength of the reaction to the scent markings. Thus, if the individual encounters a large amount of scent markings from different packs, it will have a stronger desire to head back towards the den site. If no markings are present, then the individual will have an equal probability of moving in any direction.

Using Equation (5.28) in the formula for $\mathbf{c}$ of Equation (5.20), we obtain

$$\begin{aligned} \mathbf{c} &= \lim_{\Delta t \to 0} \left( \frac{\bar{\rho}_{\Delta t}}{2\Delta t} \right) \left( b\bar{\rho}_{\Delta t} \sum_{i \neq j}^{n} p^{(j)}(\mathbf{x}, t) \right) \overrightarrow{\mathbf{x}} \\ &= \lim_{\Delta t \to 0} \left( \frac{b\bar{\rho}_{\Delta t}^{\,2}}{2\Delta t} \right) \left( \overrightarrow{\mathbf{x}} \sum_{i \neq j}^{n} p^{(j)}(\mathbf{x}, t) \right). \end{aligned} \tag{5.23}$$

Thus, the partial differential equation that models the pack density for pack $i$ is given by

$$u_t^{(i)}(\mathbf{x}, t) = -\nabla \cdot \left( \gamma u^{(i)}(\mathbf{x}, t) \left( \overrightarrow{\mathbf{x}} \sum_{i \neq j}^{n} p^{(j)}(\mathbf{x}, t) \right) \right) + d\Delta u^{(i)}(\mathbf{x}, t) \tag{5.24}$$

$$\gamma = \lim_{\Delta t \to 0} \left( \frac{b\bar{\rho}_{\Delta t}^{\,2}}{2\Delta t} \right). \tag{5.25}$$

Next, we need to consider how the scent marking density changes over time. We first assume that coyotes will produce markings even in the absense of other packs' markings, and will do so with a rate $l$. Additionally, when coming in contact with another pack's markings, an individual will increase the rate of markings proportionally to the amount encountered. We let $m$ be the parameter that gives the strength for the increase. We also assume that

the markings will decay over time with a rate of $\mu$. Then the model proposed in [81] gives the equation for the changing of the marking densities over time as

$$p_t^{(i)}(\mathbf{x},t) \;=\; Nu^{(i)}(\mathbf{x},t)\left[l + m\sum_{i\neq j}^{n} p^{(j)}(\mathbf{x},t)\right] - \mu p^{(i)}(\mathbf{x},t)\,. \tag{5.26}$$

The Equations (5.30) and (5.26) are then solved to steady-state to obtain the expected densities for the coyote packs under the conspecific avoidance model.

### 5.1.3 Steep Terrain Avoidance Plus Conspecific Avoidance (STA+CA) Model

In addition to avoiding the scent markings of other coyote packs, we wish to consider the avoidance of steep terrain. To incorporate this into the model, we use a weighted sum of two von Mises distributions for $K_{\Delta t}(\theta)$ in the redistribution kernel to determine the angular direction of motion. More specifically, [81] choose

$$k_{\Delta t}(\mathbf{x}-\mathbf{x'},\mathbf{x'}) = \frac{1}{\rho}\, f_{\Delta t}(\rho)\cdot\left[\psi K_{\Delta T}^{CA}(\phi-\phi_H,\mathbf{x},t) + (1-\psi)K_{\Delta T}^{STA}(\phi,-\phi_z,\mathbf{x})\right], \tag{5.27}$$

where $K_{\Delta t}^{CA}$ denotes the von Mises distribution for the conspecific avoidance and $K_{\Delta t}^{STA}$ for the steep terrain avoidance. The strength of bias for $K_{\Delta t}^{CA}$ remains unchanged from the (CA) model as

$$\kappa_{\Delta t}^{CA}(\mathbf{x},t) = b\rho_{\Delta t}^{-}\sum_{i\neq j}^{n} p^{(j)}(\mathbf{x},t). \tag{5.28}$$

The direction of bias for $K_{\Delta t}^{STA}$ depends on the elevation of the area, given by $z(\mathbf{x})$. We are concerned about regions where this changes quickly. Thus, the strength of the directional bias depends on $|\nabla z|$. Therefore, we take

$$\kappa_{\Delta t}^{STA}(\mathbf{x}) = a\rho_{\Delta t}^{-}|\nabla z(\mathbf{x})|. \tag{5.29}$$

Modifying Equation (5.12) to account for this new redistribution kernel, we have

$$\begin{aligned}
\mathbf{c} &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_0^{2\pi}\int_0^{\infty}(\mathbf{x}-\mathbf{x'})\,f_{\Delta t}(\rho)\left[\psi K_{\Delta T}^{CA}(\theta)+(1-\psi)K_{\Delta T}^{STA}(\phi,-\phi_z)\right]d\rho d\theta \\
&= \lim_{\Delta t\to 0}\frac{\psi}{\Delta t}\int_0^{2\pi}\int_0^{\infty}(\mathbf{x}-\mathbf{x'})\,f_{\Delta t}(\rho)\,K_{\Delta T}^{CA}(\theta)d\rho d\theta \\
&\quad + \lim_{\Delta t\to 0}\frac{1-\psi}{\Delta t}\int_0^{2\pi}\int_0^{\infty}(\mathbf{x}-\mathbf{x'})\,f_{\Delta t}(\rho)\,K_{\Delta T}^{STA}(\phi,-\phi_z)d\rho d\theta.
\end{aligned}$$

In the first integral, the $K_{\Delta T}^{CA}(\theta)$ has the same even properties as $K_{\Delta T}(\theta)$ from the previous sections. Thus, we can use the same arguments for the advection term that avoids scent-markings of other packs. The second integral, however, contains $K_{\Delta T}^{STA}(\phi,-\phi_z)$, which has a direction of bias away from areas of steep terrain. Thus, instead of taking the inner product with the vector pointing towards the home den site $\overrightarrow{\mathbf{x}}$, we use $\frac{\nabla z}{|\nabla z|}$. Repeating a similar argument, we obtain the model

$$u_t^{(i)} = d\Delta u^{(i)} - \nabla\cdot\left(\gamma\psi u^{(i)}\left(\overrightarrow{\mathbf{x}}\sum_{i\neq j}^{n}p^{(j)}(\mathbf{x},t)\right)\right)+\nabla\cdot\left((1-\psi)\eta u^{(i)}\nabla z\right). \quad (5.30)$$

We use Equation (5.26) for the time evolution of the marking densities. The authors of [81] non-dimensionalize these equations, which are displayed in the following section. It is this system of equations that will be used to model street gangs.

## 5.2 The Proposed Gang Behavioral Model

We will use the "steep terrain plus conspecific avoidance (STA+CA)" model of [81], derived in Section 5.1.3. This approach solves the following non-dimensionalized system to steady-state:

$$\frac{du^{(i)}}{dt} = \Delta u^{(i)} - \nabla\cdot\left[\beta\overrightarrow{\mathbf{x}_i}u^{(i)}\sum_{j\neq i}^{n}p^{(j)}\right]+\nabla\cdot\left[\alpha u^{(i)}\nabla z\right] \quad (5.31)$$

$$\frac{dp^{(i)}}{dt} = u^{(i)}\left[1+m\sum_{j\neq i}^{n}p^{(j)}\right]-p^{(i)}, \quad (5.32)$$

where for a given pack $i$, $u^{(i)}(x,y)$ is the expected population density, $p^{(i)}(x,y)$ is the expected marking density, and $\overrightarrow{\mathbf{x}_i}$ is the unit vector to the home den site from the current location. The parameters $\beta$, $m$, and $\alpha$ represent the strengths to avoid markings, to increase one's own markings, and to avoid steep terrain, respectively. This system also has the boundary conditions on $\partial\Omega$

$$0 = \left[\nabla u^{(i)} - \beta\overrightarrow{\mathbf{x}_i}u^{(i)}\sum_{j\neq i}^{n} p^{(j)} + \alpha u^{(i)}\nabla z\right]\cdot\overrightarrow{\mathbf{n}}, \tag{5.33}$$

where $\overrightarrow{\mathbf{n}}$ is the outward unit normal vector.

The system in Equations (5.31) and (5.32) models population densities and marking densities for a set of coyote packs. We now modify the approach for street gangs by letting $u^{(i)}(x,y)$ be the expected gang density, $p^{(i)}(x,y)$ be the expected graffiti density, and $\overrightarrow{\mathbf{x}_i}$ be the unit vector to the set space of gang $i$ from the current location. In this system, the term $\Delta u^{(i)}$ describes random motion of individuals. Since gangs typically avoid the territories of other gangs [5, 69] unless they are conducting a violent raid, they are more likely to return towards their set space when confronted by another gang's tags. This behavior is represented by the term $-\nabla\cdot\left[\beta\overrightarrow{\mathbf{x}_i}u^{(i)}\sum_{j\neq i}^{n}p^{(j)}\right]$, where individuals that come in contact with any markings that are not of their own gang move in the direction towards home, $\overrightarrow{\mathbf{x}_i}$. Equation (5.32) describes the process by which individuals increase their production of markings after discovering another gang's markings.

The $z(x,y)$ in the model is the elevation of the terrain for coyotes, but we use this term to incorporate other spatial features. We first determine geographical landmarks that could inhibit movement across them, such as rivers, freeways, and major roads. These features are not impassible, but there are limited bridges and underpasses available. Let $\Gamma\subset\mathbb{R}^2$ be the set of points where there is a landmark. Consider the minimum Euclidean distance to the set $\Gamma$,

$$g(x,y) = \min_{(x_0,y_0)\in\Gamma}\left\{\sqrt{(x-x_0)^2+(y-y_0)^2}\right\}.$$

We then define $z(x, y)$ to be

$$z(x, y) = 1 - \tanh\left(g(x, y)^2\right).$$

This gives a "steep terrain" that gang members will avoid. Geography is therefore incorporated into the movement dynamics through the final term in Equation (5.31), $\nabla \cdot \left[\alpha u^{(i)} \nabla z\right]$, where each landmark is treated equally.

This partial differential equation approach has many similar features to the agent-based model proposed in Chapter 4. Both models incorporate the geography as semi-permeable boundaries. Additionally, individuals engage in random movement, but have a bias towards their home set space. As seen in Section 5.1, the authors of [81] derive the model given in Equations (5.31) and (5.32) by using a von Mises distribution for the advection term. Thus, both models utilize the von Mises distribution to generate directionally biased motion. One major difference between the two methods is that the agent-based approach is stochastic, whereas the proposed model is deterministic with a derivation based on a stochastic process. Due to the similarities in these models, we will also use Hollenbeck as the case study and compare the results to those given in Chapter 4. We choose the same geographical features used in the agent-based approach and is given in the middle image of Figure 4.3.

## 5.3 Implementation Details

For the $G$ gangs, the model consists of a coupled system with $2G$ differential equations. We implemented this system by using finite differences and solving until steady-state. For each time step $\Delta t$, we loop through all of the gangs, updating first the $u^{(k)}$ and then $p^{(k)}$. For initialization, $u^{(k)}$ is set to be zero everywhere except at the location of the $k^{th}$ set space, where it has the value one. Alternatively, $p^{(k)}$ is initialized to be $\frac{1}{NM}$ everywhere, where the

grid is $N \times M$. The discretized equation for the interior of $u^{(k)}$ at time $n+1$ is given by

$$\begin{aligned}
\frac{u_{i,j}^{(k),n+1} - u_{i,j}^{(k),n}}{\Delta t} &= \frac{1}{(\Delta x)^2}\left(u_{i+1,j}^{(k),n} - 2u_{i,j}^{(k),n} + u_{i-1,j}^{(k),n}\right) \\
&+ \frac{1}{(\Delta y)^2}\left(u_{i,j+1}^{(k),n} - 2u_{i,j}^{(k),n} + u_{i,j-1}^{(k),n}\right) \\
&- \frac{\beta}{2\Delta x}\left(u_{i+1,j}^{(k),n}\ x_{i+1,j}^{(k)} \sum_{l\neq k} p_{i+1,j}^{(l),n} - u_{i-1,j}^{(k),n}\ x_{i-1,j}^{(k)} \sum_{l\neq k} p_{i-1,j}^{(l),n}\right) \\
&- \frac{\beta}{2\Delta y}\left(u_{i,j+1}^{(k),n}\ y_{i,j+1}^{(k)} \sum_{l\neq k} p_{i,j+1}^{(l),n} - u_{i,j-1}^{(k),n}\ y_{i,j-1}^{(k)} \sum_{l\neq k} p_{i,j-1}^{(l),n}\right) \\
&+ \frac{\alpha}{\Delta x}\left(u_{i+1,j}^{(k),n}\ \nabla_x\ z_{i+1,j} - u_{i,j}^{(k),n}\ \nabla_x\ z_{i,j}\right) \\
&+ \frac{\alpha}{\Delta y}\left(u_{i,j+1}^{(k),n}\ \nabla_y\ z_{i,j+1} - u_{i,j}^{(k),n}\ \nabla_y\ z_{i,j}\right).
\end{aligned}$$

The boundary of $u^{(k)}$ is updated using first order Neumann boundary conditions. Given this update of $u^{(k)}$, we then update both the interior and boundary of $p^{(k)}$ as

$$\frac{p_{i,j}^{(k),n+1} - p_{i,j}^{(k),n}}{\Delta t} = u_{i,j}^{(k),n+1}\left(1 + m\ \sum_{l\neq k} p_{i,j}^{(l),n}\right) - p_{i,j}^{(k),n+1}.$$

After this update, we ensure that $p^{(k)}$ is still a density that sums to one. After obtaining our final estimates for the gang and marking densities, we determine that a rivalry exists between two gangs if the regions where densities are non-negligible sufficiently overlap.

## 5.4 The Data Sets

In order to test the algorithm, we will use two different data sets. The first gives the locations of violence between two gangs from 1999 to 2002. We only use the 340 events where both the suspect and victim gang are known and both are also in the set of twenty-nine Hollenbeck gangs. In [20], the authors compared the locations of events relative to the theoretical midline boundary between gangs. They concluded that violent interactions frequently cluster about this line. From this analysis, we should expect a higher density of violent events near the boundaries of territories where there is an increase in graffiti. For our model, this is where

the markings, $p^{(i)}$, have a high density. Thus, we will compare our marking densities to this set of *violence data*.

The second set of data we have to evaluate our model gives the locations of where known or suspected gang members were stopped by the police. In such encounters, the officer fills out a Field Interview (FI) card, listing information about where the interaction took place and all individuals involved. These encounters are not initiated by violent events, and so we assume that these data are a sample of the gang densities, $u^{(i)}$. This data set of 1,079 events, covering the year of 2009, will be referred to as *FI data*. Both data sets can be seen in Figure 5.2. Due to the large number of gangs, only the aggregated data will be displayed here.

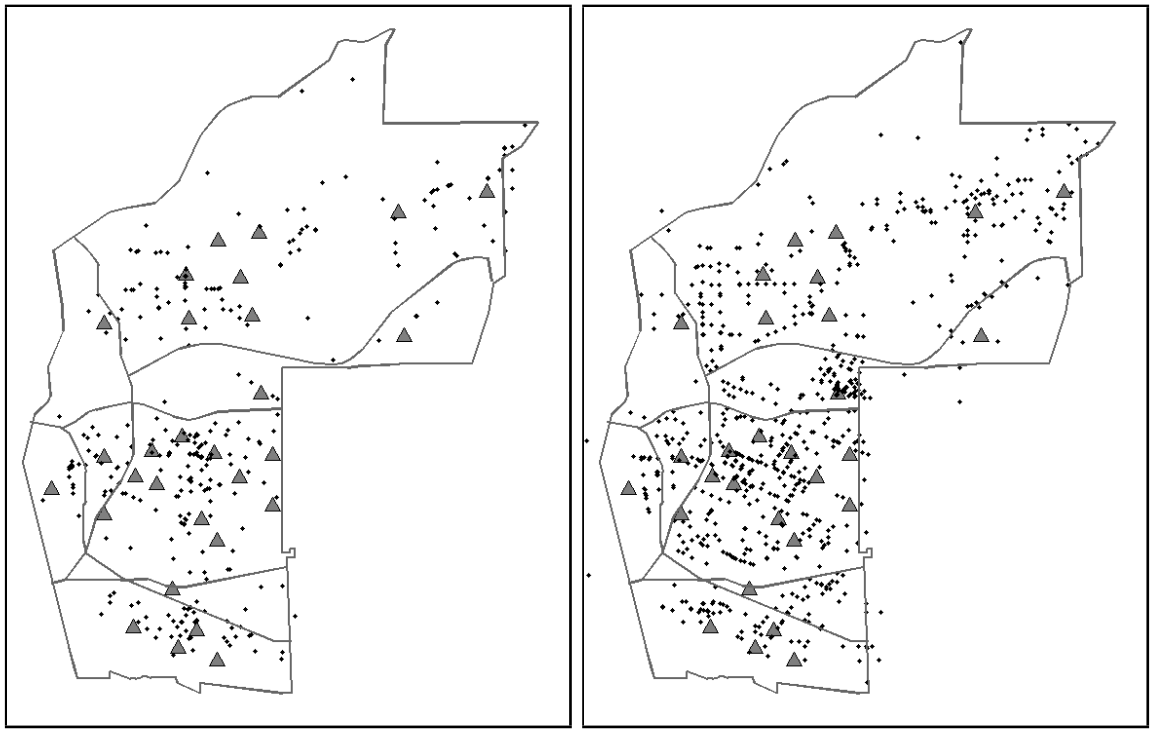

Figure 5.2: Violence and FI data in Hollenbeck. The set spaces are given by solid triangles in both images. The gang violence data (Left) and FI data (Right) are given by solid circles.

There is a possibility of errors in any data set, including inaccuracies, incompleteness, under-reporting of events, and bias [32]. Furthermore, for the data sets we are using, we should be cautious in the expectation that $p^{(i)}$ fully approximates the violence data and

$u^{(i)}$ fully approximates the FI data. The nature of the sampling frame for the FI data is not perfectly understood, and there may be significant under-reporting of all but the most serious violent crimes. It should be noted that while not every encounter may produce an FI, law enforcement needs to continually document gang members activity with their gang and fellow affiliates for prosecution in future crimes. Regularly documenting activity, especially individuals loitering together, showcases this gang acuity to criminal justice actors. Additionally, it is reasonable to assume that a large number of these events can be predicted by the densities obtained by the model.

Another potential issue to consider is the difference in time frames for each set of data. The locations of the set spaces were taken from [102] published in 2010. Some set spaces may have moved during the eight years after the violence data was collected. As cities are ever changing, gangs may have adapted to the altered environments. For example, one housing area was redeveloped, forcing several Cuatro Flats, TMC, and Primera Flats gang members to relocate. This caused these gangs to ultimately have new set spaces. For simplicity, we will only use the original set space in our model implementation.

## 5.5 Results

With this implementation of the model for the Hollenbeck gangs, we optimize the parameters with respect to the accuracy metric for the rivalry network, defined in Equation (4.7) of Section 4.4.3.1. Additionally, we must ensure that the parameter $m$ is large enough to produce a marking density that is not nearly identical to the gang density. Listed in Table 5.2 are the parameters used for this analysis.

| Parameter | Value |
|---|---|
| m | 100 |
| $\beta$ | 1 |
| $\alpha$ | 1 |
| $\Delta t$ | 0.001 |

Table 5.2: Parameter choices for the ecological territorial model.

We first compare our simulated network with that of the agent-based approach of Chapter 4 in Section 5.5.1. From here, we investigate the resulting gang and marking densities in Section 5.5.2. Given these densities, Section 5.5.3 estimates gang territory locations. One standard method for partitioning a region is through Voronoi diagrams, which we further discuss in Section 5.5.4. We then use the estimated territories to investigate the relationship with the two data sets in Section 5.5.5. Finally, we examine pairs of rival gangs and the locations of events with respect to the theoretical territories in Sections 5.5.6 and 5.5.7.

### 5.5.1 Rivalry Networks

To create a rivalry network from the ecological territorial model, we must determine what quantifies a rivalry. If the violence among gangs in Hollenbeck is primarily attributed to territorial issues, then we would expect a rivalry to exist if the densities of two gangs sufficiently overlap. We thus threshold the densities to produce an approximate territory, and then find the overlapping area between pairs of territories. Since there are 69 empirically known rivalries among the gangs in Hollenbeck, we select the top 69 area pairs to construct a network. The resulting rivalry network obtained from the simulation is given in Figure 5.3, together with the observed network and the Simulated Biased Lèvy walk Network (SBLN) obtained in Chapter 4.

We notice that the ecological territorial model correctly identifies 45 of the 69 edges in the observed network. Additionally, this simulated network has similar graph properties with both of the other networks. To better quantify this similarity, we examine a few graph metrics that were previously used in Chapter 4. We would first like to inspect the accuracy of the simulated network. We calculate the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), as well as the *accuracy* score (ACC), all defined in Section 4.4.3.1. These values are presented in Table 5.3.

While these accuracy measures are important, we would also like our rivalry network to have similar graph properties as the observed network. For example, the number of rivals that a gang has should be approximately the same in the simulation. Thus, we investigate

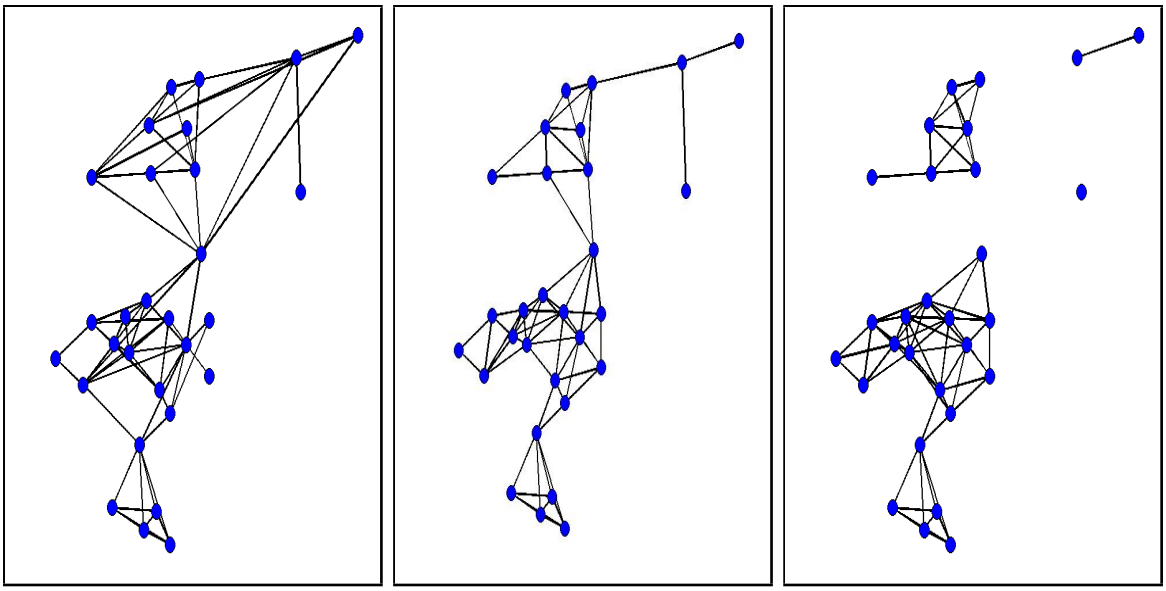

Figure 5.3: Comparison of rivalry networks. Each node in the network represents a gang's set space. An edge in the graph exists if there is a rivalry between the two gangs. The observed network (Left) is plotted with both the SBLN network from Chapter 4 (Middle) and the ecological territorial model from this paper (Right).

| | TP | TN | FP | FN | ACC |
|---|---|---|---|---|---|
| SBLN | 50 | 320 | 17 | 19 | 0.9113 |
| Ecological Territorial Model | 45 | 313 | 24 | 24 | 0.8818 |

Table 5.3: Accuracy measures for the SBLN model of Chapter 4 and the ecological territorial model.

features of the network related to the degree of a node, representing the number of rivals of a gang $i$. Three measures we use are the *graph density*, *degree variance*, and *Freeman's centrality measure*, all defined in Section 4.4.3.2. The density of the graph is a scalar multiple of the average degree. Thus, if two graphs have the same number of edges, then they will have the same density. The degree variance provides information on the spread of the degrees. Thus, a larger degree variance indicates a wider range of degrees, which translates to a larger spread in the number of rivals of each gang. The centrality measures whether there are key gangs that are more connected than the rest, indicating they are more central to the network. If all gangs have the same degree, i.e. no gang is more influential to the network,

then the centrality measure would be zero. Alternatively, if one gang is connected to every other gang, and these are the only edges in the graph, then the network would have a high centrality measure.

These graph shape measures for the observed network, the SBLN network of Chapter 4, and the ecological territorial model network are shown in Table 5.4. We note that the three networks have close values for the three metrics.

| | Density | Degree Variance | Centrality |
|---|---|---|---|
| Observed Network | 0.16995 | 4.32105 | 0.20106 |
| SBLN | 0.16503 | 3.54578 | 0.16799 |
| Ecological Territorial Model | 0.16995 | 5.70036 | 0.16270 |

Table 5.4: Graph metrics comparing the observed network with the SBLN model of Chapter 4 and the ecological territorial model.

While it is of interest to compare the rivalry network from the ecological territorial model to the observed network and the ones existing in the literature, we must also recognize that the method for obtaining the rivalry network from overlapping densities has limitations. More specifically, the edges that cross boundaries in the upper portion of the region are absent. This is expected since the model limits the smoothing of the gang densities across these boundaries. Thus, we do not anticipate much overlap in the territories of gangs in separate regions. Hence, we should examine other features of the model to determine the validity of the model.

### 5.5.2 Gang and Marking Densities

Modeled gang and marking densities display some promising features. First, the gang densities $u^{(i)}$ remain centrally located about the known set spaces, with minimal overlap between gangs. Geographic boundaries limit the spread of gang density. Marking densities are highest between adjacent gangs that have a relatively small distance between them. Otherwise, the marking densities mimic the gang densities. This is expected as the steady-state solution

of Equation (5.32) is given by

$$p^{(i)} = u^{(i)} \left[ 1 + m \sum_{j \neq i}^{n} p^{(j)} \right]$$

The plots of all gang and marking densities are provided together in Figure 5.4.

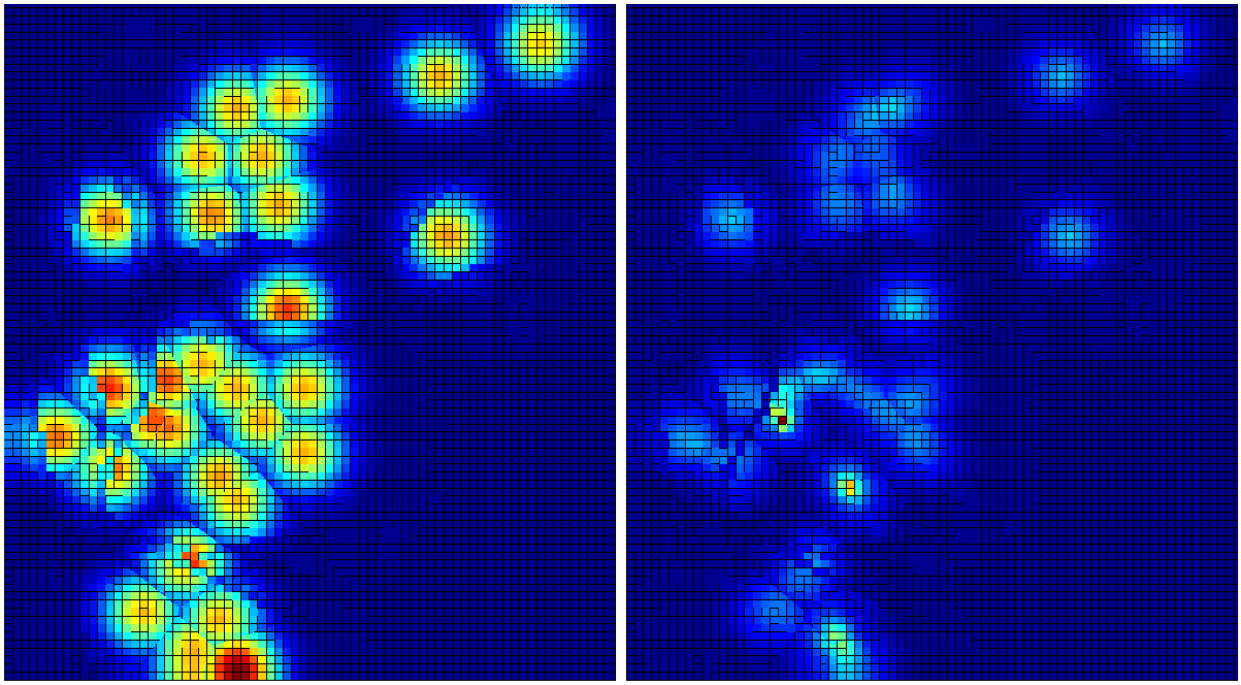

Figure 5.4: Simulated gang (Left) and marking (Right) densities in Hollenbeck. The densities for all of the gangs are plotted together. Red indicates a higher density and blue a lower density. The gang densities demonstrate the formation of territories and the impact of boundaries within the region. The marking densities are the highest between two close set spaces. Otherwise, the marking densities are similar to the gang densities.

To better demonstrate the impact of the semi-permeable boundaries that represent the major roads and the river, the gang density for a central gang is shown in Figure 5.5. The set space for this particular gang is situated between two major freeways. To illustrate the impact of the southern boundary, we view the density facing East. Notice the dramatic drop in density over the boundary. Additionally, we see that some density still flows beyond the freeway as expected, since movement is not prohibited across boundaries, only discouraged.

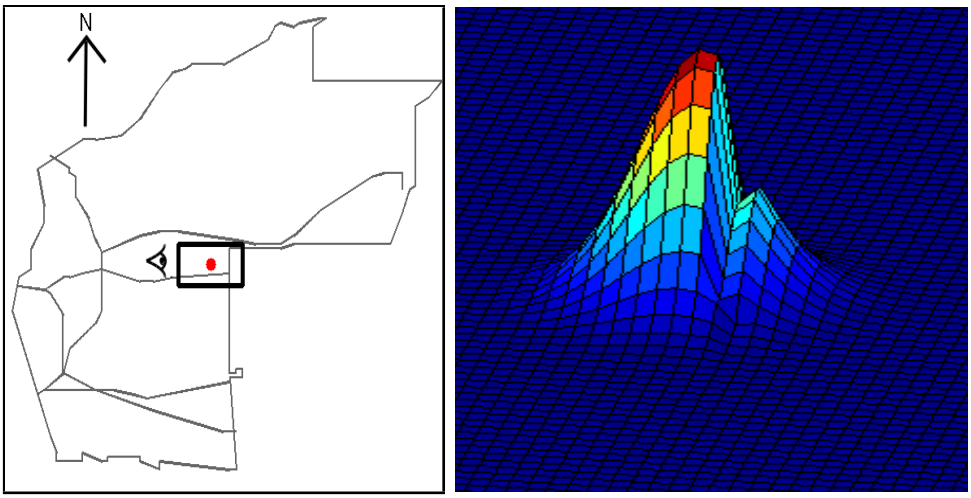


Figure 5.5: Impact of boundary on gang densities. The left figure shows the location of the set space for the gang of interest, represented by the dot situated between major roads. The eye symbol (Left) shows the lateral view for this gang's approximated density (Right). To demonstrate the impact of the southern boundary, the density is viewed facing East.

### 5.5.3 Territories

The ability to approximate gang territories is a desirable feature of this model. To do this, we classify a point $(x, y)$ in space as belonging to the territory of gang $i$ if

$$\begin{aligned} u^{(i)}(x,y) &> u^{(j)}(x,y) \qquad \text{for all } j \neq i \\ u^{(i)}(x,y) &> \delta. \end{aligned}$$

Using this thresholding with $\delta$, we limit the size of the territories so they do not cover the entire region. We took $\delta = 0.001$. The case where $\delta = 0$ is discussed in Section 5.5.4. The resulting territories are shown together in Figure 5.6. Although the semi-permeable boundaries highly influence the gang densities, some territory plots seem to extend beyond these boundaries. This is particularly evident for the southern boundary of the gang plotted in Figure 5.5.

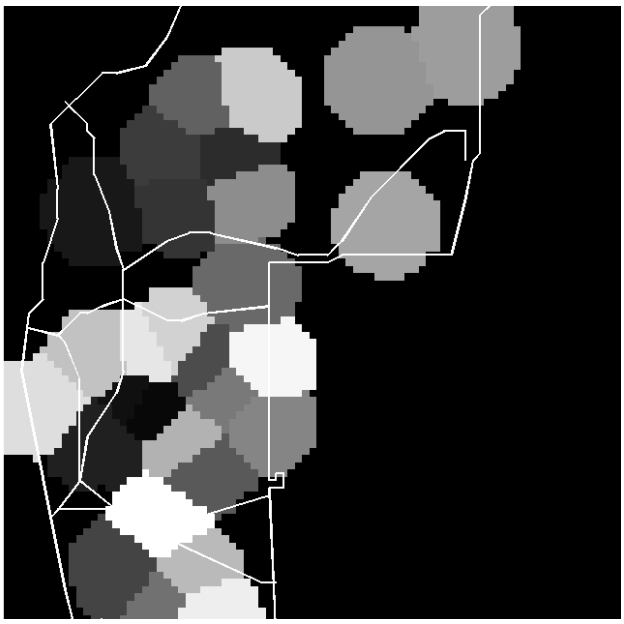

Figure 5.6: Estimated gang territories. The semi-permeable boundaries of the model and the county line are plotted over the territories in white.

### 5.5.4 Voronoi Diagrams

Given a set of points, a Voronoi diagram partitions the area of interest into smaller regions that each contain only one of the points [39]. For a point $x_A$, we consider every pair $(x_A, x_B)$ where $A \neq B$. We then construct a line between them that is equidistant to both points. After completing this for each point $x_B$, we take the resulting polygon that contains $x_A$. This is repeated for all points, decomposing the space into smaller regions. We compare the Voronoi diagram obtained from using the set spaces to the territories constructed with the ecological territorial model. We obtain interesting results if we consider $\delta = 0$ instead of $\delta = 0.001$ (i.e. no thresholding of the densities). Figure 5.7 overlays the partitioning line segments of the Voronoi diagram on the constructed territories with no thresholding. We notice that they align almost exactly.

For comparison, we do a simple approximation to the gang densities by using a Gaussian distribution centered at the set space. We used the parameter $\sigma = 3$ for the standard deviation of the distribution. Using the same technique for determining territories as above, we created a territory plot with no thresholding and plotted the Voronoi diagram over it as

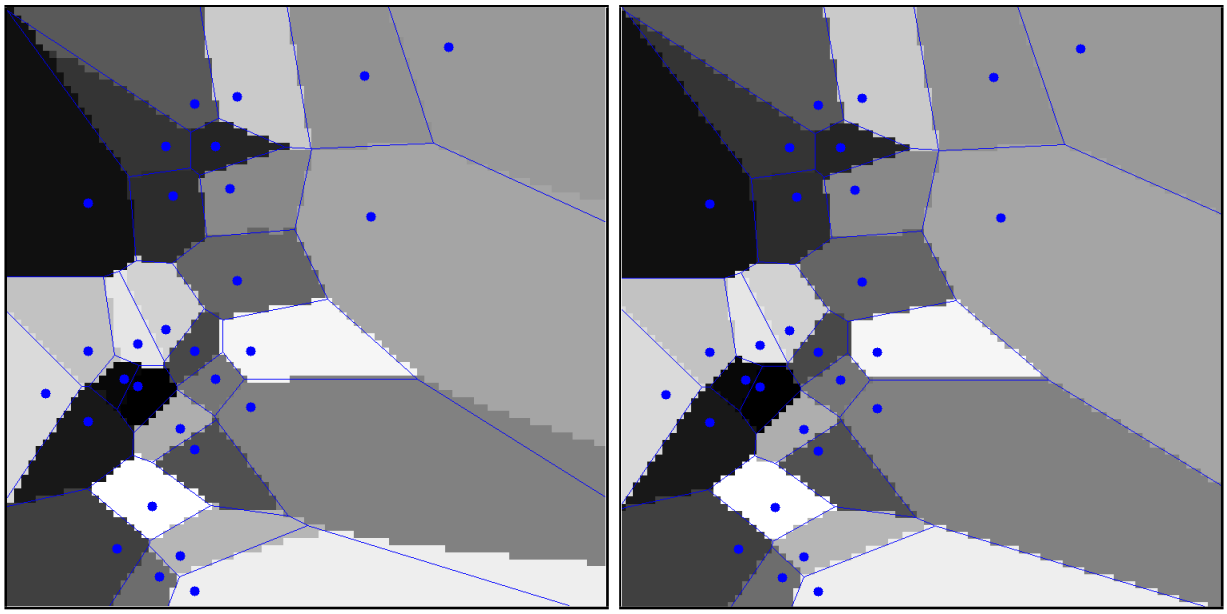

Figure 5.7: Voronoi plots compared to simulated territories. In both images, the set spaces for the gangs are represented by the solid circle. The solid lines give the Voronoi diagram for the region with input of the set spaces. The shaded regions provide the estimated gang territories with a Gaussian distribution for the gang densities (Right) and the $u^{(k)}$ gang densities obtained from the ecological territorial model (Left). Notice that the Gaussian distribution territories line up directly with the Voronoi plot, and the ecological territorial model only differs significantly on the far right portion of the region.

well. This is also shown in Figure 5.7. The boundaries of the territories match up exactly with the Voronoi diagram.

Both territory plots of Figure 5.7 and the Voronoi diagram produce a similar partitioning of the region. This suggests that the territories created using the ecological territorial model produces territories similar to using a standard normal distribution about the set space. However, we note that the normal distributions of separate gangs has significantly more overlap than the ecological territorial model gang densities. This does not show in these images as we only indicate which gang has a higher density here.

### 5.5.5 Comparison of the Densities with the Data

As discussed in Section 5.4, we have two data sets with which we will compare our model output. The FI data provides the locations of where individuals have been stopped by the

police. We assume this to be a reasonable sample of the gang densities. Thus, we plot the FI data over the gang densities in Figure 5.8. Additionally, we have violence data for interactions between gangs. These events are most likely to occur in locations where the marking density is high. Therefore, we plot this data over the marking densities in Figure 5.8.

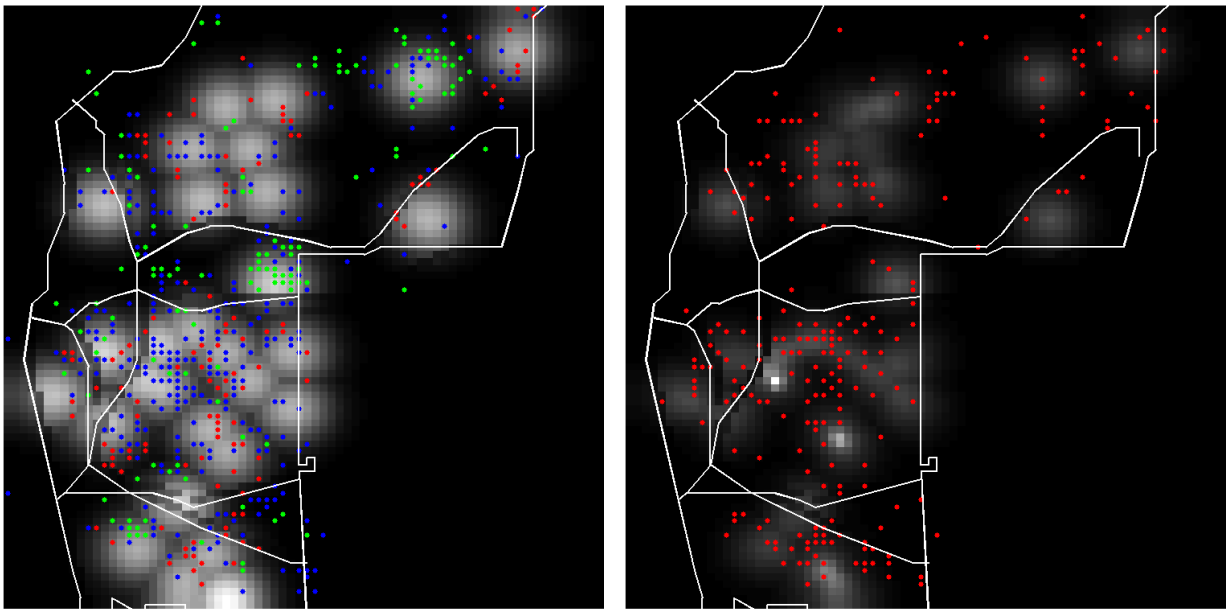

Figure 5.8: Data points and the simulated densities. The gang densities and the FI data are plotted together (Left), and the marking densities are given with the violence data (Right). Here, the densities are provided as the background gray-scale image with whiter colors representing higher values. In the FI data image (Left), the green dots are for individuals in their own territory, the red dots are for individuals in their rival gang's territory, and the blue dots are for the other cases.

To see how well the model fits the data, we examine the Akaike information criterion (AIC), given by

$$AIC = 2k - 2\ln(L), \tag{5.34}$$

where $L$ is the likelihood function and $k$ is the number of model parameters. The results are given in Table 5.5. In addition to the AIC values for the ecological territorial model, we examine the Gaussian distribution model. We find that the ecological territorial model performs better than the Gaussian distribution model when $\sigma = 3$, the standard deviation used in Figure 5.7. We optimized the $\sigma$ to give the best AIC measure, resulting in a value of $\sigma = 11.5$. However, theoretical territory plots for this value resulted in gangs having

set spaces in other gangs' theoretical territories. This is an undesirable property, which is why the image is not included in the paper, but we include the AIC value in the table for comparison.

| Method | Data Set | AIC |
|---|---|---|
| Ecological Territorial Model | FI Data | 31150 |
| Gaussian Distribution Model ($\sigma = 3$) | FI Data | 41512 |
| Gaussian Distribution Model ($\sigma = 11.5$) | FI Data | 18207 |
| Ecological Territorial Model | Violence Data | 10526 |

Table 5.5: Akaike information criterion values for the ecological territorial model and the Gaussian distribution model. For the Gaussian distributions, we use $\sigma = 3$ as in Section 5.5.4 and $\sigma = 11.5$. The latter value gives the best AIC value, but creates territories that does not contain the set spaces for some gangs. Lower AIC values are desired.

Overall, we notice some interesting properties of the locations of the individuals in the FI data. Figure 5.8 plotted these data in different colors to correspond to the various cases. In particular, green dots represent gang members that are located in their own theoretical territory as estimated in Figure 5.6. The red dots show gang members located in a rival gang's theoretical territory, and the blue dots represent individuals that are neither in their own territory nor a rival gang's territory. Here, we determine a rival gang from the observed rivalry network. To help examine the individual cases, we plot them separately in Figure 5.9. We notice that the data points corresponding to locations where the gang members are not in their own theoretical territory are more likely to occur in areas near the boundaries of other gangs' theoretical territories. Perhaps these individuals were passing through the region by slipping between territories.

We highlight the FI data for three individual gangs in Figure 5.10. We notice that the densities visually indicate where to find the majority of the gang members. Instances where the individuals are not in the general vicinity of the home territory often occur near the major roads.

Next, we examine the violence data and how it compares to the theoretical marking densities produced by the model. Table 5.5 provides the AIC value for the violence data.

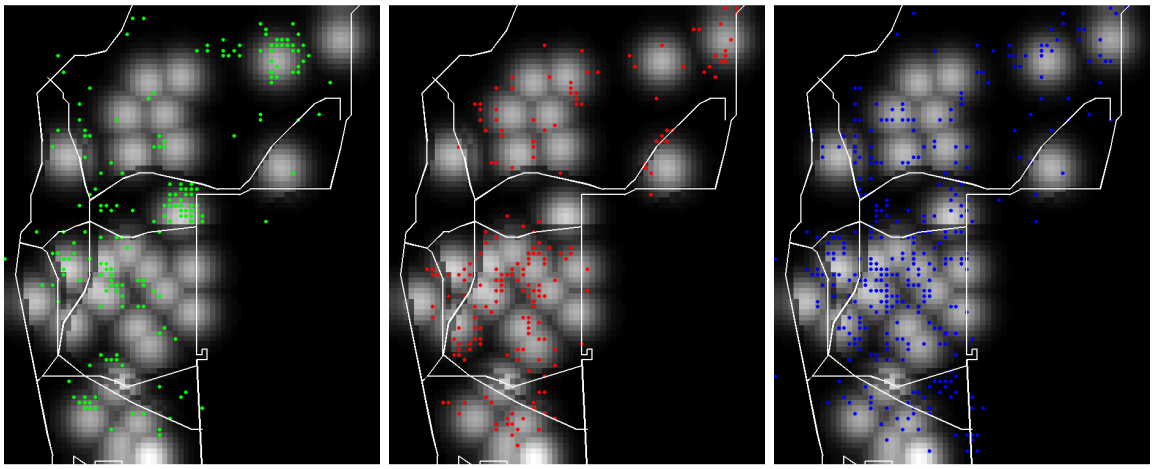

Figure 5.9: FI data plotted over the gang densities determined by the ecological territorial model. The gang members in their own territories (Left), the gang members in their rival's territory (Middle), and the other cases (Right) are plotted here. A dot represents a location where a gang member was found. The gray-scale image in the background shows the gang densities estimated by the ecological territorial model.

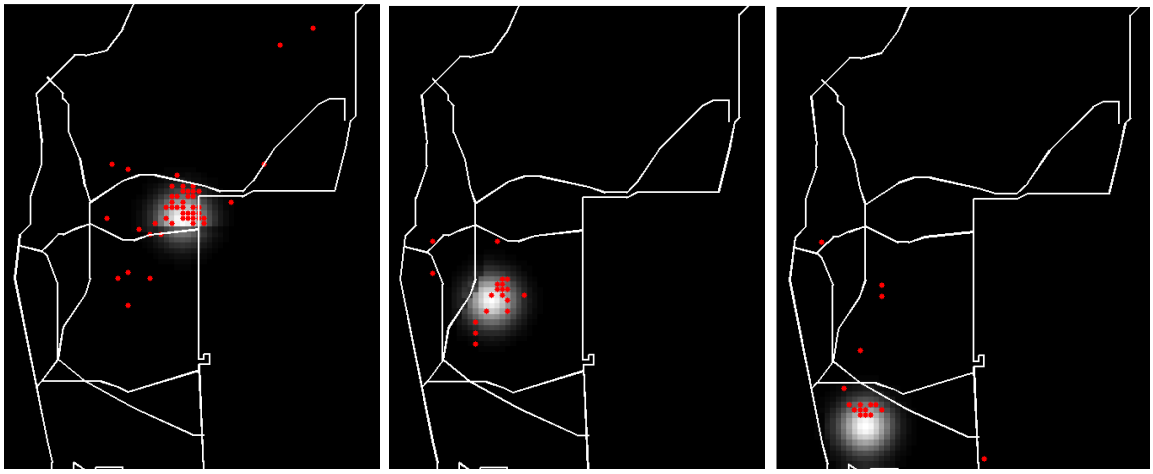

Figure 5.10: FI data for three individual gangs. The red dots indicate the position of where a gang member was located. The gray-scale image provides the gang's density information estimated by the ecological territorial model.

Since we do not have another model to compare this with, we instead flag a certain portion of the city and find how many events occurred in this region. More specifically, we take all

points such that for any $i$,

$$p^{(i)}(x,y) > \gamma \cdot \max_{j,x,y} \left(p^{(j)}(x,y)\right).$$

We took $\gamma$ to be $0.2, 0.1$, and $0.05$. The flagged regions are shown in Figure 5.11, and more information on the results are provided in Table 5.6. By flagging less than 14% of the cells, we predict more than half of the data.

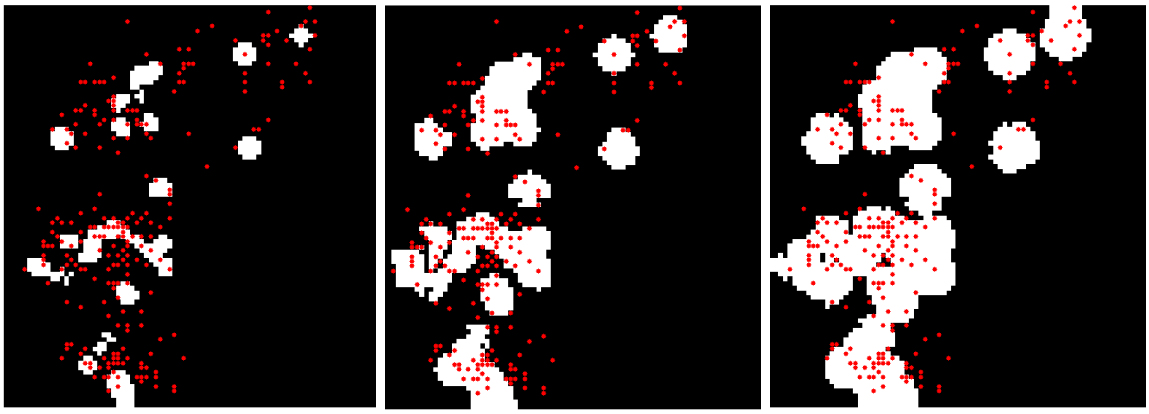

Figure 5.11: Violence data compared to flagged regions of the city. From left to right, we have $\gamma = 0.2, 0.1, 0.05$. The red dots indicate locations of violence between two gangs. The white regions are the flagged cells of the marking densities where we expect the events are more likely to occur.

| $\gamma$ | Percent of City Flagged | Percent of Violence Data Predicted |
|---|---|---|
| 0.2 | 5.02% | 20.61% |
| 0.1 | 13.94% | 50.91% |
| 0.05 | 22.46% | 71.21% |

Table 5.6: Proportion of data predicted given a flagged region of the city.

### 5.5.6 Rival Gang Territories and Events

Given approximate territories for gangs, we can investigate the locations of events with respect to these regions. Figure 5.12 gives the theoretical territories of three sets of rival gangs and the locations of the known violent interactions between them. Since we have

information about which gang is the suspect and which is the victim, we illustrate which gang is the victim by matching the interior color of the marker with the territory color. In these three rivalries, we notice that all events occur near the boundaries of the theoretical territories. Events slightly to the interior of the theoretical territory boundary correspond with victim gangs in their own territory.

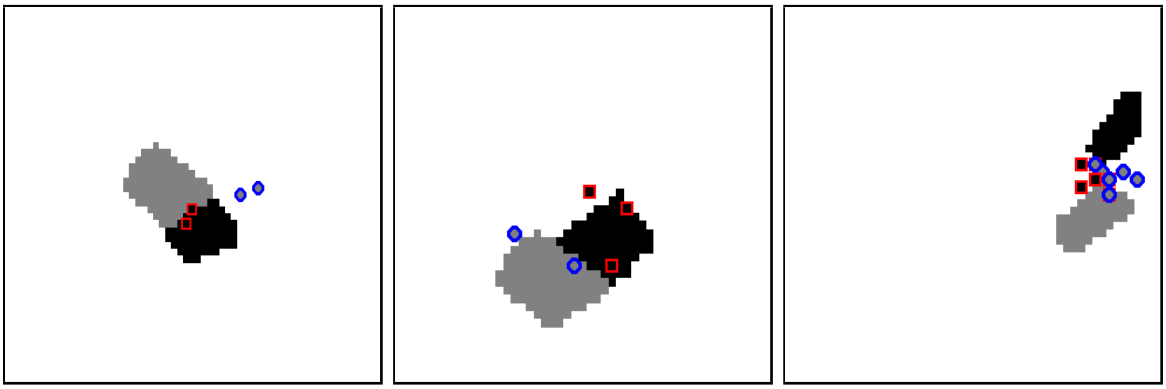

Figure 5.12: Territories and events of rival gangs. The three plots represent three pairs of rival gangs and the violent interactions between them. The interior color of the markers (circles and squares) indicates which gang is the victim by matching the color of its territory. All events occur near the boundaries of the territories.

Sometimes there exists a rivalry between gangs of varying sizes. One such pair is illustrated in Figure 5.13, with one gang significantly larger than the other. Counterintiuitively, the gang with the black set space is the larger gang and has a much greater number of events where it is the victim. We also notice that the smaller gang is only the victim on the boundary of the larger gang's theoretical territory or on the interior of its own theoretical territory. Since this data was recorded, the smaller gang collapsed after a key gang member was killed. Thus, the smaller number of attacks by the larger gang may have had more impact than the larger number made by the smaller gang.

### 5.5.7 Suspect versus Victim

In the previous section, we noticed a pattern where the locations of events occur with respect to theoretical territory boundaries. Figure 5.14 shows six pairs of rival gangs and where their violent interactions occurred. We notice a trend where the events typically fall either near

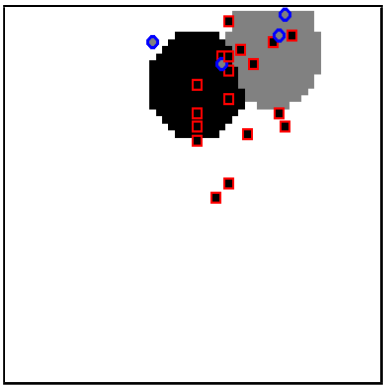

Figure 5.13: A rival pair of gangs with different sizes. The larger gang, whose theoretical territory is in black, is significantly larger than the smaller gang. The smaller gang is only the victim on the boundary of the larger gang's territory or on the interior of its own territory.

the boundary of the approximated territories of either gang, on the interior of the victim's territory, or are far from either theoretical territory.

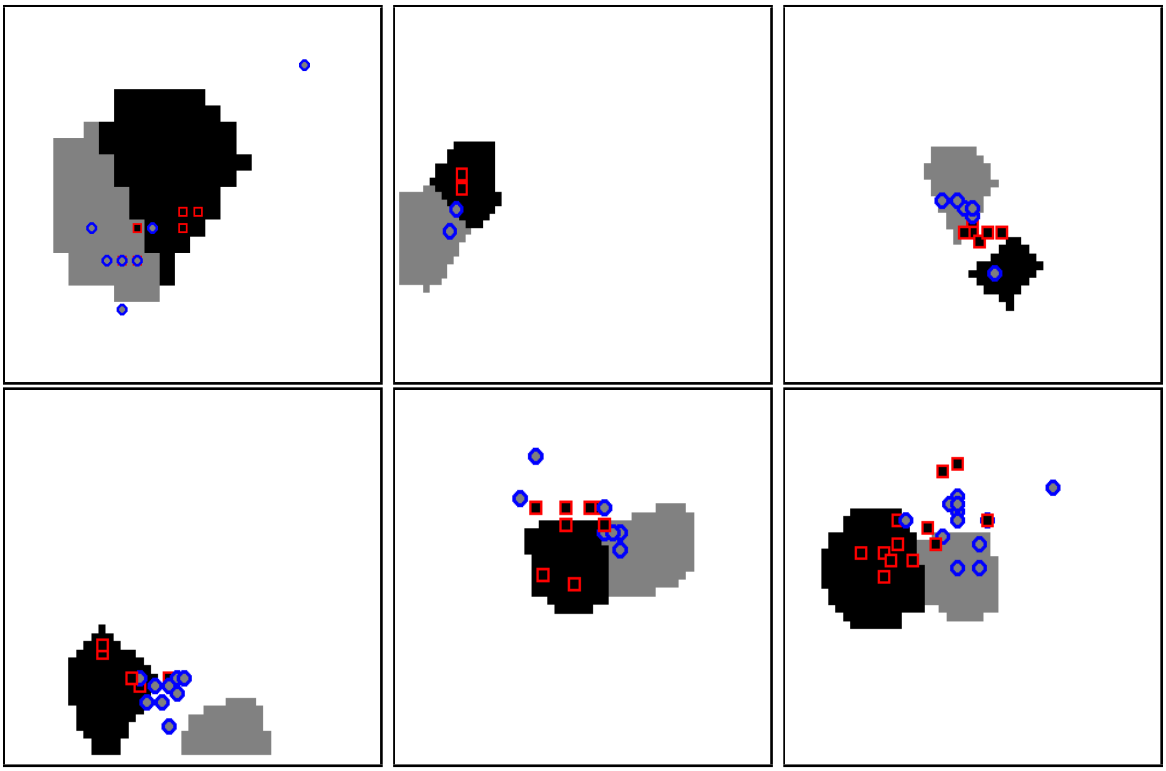

Figure 5.14: Rival gang territories and the events between them. The victim's gang's territory color matches the interior of the marker. The majority of events occur near the approximate territory boundaries, on the interior of the victim's territory, or away from both territories.

With this observation, we consider the location of events with respect to the boundaries of the approximate territories for both the suspect and the victim gangs. Previously, the authors of [20] investigated the positioning of the events with respect to a boundary line between gangs, equidistant to both set spaces. With the theoretical boundaries from the ecological territorial model, we are able to further analyze this data with respect to all boundaries. Given the victim gang's theoretical territory, we find the distance from an event to the territory boundary. We let this value be negative if the event occurs in the victim's territory and positive if it is outside. A histogram of these distances is provided in Figure 5.15, where one unit is equivalent to 0.1 km. We see that the events strongly cluster

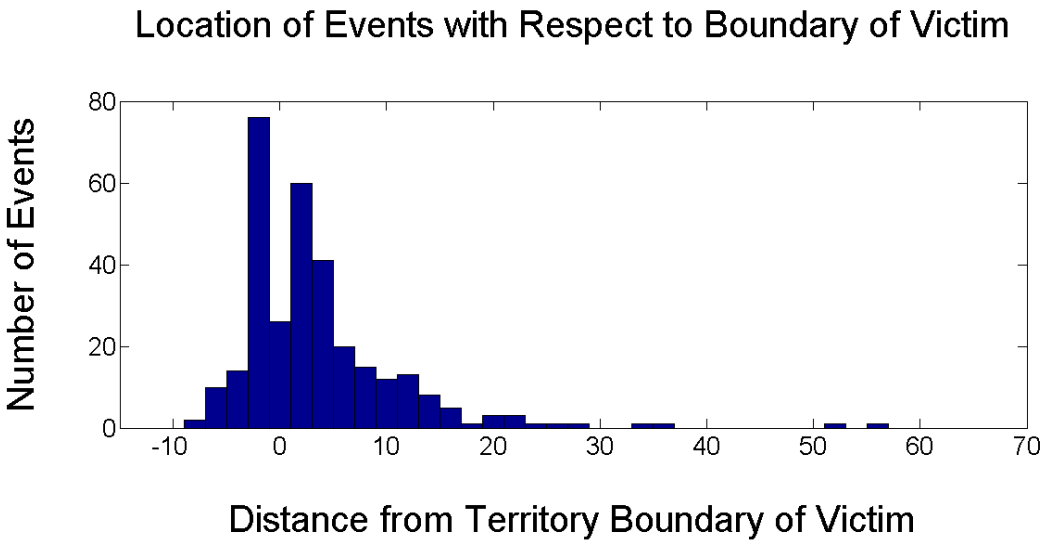


Figure 5.15: Histogram of the distance of an event to the boundary of the victim's territory. Negative values indicate the event occurred on the interior of the victim's territory, and positive values are for events on the exterior of the region. The distance is the number of grid units to the closest boundary point. 1 unit = 0.1 km.

near the boundary of the victim. We also note that the distribution is slightly skewed to have a larger number of events just interior to the victim's territory. This agrees with our observation in Figure 5.14.

Since the rival gangs are not necessarily adjacent to each other, we also repeat this analysis for the location of events with respect to the suspect's theoretical territory. Again, negative

values indicate the event occurs on the interior of the suspect's territory, and positive values for the outside. The histogram is included in Figure 5.16. From this figure, we observe that

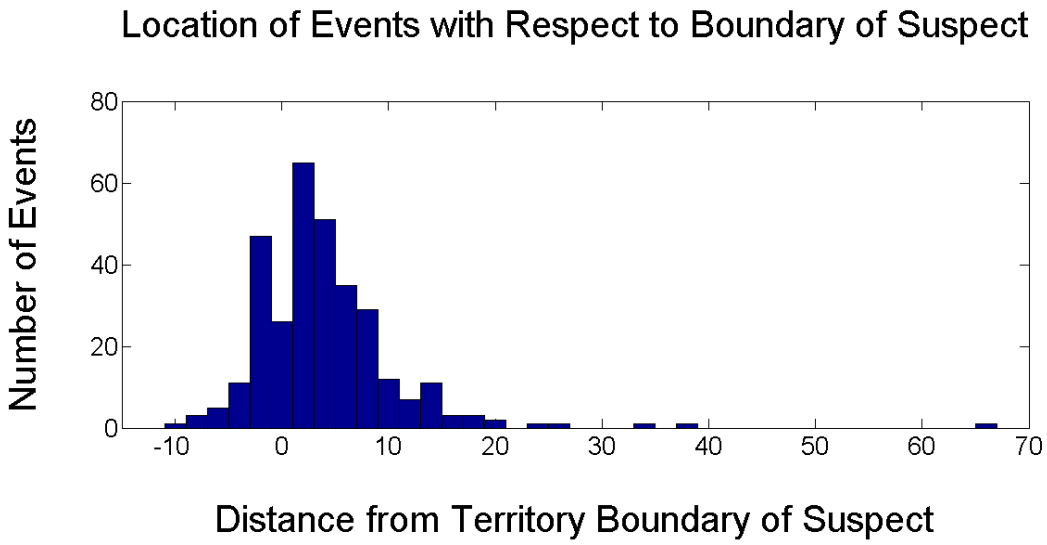


Figure 5.16: Histogram of the distance of an event to the boundary of the suspect's territory. Negative values indicate the event occurred on the interior of the suspect's territory, and positive values are for events on the exterior of the region. The distance is the number of grid units to the closest boundary point. 1 unit = 0.1 km.

the majority of events occur near the boundary of the suspect's territory, but skewed to the exterior of the region.

Quantitatively, we find 24.7% of the events occur within three grid units of the theoretical boundaries of both suspect and victim gangs. Additionally, 66.8% of the events occur within three grid units of either the suspect or victim gang's territory boundary.

## 5.6 Discussion

Using the framework of the coyote model proposed in [81], we formulated a model for street gang spatial behavior. This model was able to incorporate spatial features of the region that limit the movement individuals across these semi-permeable boundaries. For each gang, the model produced a gang density and marking density, indicating where gang members are

expected to be and where their tags are most likely to be found.

As a first approach to evaluating this model, we compared the rivalry network from this model to the observed network and the SBLN network of Chapter 4. While this model's network gave similar metric values, it did not outperform the SBLN network. This is due to the fact that we constructed the rivalry network from overlapping gang densities, which are discouraged from spreading beyond the semi-permeable boundaries encoded in the model. Thus, rivalries that cross these boundaries are not likely to be captured by the method we used to produce a network. However, this method is able to provide more information than just a rivalry network. In particular, it can be used to approximate gang territories and the locations of violent interactions.

From the gang densities produced by the model, we approximated gang territories by classifying a point based on which gang has the highest density at that location. Without any thresholding, the resulting territory plot closely approximates a Voronoi diagram. With thresholding, the resulting territory plots give reasonable estimates for each gang's turf.

By flagging the cells with the highest theoretical marking densities, we were able to identify the majority of the violent events with a small portion of the city. Additionally, we were able to identify interesting behaviors of rival gangs. Events were more likely to occur in the interior of the victim gang's theoretical territory or on the boundary of either the victim or suspect gang's theoretical territory.

While comparing the gang densities to the FI data, we found the gang members were often not in the general vicinity of their approximated home territory. These individuals were frequently traveling on the boundaries of other gang's territories. Despite this, we demonstrated with three gangs that we were still able to give a good density estimate to the location of gang members.

# References


[1] A. S. Ackleh, B. G. Fitzpatrick, R. Scribner, N. Simonsen, and J. J. Thibodeaux. Ecosystem modeling of college drinking: Parameter estimation and comparing models to data. Mathematical and Computer Modelling, 50:481–497, 2009.

[2] R. Albert and A. Barabási. Statistical mechanics of complex networks. Reviews of Modern Physics, 74(1):47–97, January 2002.

[3] A. Alonso. Urban graffiti on the city landscape. Western Geography Graduate Conference, February 1998.

[4] L. Ambrosio and V. M. Tortorelli. Approximation of functional depending on jumps by elliptic functional via Γ-convergence. Communications on Pure and Applied Mathematics, 43(8):999–1036, 1990.

[5] E. Anderson. Code of the Street: Decency, Violence, and the Moral Life of the Inner City. W. W. Norton & Company, Inc., 2000.

[6] D. G. Aronson. Density-dependent interaction-diffusion systems. In W. E. Stewart, W. H. Ray, and C. C. Conley, editors, Dynamics and Modelling of Reactive Systems, pages 161–176. Academic Press, 1980.

[7] J. Austin, E. Smith, S. Srinivasan, and F. Sanchez. Social dynamics of gang involvement: A mathematical approach. Mathematical and Theoretical Biology Institute, MTBI-08-08M, 2011.

[8] P. Baldi, S. Brunak, Y. Chauvin, C. A. Andersen, and H. Nielsen. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics, 16(5):412–424, May 2000.

[9] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rougé. A variational model for P+XS image fusion. International Journal of Computer Vision, 69(1):43–58, 2006.

[10] H. T. Banks, P. M. Kareiva, and L. Zia. Analyzing field studies of insect dispersal using two-dimensional transport equations. Environmental Entomology, 17:815–820, 1988.

[11] A. Barbaro, L. Chayes, and M. R. D'Orsogna. Territorial development based on graffiti: a statistical mechanics approach. UCLA CAM Report 11-70, 2011.

[12] A. Barbaro, B. Einarsson, B. Birnir, S. Sigurdsson, H. Valdimarsson, O. K. Palsson, S. Sveinbjornsson, and T. Sigurdsson. Modelling and simulations of the migration of pelagic fish. ICES Journal of Marine Science, 66(5):826–838, 2009.

[13] R. Berk. How you can tell if the simulations in computational criminology are any good. Journal of Experimental Criminology, 4:289–308, 2008. 10.1007/s11292-008-9053-5.

[14] D. J. Best and N. I. Fisher. Efficient simulation of the von Mises distribution. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(2):152–157, 1979.

[15] L. M. A. Bettencourt, A. Cintròn-Arias, D. I. Kaiser, and C. Castillo-Chàvez. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. Physica A: Statistical Mechanics and its Applications, 364:513–536, 2006.

[16] C. R. Block and R. Block. Street gang crime in Chicago. U.S. Department of Justice, Office of Justice Programs, National Institute of Justice: Research in Brief, December 1993.

[17] M. Bradonjić, A. Hagberg, and A. G. Percus. Giant component and connectivity in geographical threshold graphs. In Algorithms and Models for the Web-Graph (WAW 2007), pages 209–216, San Diego, CA USA, 2007.

[18] M. Bradonjić, A. Hagberg, and A. G. Percus. The structure of geographical threshold graphs. Internet Mathematics, 5:113–140, 2009.

[19] P. J. Brantingham and G. E. Tita. Offender mobility and crime pattern formation from first principles. In L. Liu and J. Eck, editors, Artificial Crime Analysis Systems: Using Computer Simulations and Geographic Information Systems, pages 193–208. Idea Press, 2008.

[20] P. J. Brantingham, G. E. Tita, M. B. Short, and S. E. Reid. The ecology of gang territorial boundaries. Criminology, in press.

[21] L. Bregman. The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. USSR Computational Mathematics and Mathematical Physics, 7:200–217, 1967.

[22] J. S. Brew. Competition and niche dynamics from steady-state solutions of dispersal equations. Theoretical Population Biology, 32:240–261, 1987.

[23] B. Briscoe, M. A. Lewis, and S. Parrish. Home range formation in wolves due to scent marking. Bulletin of Mathematical Biology, 64:261–284, 2002. 10.1006/bulm.2001.0273.

[24] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. Nature, 439:462–465, 2006.

[25] W. K. Brown. Graffiti, identity and the delinquent gang. Internal Journal of Offender and Comparative Criminology, 22(1):46–48, 1978.

[26] J. A. Canizo, J. A. Carrillo, and J. Rosado. A well-posedness theory in measures for some kinetic models of collective motion. Mathematical Models and Methods in Applied Sciences, 21(3):515–539, 2011.

[27] H. C. Covey. Street Gangs Throughout the World. Charles C Thomas, Publisher, 2010.

[28] P. L. Davies and A. Kovac. Densities, spectral densities and modality. The Annals of Statistics, 32:1093–1136, 2004.

[29] S. H. Decker and B. Van Winkle. Life in the Gang: Family, Friends, and Violence. Cambridge University Press, 1996.

[30] M. R. D'Orsogna, Y. Chuang, A. L. Bertozzi, and L. Chayes. Self-propelled particles with soft-core interactions: patterns, stability, and collapse. Physical Review Letters, 96, 2006. 104302.

[31] P. K. Eason, G. A. Cobbs, and K. G. Trinca. The use of landmarks to define territorial boundaries. Animal Behaviour, 58(1):85 – 91, 1999.

[32] J. Eck and L. Liu. Contrasting simulated and empirical experiments in crime prevention. Journal of Experimental Criminology, 2008.

[33] L. Edelstein-Keshet. Mathematical models in biology. Random House, 1986.

[34] M. Egesdal, C. Fathauer, K. Louie, and J. Neuman. Statistical modeling of gang violence in Los Angeles. SIAM Undergraduate Research Online, 3, 2010.

[35] P. P. B. Eggermont and V. N. LaRiccia. Maximum Penalized Likelihood Estimation. Springer, 2001.

[36] S. Esedoglu and J. Shen. Digital inpainting based on the Mumford-Shah-Euler image model. European Journal of Applied Mathematics, 13:353–370, 2002.

[37] S. Eubank, H. Guclu, V. S. A. Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. Modelling disease outbreaks in realistic urban social networks. Nature, 429:180–184, 2004.

[38] J. D. Farmer and D. Foley. The economy needs agent-based modelling. Nature, 460:685–686, August 2009.

[39] S. Fortune. Voronoi diagrams and Delaunay triangulations. In D. Du and F. Hwang, editors, Computing in Euclidean Geometry, volume 4 of Lecture Notes Series on Computing, pages 225–236. World Scientific, 1995.

[40] L. C. Freeman. Centrality in social networks conceptual clarification. Social Networks, 1(3):215 – 239, 1978-1979.

[41] C. Fremon. G-Dog and the Homeboys: Father Greg Boyle and the Gangs of East Los Angeles. University of New Mexico Press, 2008.

[42] M. T. Gastner and M. E. J. Newman. The spatial structure of networks. European Physical Journal B - Condensed Matter and Complex Systems, 49:247–252, 2006. 10.1140/epjb/e2006-00046-8.

[43] L. Giuggioli, J. R. Potts, and S. Harris. Animal interactions and the emergence of territoriality. PLoS Computational Biology, 7:1–9, 2011.

[44] T. Goldstein and S. Osher. Split Bregman method for L1 regularized problems. SIAM Journal on Imaging Sciences, 2:323–343, 2009.

[45] M. González, C. A. Hildalgo, and A. L. Barabási. Understanding individual human mobility patterns. Nature Letters, 453:779–782, 2008.

[46] I. J. Goodd and R. A. Gaskins. Nonparametric roughness penalties for probability densities. Biometrika, 58(2):255–277, 1971.

[47] M. B. Gordon. A random walk in the literature on criminality: A partial and critical view on some statistical analyses and modelling approaches. European Journal of Applied Mathematics, 21(Special Double Issue 4-5):283–306, 2010.

[48] E. R. Groff. 'situating simulation to model human spatio-temporal interactions: An example using crime events. Transactions in GIS, 11(4):507–530, 2007.

[49] W. S. C. Gurney and R. M. Nisbet. The regulation of inhomogeneous populations. Journal of Theoretical Biology, 52:441–457, 1975.

[50] K. Hawkes, K. Hill, and J. F. O'Connell. Why hunters gather: Optimal foraging and the achè of eastern paraguay. American Ethnologist, 9:379–398, 1982.

[51] R. A. Hegemann, E. A. Lewis, and A. L. Bertozzi. An "estimate & score algorithm" for simultaneous parameter estimation and reconstruction of missing data on social networks. UCLA CAM Report 12-12, 2012.

[52] R. A. Hegemann, L. M. Smith, A. Barbaro, A. L. Bertozzi, S. E. Reid, and G. E. Tita. Geographical influences of an emerging network of gang rivalries. Physica A: Statistical Mechanics and its Applications, 390(21-22):3894 – 3914, 2011.

[53] I. S. Helland, J. M. Hoff, and G. Anderbrant. Attraction of bark beetles (Coleoptera:Scolytidae) to a phermone trap: experiment and mathematical models. Journal of Chemical Ecology, 10:723–752, 1984.

[54] C. K. Hemelrijk and K. Hanspeter. Density distribution and size sorting in fish schools: an individual-based model. Behavioral Ecology, 16(1):178–187, 2004.

[55] E. E. Holmes, M. A. Lewis, J. E. Banks, and R. R. Veit. Partial differential equations in ecology: Spatial interactions and population dynamics. Ecology, 75(1):17–29, 1994.

[56] Information Technology Division Management Report Unit. Statistical digest. Los Angeles Police Department, 2008.

[57] S. R. Jammalamadaka and A. SenGupta. Topics in Circular Statistics, volume 5 of Multivariate Analysis. World Scientific Publishing Co. Pte. Ltd., 2001.

[58] J. E. Jankowski, S. K. Robinson, and D. J. Levey. Squeezed at the top: Interspecific aggression may constrain elevational ranges in tropical birds. Ecology, 91(7), 2010.

[59] B. Jiang, J. Yin, and S. Zhao. Characterizing the human mobility pattern in a large street network. Physical Review E, 2009.

[60] P. A. Jones, P. J. Brantingham, and L. Chayes. Statistical models of criminal behavior: The effects of law enforcement actions, 2010.

[61] M. J. Keeling and P. Rohani. Modeling Infectious Diseases in Humans and Animals. Princeton University Press, 2008.

[62] M. Klein, H. Kerner, C. L. Maxson, and E. G. M. Weitekamp. Euro Gang Paradox: Street Gangs and Youth Groups in the U.S. and Europe. Kluwer Academic Publisher, 2001.

[63] M. Klein and C. Maxson. Street Gang Patterns and Policies. Oxford University Press, 2006.

[64] M. Klein, F. M. Weerman, and T. P. Thornberry. Street gang violence in Europe. European Journal of Criminology, 3(4):413–437, 2006.

[65] R. Koenker and I. Mizera. Density estimation by total variation regularization. Advances in Statistical Modeling and Inference, Essays in Honor of Kjell A. Doksum, World Scientific:613–634, 2007.

[66] C. Kooperberg and C. J. Stone. Logspline density estimation with free knots. Computational Statistics and Data Analysis, 12:327–347, 2002.

[67] M. A. Lewis and J. D. Murray. Modelling territoriality and wolf-deer interactions. Nature, 366:738–740, 1993.

[68] M. A. Lewis, K. A. J. White, and J. D. Murray. Analysis of a model for wolf territories. Journal of Mathematical Biology, 35:749–774, 1997. 10.1007/s002850050075.

[69] D. Ley and R. Cybriwsky. Urban graffiti as territorial markers. Annals of the Association of American Geographers, 64(4):491–505, 1974.

[70] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. Proceedings of the National Academy of Sciences, 102(33):1162311628, August 2005.

[71] W. Liu, M. B. Short, Y. E. Taima, and A. L. Bertozzi. Multiscale collaborative searching through swarming. Proceedings of the 7th International Conference on Informatics in Control, Automation and Robotics (ICINCO), June 2010.

[72] M. W. Macy and R. Willer. From factors to actors: computational sociology and agent-based modeling. Annual Review of Sociology, 28:143–166, 2002.

[73] K. V. Mardia and P. E. Jupp. Directional Statistics. John Wiley & Sons Ltd, 2000.

[74] N. Masuda, H. Miwa, and N. Konno. Geographical threshold graphs with small-world and scale-free properties. Physical Review E, 71(3):036108, March 2005.

[75] B. W. Matthews. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta, 405(2):442–451, 1975.

[76] S. M. Mniszewski, S. Y. Del Valle, P. D. Stroud, J. M. Riese, and S. J. Sydoriak. EpiSims simulation of a multi-component strategy for pandemic influenza. In SpringSim '08: Proceedings of the 2008 Spring simulation multiconference, pages 556–563, San Diego, CA, USA, 2008. Society for Computer Simulation International.

[77] M. Moeller, T. Wittman, and A. L. Bertozzi. Variational wavelet pan-sharpening. UCLA CAM Report 08-81, December 2008.

[78] G. O. Mohler, A. L. Bertozzi, T. A. Goldstein, and S. J. Osher. Fast TV regularization for 2D maximum penalized likelihood estimation. Journal of Computational and Graphical Statistics, 20, 2011.

[79] P. R. Moorcroft and A. Barnett. Mechanistic home range models and resource selection analysis: A reconciliation and unification. Ecology, 89:1112–1119, 2008.

[80] P. R. Moorcroft, M. A. Lewis, and R. L. Crabtree. Home range analysis using a mechanistic home range model. Ecology, 80:1656 – 1665, 1999.

[81] P. R. Moorcroft, M. A. Lewis, and R. L. Crabtree. Mechanistic home range models capture spatial patterns and dynamics of coyote territories in yellowstone. Proceedings: Biological Sciences, 273(1594):1651–1659, 2006.

[82] J. Moore. Homeboys: Gangs, Drugs, and Prison in the Barrios of Los Angeles. Temple University Press, 1978.

[83] J. Moore. Going Down to the Barrio: Homeboys and Homegirls in Change. Temple University Press, 1991.

[84] J. Moore and R. Pinderhughes, editors. In the Barrios: Latinos and the Underclass Debate. Russel Sage Foundation, 1993.

[85] J. Moore, D. Vigil, and R. Garcia. Residence and territoriality in chicano gangs. Social Problems, 31(2):pp. 182–194, 1983.

[86] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics, 42(5):577–685, 1989.

[87] J. D. Murray. Mathematical Biology, volume 19. Springer-Verlag, 1989.

[88] M. E. J. Newman. Scientific collaboration networks. I. network construction and fundamental results. Physical Review E, 64(1):016131, June 2001.

[89] M. E. J. Newman. Scientific collaboration networks. II. shortest paths, weighted networks, and centrality. Physical Review E, 64(1):016132, June 2001.

[90] M. E. J. Newman. The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences, 98(2):404–409, 2001.

[91] M. E. J. Newman. The structure and function of complex networks. SIAM Review, 45(2):167–256, 2003.

[92] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. Physical Review E, 64(2):026118, July 2001.

[93] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks. Proceedings of the National Academy of Sciences, 99(suppl. 1):2566–2572, February 2002.

[94] A. Okubo. Differential and Ecological Problems: Mathematical Models. Springer-Verlag, 1980.

[95] A. Okubo. Dynamical aspects of animal grouping: swarms, schools, flocks, and herds. Advances in Biophysics, 22:1–94, 1986.

[96] M. O'Leary. The mathematics of geographic profiling. Journal of Investigative Psychology and Offender Profiling, 6:253 –265, 2009.

[97] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. Multiscale Modeling and Simulation, 4(2):460–489, 2005.

[98] A. V. Papachristos. Murder by structure: Dominance relations and the social structure of gang homicide. American Journal of Sociology, 115(1):74–128, 2009.

[99] R. P. Peters and L. D. Mech. Scent-marking in wolves: Radio-tracking of wolf packs has provided definite evidence that olfactory sign is used for territory maintenance and may serve for other forms of communication within the pack as well. American Scientist, 63(6):pp. 628–637, 1975.

[100] A. B. Pitcher. Adding police to a mathematical model of burglary. European Journal of Applied Mathematics, 21(Special Double Issue 4-5):401–419, 2010.

[101] Product No. 2009-M0335-001. National gang threat assessment. National Gang Intelligence Center, January 2009.

[102] S. Radil, C. Flint, and G. E. Tita. Spatializing social networks: Using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles. Annals of the Association of American Geographers, 100(2):307–326, 2010.

[103] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong. On the Lévy-walk nature of human mobility: Do humans walk like monkeys? In IEEE INFOCOM 2008 - IEEE Conference on Computer Communications, pages 924–932. IEEE, April 2008.

[104] D. M. Romero, C. M. Kribs-Zaleta, A. Mubayi, and C. Orbe. An epidemiological approach to the spread of political third parties. Social Sciences Research Network, 2009.

[105] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. Physica D, 60:259–268, 1992.

[106] P. E. Sandstrom. An optimal foraging approach to information seeking and use. The Library Quarterly, 64:414–449, 1994.

[107] S. Sardy and P. Tseng. Density estimation by total variation penalized likelihood driven by the sparsity $l_1$ information criterion. Scandinavian Journal of Statistics, 37:321–337, 2010.

[108] F. Schweitzer. Brownian Agents and Active Particles: Collective Dynamics in the Natural and Social Sciences. Springer, 2003.

[109] F. Schweitzer and B. Tilch. Self-assembling of networks in an agent-based model. Physical Review E, 66(2):026113, August 2002.

[110] F. Sebastiani. Machine learning in automated text categorization. ACM Computing Surveys, 34(1):1–47, 2002.

[111] N. Shigesada. Spatial distribution of dispersing animals. Journal of Mathematical Biology, 9:85–96, 1980.

[112] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segredation of interacting species. Journal of Theoretical Biology, 79:83–99, 1979.

[113] M. B. Short, P. J. Brantingham, A. L. Bertozzi, and G. E. Tita. Dissipation and displacement of hotspots in reaction-diffusion models of crime. Proceedings of the National Academy of Sciences, 107(9):3961–3965, 2010.

[114] M. B. Short, P. J. Brantingham, and M. R. D'Orsogna. Cooperation and punishment in an adversarial game: How defectors pave the way to a peaceful society. Physical Review E, 82(6):066114, Dec 2010.

[115] 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. Mathematical Models and Methods in Applied Sciences (M3AS), 18:1249–1267, 2008.

[116] M. B. Short, G. Mohler, P. J. Brantingham, and G. E. Tita. Gang rivalry dynamics via coupled point process networks. http://math.scu.edu/ gmohler/gangnetwork.pdf, 2011.

[117] B. W. Silverman. Kernel density estimation using the Fast Fourier Transform. Statistical Algorithm, AS176, Applied Statistics, 31:93–97, 1982.

[118] B. W. Silverman. Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC, April 1986.

[119] H. M. Singer, I. Singer, and H. J. Herrmann. Agent-based model for friendship in social networks. Physical Review E, 80(2):026113, August 2009.

[120] L. M. Smith, A. L. Bertozzi, P. J. Brantingham, G. E. Tita, and M. Valasik. Adaptation of an ecological territorial model to street gang spatial patterns in los angeles. Discrete and Continuous Dynamical Systems, 32(9):3223–3244, 2012.

[121] L. M. Smith, M. Keegan, T. Wittman, G. Mohler, and A. L. Bertozzi. Improving density estimation by incorporating spatial information. EURASIP Journal on Advances in Signal Processing - Special issue on advanced image processing for defense and security, 2010.

[122] A. Stomakhin, M. B. Short, and A. L. Bertozzi. Reconstruction of missing data in social networks based on temporal patterns of interactions. Inverse Problems, 27(11):115013, 2011.

[123] T. A. Taniguchi, J. H. Ratcliffe, and R. B. Taylor. Gang set space, drug markets, and crime around drug corners in Camden. Journal of Research in Crime and Delinquency, 48(3):327–363, August 2011.

[124] L. Tesfatsion. Chapter 16 agent-based computational economics: A constructive approach to economic theory. In L. Tesfatsion and K. L. Judd, editors, Handbook of Computational Economics, volume 2, pages 831–880. Elsevier, 2006.

[125] F. M. Thrasher. The Gang: A Study of 1313 Gangs in Chicago. University of Chicago Press, 1927.

[126] G. E. Tita, J. Cohen, and J. Engberg. An ecological study of the location of gang "set space". Social Problems, 52(2):272–299, 2005.

[127] G. E. Tita and S. Radil. Making space for theory: The challenges of theorizing space and place for spatial analysis in criminology. Journal of Quantitative Criminology, pages 1–13, 2010. 10.1007/s10940-010-9115-5.

[128] G. E. Tita, K. Riley, G. Ridgeway, C. Grammich, A. Abrahamse, and P. Greenwood. Reducing gun violence: Results from an intervention in East Los Angeles. National Institute of Justice, RAND, 2003.

[129] M. Tomassini, E. Pestelacci, and L. Luthi. Mutual trust and cooperation in the evolutionary hawks-doves game. Biosystems, 99(1):50 – 59, 2010.

[130] Z. Toroczkai and H. Guclu. Proximity networks and epidemics. Physica A: Statistical Mechanics and its Applications, 378(1):68 – 75, 2007.

[131] P. Torrens. Geography and computational social science. GeoJournal, 75:133–148, 2010. 10.1007/s10708-010-9361-y.

[132] J. Tseng, C. H. Lin, C. T. Lin, S. Wang, and S. Li. Statistical properties of agent-based models in markets with continuous double auction mechanism. Physica A: Statistical Mechanics and its Applications, 389(8):1699–1707, 2010.

[133] P. Turchin. Population consequences of aggregative movement. Journal of Animal Ecology, 58:75–100, 1989.

[134] J. Vigil. Barrio Gangs: Street Life and Identity in Southern California. University of Texas Press, 1988.

[135] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. SIAM Journal on Imaging Sciences, 1(3):248–272, 2008.

[136] S. Wasserman and K. Faust. Social Network Analysis: Methods and Applications. Cambridge University Press, 2009.

[137] D. P. Watts and J. C. Mitani. Boundary patrols and intergroup encounters in wild chimpanzees. Behaviour, 138(3):299–327, 2001.

[138] B. Wellman. Are personal communities local? A Dumptarian reconsideration. Social Networks, 18(4):347 – 354, 1996.

[139] K. A. J. White, M. A. Lewis, and J. D. Murray. A model for wolf-pack territory formation and maintenance. Journal of Theoretical Biology, 178(1):29 – 43, 1996.

[140] K. A. J. White, M. A. Lewis, and J. D. Murray. On wolf territoriality and deer survival. In J. Bascompte and R. V. Sole, editors, Modeling Spatiotemporal Dynamics in Ecology, pages 105 – 126. Springer - Verlag, 1998.

[141] Allen Wilhite. Chapter 20 economic activity on fixed networks. In L. Tesfatsion and K. L. Judd, editors, Handbook of Computational Economics, volume 2, pages 1013–1045. Elsevier, 2006.

[142] L. H. Wong, P. Pattison, and G. Robins. A spatial model for social networks. Physica A: Statistical Mechanics and its Applications, 360(1):99 – 120, 2006.

[143] Y. Yang and X. Liu. A re-examination of text categorization methods. In SIGIR '99: Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval, pages 42–49, New York, NY, USA, 1999. ACM.

[144] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for $l_1$-minimization with applications to compressed sensing. SIAM Journal on Imaging Sciences, 1(1):143–168, 2008.