

# TEMPORAL CONSIDERATIONS FOR SPATIAL EVENT PREDICTION

---

A Dissertation

Presented to

The Faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

in Systems and Information Engineering

---

by

Jonathan Fox

2012

## APPROVAL SHEET


This dissertation is submitted in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy  
(Systems Engineering):



---

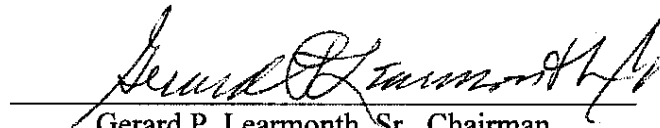
Jonathan Fox, Author

This dissertation has been read and approved by the examining committee:



---

Donald E. Brown, Advisor



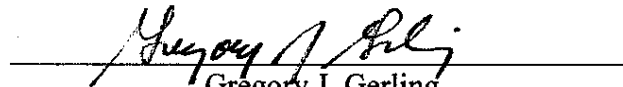
---

Gerard P. Learmonth, Sr., Chairman



---

Matthew Gerber



---

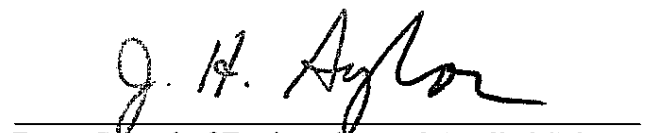
Gregory J. Gerling



---

Brian L. Smith

Accepted for the School of Engineering and Applied Science:



---

Dean, School of Engineering and Applied Science

May

2012

To the families of Captain Robinson, Staff Sergeant Alexander, and Staff Sergeant  
Piper.

## ABSTRACT

A significant amount of research and practice in the law enforcement arena focuses on spatial and temporal event analysis. Although some efforts integrate spatial and temporal analysis, the majority of the previous work focuses on modeling events in one area across time periods or modeling multiple areas within single temporal intervals. This dissertation attempts to close the gap between spatial and temporal analysis by focusing on the challenges of integrating spatial *and* temporal features in criminal site selection problems. Since the criminal environment contains multiple actors intermingled in a spatial region, the challenge for the analyst is identifying patterns *both* spatially *and* temporally based on the criminal’s site-selection criteria. This dissertation uses a three-pronged approach to study (1) the identification and inclusion of temporal distances into a feature-space model; (2) the application of hierarchical modeling for temporal intervals; and (3) the consideration of a random field construct for modeling site-selection event data. We evaluate each approach on both a simulated data set and a sample data set from a small US city. We compare results of the different methodologies using a test statistic based on measuring performance across a space-time surveillance plot. We found that a two-stage analytical approach assists us in identifying when spatial-temporal modeling might offer significant improvements in predictive performance. Additionally, the integration of a hierarchical structure with the feature-space generalized linear model provides a method that accounts for temporal considerations in the criminal’s site-selection process. These temporal features can be “pulse events” that indicate temporary shifts in spatial-temporal patterns or transitions at the discrete temporal interval level that might account for low-level seasonality within the criminal’s site-selection process. Finally the development of an expanded feature-space model provides benefits similar to those of a Markov random field model without the computational demands of the Bayesian construct for hierarchical modeling. These methods for identifying and including temporal information into spatial analysis methodologies offer benefits not only in the law enforcement and military communities, but also to the business and environmental sectors.

## Acknowledgements

When one takes as long as I did to complete a dissertation, the list of people to thank becomes rather extensive. In the interest of keeping my acknowledgements shorter than my literature review, I would like to highlight three groups. First of all, I am very thankful for my advisor, Donald Brown, and my committee members, Gerry Learmonth, Brian Smith, Greg Gerling, and Matt Gerber. Their individual efforts, assistance, and patience assisted me while in residence at the University, and, even more so, while I was working away from the Grounds. Their collective flexibility and support made completion possible as the Army decided to move me away from Charlottesville a second time.

In addition to my advisor and committee, God blessed me with two unique individuals who coached, taught, and mentored me throughout this effort. I am forever indebted to Donald Robinson and Sam Huddleston.

Finally, to my wife, Laura, and children, Savannah, and Gideon. Thank you. After many years, God answered our prayers – we finally “impressed our professors.”

Jon Fox  
Charlottesville, Virginia  
28 April 2012

## CONTENTS

<b>ABSTRACT . . . . .</b>	<b>iii</b>
<b>LIST OF TABLES . . . . .</b>	<b>ix</b>
<b>LIST OF FIGURES . . . . .</b>	<b>xii</b>
<b>CHAPTER</b>	<b>1</b>
<b>1 Introduction . . . . .</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Problem Statement . . . . .	3
1.3 Spatial-Temporal Events . . . . .	4
1.4 Contributions and Organization . . . . .	8
<b>2 Literature Review . . . . .</b>	<b>13</b>
2.1 Spatial-Temporal Events . . . . .	15
2.2 Spatial Analysis . . . . .	16
2.2.1 Spatial Considerations . . . . .	17
2.2.2 Spatial Challenges . . . . .	19
2.2.3 Spatial Applications . . . . .	26
2.2.4 Spatial Analysis of Crime Incidents . . . . .	27
2.3 Temporal Analysis . . . . .	29
2.3.1 Temporal Considerations . . . . .	30
2.3.2 Temporal Challenges . . . . .	32
2.3.3 Temporal Applications . . . . .	36

2.3.4	Temporal Analysis of Crime Incidents . . . . .	37
2.4	Spatial-Temporal Analysis . . . . .	39
2.4.1	Spatial-Temporal Considerations . . . . .	40
2.4.2	Spatial-Temporal Challenges . . . . .	44
2.4.3	Spatial-Temporal Applications . . . . .	45
2.4.4	Spatial-Temporal Analysis of Crime Incidents . . . . .	48
<b>3</b>	<b>Methodology . . . . .</b>	<b>53</b>
3.1	Overview of Methodology . . . . .	54
3.2	Spatial and Temporal Analysis . . . . .	55
3.3	Hierarchical Feature-Space Modeling . . . . .	58
3.4	Of Seasonality and Patterns – Temporal Conditioning . . . . .	62
3.5	Of Indicators and Interactions – Temporal Pulse Events . . . . .	63
3.6	Of Neighbors and Recent Past . . . . .	64
3.6.1	Accounting for Neighbors . . . . .	64
3.6.2	Accounting for the Past . . . . .	65
3.7	Significance Testing . . . . .	66
3.7.1	Receiver Operating Characteristic Curves . . . . .	67
3.7.2	Surveillance Plots . . . . .	68
3.7.3	Area Under the Curve . . . . .	69
3.7.4	Performance Limit Ratio . . . . .	69
<b>4</b>	<b>Method 1 – Of Seasonality and Patterns . . . . .</b>	<b>70</b>
4.1	Formulation . . . . .	71
4.2	Methodology . . . . .	73
4.2.1	Two-stage Analysis . . . . .	73
4.2.2	Hierarchical Modeling . . . . .	77
4.3	Simulated Data . . . . .	80

4.3.1	Introduction of Crime Simulation . . . . .	80
4.3.2	Two-Stage Analysis . . . . .	81
4.3.3	Hierarchical Modeling . . . . .	85
4.3.4	Significance Testing . . . . .	89
4.4	Discussion . . . . .	89
<b>5</b>	<b>Method 2 – Of Indicators and Interactions . . . . .</b>	<b>92</b>
5.1	Formulation . . . . .	93
5.2	Methodology . . . . .	96
5.2.1	Two-stage Analysis . . . . .	96
5.2.2	Hierarchical Modeling . . . . .	97
5.3	Simulated Data . . . . .	99
5.3.1	Modification of Crime Simulation . . . . .	99
5.3.2	Two-stage Analysis . . . . .	102
5.3.3	Hierarchical Modeling . . . . .	104
5.3.4	Significance Testing . . . . .	105
5.4	Discussion . . . . .	108
<b>6</b>	<b>Method 3 – Of Neighbors and Recent Past . . . . .</b>	<b>110</b>
6.1	Formulation . . . . .	111
6.2	Methodology . . . . .	114
6.2.1	Bayesian Hierarchical Modeling . . . . .	114
6.2.2	Expanded Feature-Space Model . . . . .	120
6.2.3	Model Comparison . . . . .	121
6.3	Simulated Data . . . . .	122
6.3.1	Modifications of Crime Simulation . . . . .	122
6.3.2	Two-Stage Analysis . . . . .	122
6.4	Discussion . . . . .	123



<b>7</b>	<b>Crime Event Modeling</b>	<b>125</b>
7.1	Introduction of Dataset	126
7.2	Two-stage Analysis	130
7.3	Seasonality and Patterns	141
7.3.1	Temporal Conditioning	141
7.3.2	Significance Testing	142
7.4	Indicators and Interactions	150
7.4.1	Temporal Pulse Events	150
7.4.2	Significance Testing	151
7.5	Neighbors and Recent Past	154
7.5.1	Data Aggregation and Preparation	155
7.5.2	Significance Testing	157
7.6	Discussion	162
<b>8</b>	<b>Conclusions</b>	<b>164</b>
8.1	Contributions	165
8.2	Future Work	168
8.2.1	Computational Alternatives	168
8.2.2	Vehicle Crash Modeling	169
8.2.3	Other Areas of Future Research	170
<b>VITA</b>		<b>183</b>

## LIST OF TABLES

0.1	Notation and Key Symbols . . . . .	xii
2.1	Types of Mapped Data . . . . .	18
2.2	Components of Cartographic Time . . . . .	31
3.1	Contingency Table . . . . .	68
4.1	Predictor Variables for Simulation Dataset . . . . .	87
5.1	Partial Area Under the Curve for Surveillance Plots . . . . .	107
7.1	Predictor Variables for Crime Dataset . . . . .	128
7.2	Average Nearest Neighbor Index Results – Daily . . . . .	136
7.3	Average Nearest Neighbor Index Results – Hourly . . . . .	136
7.4	Potential Spatial-Temporal Pulse Events – Charlottesville Assaults .	140
7.5	Varying Coefficients for HFSM – Daily Interval . . . . .	144
7.6	Varying Coefficients for HFSM – Hourly Interval . . . . .	146
7.7	Area Under the Curve (AUC) Results for Hourly Intervals . . . . .	148
7.8	Varying Coefficients for Foxfield Pulse Event – HFSM . . . . .	152
7.9	Bayesian Hierarchical Feature-Space Model Development for Long Term Data Study . . . . .	158
7.10	Bayesian Hierarchical Feature-Space Model Development for Short Term Data Study . . . . .	160

## LIST OF FIGURES

2.1	Spatial Categories . . . . .	19
2.2	Three-Dimensional Space of Graphical Displays . . . . .	24
2.3	Sample Results from Brushing and Linking . . . . .	24
2.4	Sample Predictive Surface . . . . .	25
2.5	Temporal Categories . . . . .	31
2.6	First-Order Markov Chain . . . . .	43
2.7	Neighborhood Structures . . . . .	43
2.8	Irregular Lattice with Second-Order Neighbors . . . . .	44
2.9	Undirected Labeled Graph for Simulation Region . . . . .	45
2.10	Space-Time Neighborhood Framework for Modeling Event Activity . . . . .	46
2.11	Original Feature-Space Model . . . . .	49
2.12	Geographic Information System (GIS) View of Feature-Space . . . . .	50
3.1	Overview of Dissertation Methodology . . . . .	56
3.2	Appropriateness of Hierarchical Modeling . . . . .	57
3.3	Sample Space-Time Lattice Structures . . . . .	66
3.4	Flattened Space-Time Lattice for Study Domain . . . . .	67
4.1	Method 1 – Temporal Conditioning . . . . .	72
4.2	Graphical Depiction of Events at Different Temporal Resolutions . . . . .	75
4.3	Sample Daily Prediction Surfaces . . . . .	76
4.4	Feature-Space Model of Temporal Intervals . . . . .	78
4.5	Sample Simulation Run with Weekly Crime Patterns . . . . .	82
4.6	Count of Crimes Per Day of Week in Simulation . . . . .	83
4.7	Autocorrelation Plots – Simulated Crimes . . . . .	84
4.8	Geographic Mean Centers of Temporal Intervals from Crime Simulation. . . . .	86
4.9	Predictive Surfaces for Simulation . . . . .	88
4.10	Surveillance Plot for Simulated Crime Data – HFSM vs FSM . . . . .	90

5.1	Method 2 – Temporal Pulse Events . . . . .	94
5.2	Feature-Space Model with Cultural Events . . . . .	96
5.3	Sample Simulation Run with Pulse Events . . . . .	101
5.4	Time Series and Cross Correlation Plots . . . . .	103
5.5	Geographic Mean Centers of Simulated Crime Clusters . . . . .	104
5.6	Comparison of Threat Surfaces – Simulation with Pulse Events . . . .	106
5.7	Surveillance Plot – HFSM versus Base FSM without Temporal Infor- mation . . . . .	107
6.1	Method 3 – Feature-Space Modeling . . . . .	112
6.2	Directed Acyclic Graph for Bayesian Hierarchical Feature-Space Model	118
6.3	Predictive Threat Surfaces . . . . .	123
7.1	Mapping Steps for Feature-Space Analysis . . . . .	129
7.2	Distribution of Crimes Across Different Temporal Resolutions . . . .	131
7.3	Monthly Average of Assaults Conditioned upon Number of Days . . .	131
7.4	Correlation Plots for Assaults Viewed at Daily Temporal Resolution .	132
7.5	Point Plot of All Assaults within Study Region and Horizon . . . . .	133
7.6	Monthly Distribution of Geographic Mean Centers . . . . .	134
7.7	Daily Distribution of Geographic Mean Centers . . . . .	134
7.8	Hourly Distribution of Geographic Mean Centers . . . . .	135
7.9	Comparison of Hourly Interval Dispersion . . . . .	137
7.10	Time Series Plot of All Assaults . . . . .	138
7.11	Time Series Plot with Zoomed Window for Possible Pulse Events . .	139
7.12	Cross Correlation Results – Assaults . . . . .	141
7.13	Surveillance Plot – Assaults – Daily Interval . . . . .	143
7.14	Comparison of Predictive Surfaces from HFSM – Daily Intervals . . .	145
7.15	Surveillance Plots for Comparison of HFSM and FSM During Periods of High and Low Activity . . . . .	147
7.16	Varying Coefficients for HFSM – Hourly Interval . . . . .	149
7.17	Cross-Correlation Results – Foxfield Races and Assaults . . . . .	151
7.18	Base FSM for Foxfield Races with Incidents for Race Day . . . . .	153
7.19	HFSM for Foxfield Races with Incidents for Race Day . . . . .	154
7.20	Evolution of Spatial Patterns Over a Continuous Temporal Horizon .	156
7.21	Space-time Lattice for Study Domain . . . . .	156
7.22	Undirected Labeled Graph for Study Region . . . . .	157
7.23	Performance Limit Ratio . . . . .	162

**TABLE 0.1: Notation and Key Symbols**

Notation	Description
$\Pr$	Probability
$Y$	Observation of process
$S, s$	Locations of observation
$T, t$	Time of observation
$J, j$	Range of temporal intervals
$\mathbf{X}$	Vector of features
$\beta$	Regression coefficient
$K, k$	Index for number of features
$\alpha_j$	Varying intercept for hierarchical component $j$
$W_t$	Special event time series
$\rho_{wy}$	Cross-correlation of time series
$\omega_s$	Markov random field accounting for neighboring effects
$g_t$	Time varying trend for Bayesian model
$v_s$	Error term for Bayesian model
$\psi_{s,t}$	Spatial-temporal interaction term for Bayesian model
$N()$	Normal distribution

# CHAPTER 1

## Introduction

*It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories instead of theories to suit facts.*

– Sherlock Holmes in *The Five Orange Pips*

This chapter introduces the research conducted for this dissertation beginning with one motivating scenario related to historical spatial-temporal modeling. From the motivating scenario, we derive the problem statement and present the supporting questions that formed the framework for the dissertation. The chapter concludes with an overview of the contributions of the research and the structure found in the remaining chapters of the dissertation.

## 1.1 Motivation

During the summer of 2007, the city of Charlottesville, Virginia, experienced a string of criminal incidents involving similarly attired youth. These incidents included violent attacks in the vicinity of the city’s “downtown mall” and disturbances at major shopping venues. Although none of the events resulted in serious injury, they did disrupt the public’s sense of peace and security. Public officials quickly released statements claiming that the incidents may or may not have been related and that these incidents did not portend an increase in criminal activity (Provence, 2007). For the average citizen, these statements did not necessarily provide the reassurance officials might have intended. Most citizens want to know information that includes the following: How safe is the downtown mall? Does criminal activity increase along the mall during the summertime? What about during special events such as concerts? Does an athletic event on the other side of town make the mall a safer place, or does criminal activity actually increase after the event concludes? Should we avoid the east side of the mall on weekends? Is the parking area on the south side of the mall safe during the weekdays? What about during the evening hours?

Although much work has been done in the geostatistical and environmental sciences community on spatial-temporal prediction problems, the majority of this previous research focuses on the non-intelligent, or physical elements, of spatial-

temporal prediction problems (Jagger et al., 2002; Wickle, 2003; Lawson, 2009). The scenario presented above is one example of spatial-temporal prediction problems that are the result of an actor’s site-selection process. And although the site-selection analytical community has a variety of tools available to answer questions regarding spatial hotspots and temporal patterns of event rates, most of these tools lack a methodical approach for analyzing and identifying spatial-temporal patterns. Integrating temporal factors into current analytical methodologies offers the potential for improved understanding of intelligent site-selection problems and an improved capability to predict future incidents.

## 1.2 Problem Statement

Willie Sutton, an infamous bank robber in the 1930s, replied simply when asked why he robbed banks: “because that’s where the money is.” Although the spatial choice might seem obvious, what is less obvious is Mr. Sutton’s temporal preference for robbing the banks early in the day to take advantage of the morning disorder often present in business openings (FBI, 2010).

The main goal of this research is to improve our understanding of the spatial-temporal patterns found within an actor’s site selection process. Using the definition of intelligent site-selection provided by Porter (2006), we consider processes in which an actor decides to initiate an event at a specific time and at a specific spatial location. For this dissertation, we focus on the criminal as the intelligent actor and use the term introduced by Xue and Brown (2006), criminal site-selection (CSS). We assume that today’s criminal executes a decision process based on preferences associated with the spatial-temporal features of the incident, much like Mr. Sutton’s preference to rob banks (*where the money is*) in the morning (*perceived higher probability of success*). In reviewing the literature related to site-selection



and crime analysis, we find that most work examining site selections as a discrete choice problem focuses purely on the spatial components of an actor’s decision process while assuming that the temporal density remains constant across the study region for all actors. Even though previous research has shown that temporal patterns exist in criminal activity (Rossmo et al., 2005; Groff, 2007), early studies on CSS problems examined data aggregated across temporal intervals or within small two-week or monthly windows (Brown and Liu, 1999; Liu and Brown, 2003; Brown et al., 2004). To account for temporal patterns in the CSS process, we merge the traditional feature-space model with a hierarchical specification to improve current forecasting capabilities. The questions of “how can we include temporal information into feature-space modeling?” and “how can we improve our understanding of discrete choices in space *and* time?” are the motivating questions of the dissertation. The final work should be particularly relevant for law enforcement activities and resource-allocation planning.

### 1.3 Spatial-Temporal Events

We define a spatial-temporal event as a realization of a process in space and time. Following the work of Liu and Brown (1998), we assume that the spatial-temporal event is a realization of a stochastic spatial-temporal process  $\mathbf{Z}$  observed at times  $t_1, \dots, t_m \in T$  at locations  $s_1, \dots, s_n \in S$ . We further define the random variable  $Y(s, t)$  as the observed realization of our spatial-temporal process  $\mathbf{Z}$ . Work by Brown et al. (2004) provides for the likelihood of a criminal event at a given location given a crime occurs:

$$\Pr(Y_{s_i} = 1|X) = \text{logit}^{-1}(\beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}) \quad (1.1)$$

Building on this notation, we label the response  $Y_{s_i} = 1$  for instances where an event occurred at a location  $s$ .  $\mathbf{X}$  is a vector of predictor variables of length  $k$  as used in feature-space modeling, and  $\boldsymbol{\beta}$  is a vector of coefficients that can be used to estimate the criminal’s preferences on the selected features. The model above, and some of the previous work, assumes that the temporal density remains constant over time. In order to identify periods of study that might not hold to this assumption, we first analyze the incident set temporally to identify cyclic changes at different temporal resolutions (Chatfield, 1975). After identifying these distinct temporal resolutions, we use hierarchical modeling to expand the feature-space model with the inclusion of temporal intervals:

$$\begin{aligned} \Pr(Y_{s,j} = 1) &= \text{logit}^{-1}(\alpha_j + \boldsymbol{\beta}\mathbf{X}_s), \text{ for } s = 1, \dots, n \\ \alpha_j &\sim N(\mu_\alpha + \gamma_j, \sigma_{\text{temporal interval}}^2), \\ &\text{for } j = \text{range of temporal interval} \end{aligned} \tag{1.2}$$

This model provides a varying intercept model to examine the changing base likelihood of criminal activity during a specific temporal period. The varying intercept  $\alpha_j$  represents the intercept at each temporal unit  $j$  where  $j$  corresponds to the range of the specified temporal resolution. The multilevel modeling approach allows the analyst to explore spatial and temporal information in a “conditioned” approach in order to hopefully improve both understanding and prediction (Schabenberger and Gotway, 2005). Chapter 4 provides additional details on hierarchical modeling.

Although the hierarchical model allows us to look at the variation in a criminal’s preference for certain temporal intervals periods, we lack the ability to account for how certain temporal events might impact on the criminal’s site selection process regarding a set of feature-space predictor variables in vector  $\mathbf{X}$ . Using the cross-correlation function (CCF) as a tool to determine the relationship between two time

series, we identify temporal windows for further study (Chatfield, 1975).

Identifying potential pulse events enables us to expand the covariate selection process to include a hierarchical component representing the temporal distance to the pulsed events. Changing the notation slightly to reflect this new hierarchical component for a varying-slope and varying-intercept model, we model the coefficients to determine the impact of the temporal proximity to special events on the CSS process:

$$\begin{aligned} \Pr(Y_{s,j} = 1|X) &= \text{logit}^{-1}(\alpha_j + \beta_{j[k]}\mathbf{X}_s), \\ &\text{for } s = 1, \dots, S \\ \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} &\sim N \left( \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix} \right), \\ &\text{for } j = \text{index for the temporal distance to pulse event} \end{aligned} \quad (1.3)$$

Hierarchical modeling provides us a method for including temporal features in the feature-space model. Doing so allows us to account for spatial-temporal patterns within the CSS process process. Equation 1.3 uses a set of features  $\mathbf{X}$  as a vector of length  $k$  for each location  $s$  in order to estimate  $\beta$ , a vector of coefficients that represents the criminal's preference for executing a crime at a location with the selected spatial features. For this dissertation, we assume that the feature vector  $\mathbf{X}$  remains constant over time, but the criminal's preferences for those features adjust based on the temporal proximity to the special event.

Finally, the challenge of looking at the criminal's decision-making as a continuous process adds additional theoretical and computational challenges. Chapter 6 expands the discussions of neighborhood structures and introduces the feature-space model for examining criminal preferences across a continuous temporal horizon. Chapter 6 outlines the assumptions of the MRF and the computational construct

used to estimate the coefficients and to provide insight on the impact of temporal variables within the CSS process.

The considerations for including temporal components in feature-space modeling help provide the main questions answered by this dissertation:

- How do we identify temporal features that might be significant to an actor's decision process?
  - Do temporal patterns related to the actor's decision process exist in the event data?
  - Do other event patterns cause shifts in the criminal's "normal" spatial-temporal decision process?
- How do we include this temporal information in feature-space modeling?
  - How can we modify the feature-space model to account for identified temporal patterns?
  - How can we model the impact of temporal "pulse events" on the criminal's spatial-temporal decision process?
- How do we model an actor's spatial-temporal preferences for site-selection across a continuous temporal horizon?
  - In the feature-space model, can we model the impact of neighborhood patterns to account for local activity on the criminal's spatial-temporal decision process?
  - Can we model the impact of the recent past to account for local temporal patterns on the criminal's spatial-temporal decision process?

## 1.4 Contributions and Organization

This dissertation contributes to the understanding of the spatial-temporal patterns related to criminal site selection (CSS). Contrary to work in geostatistics and disease mapping that focuses on exclusively on spatial *or* temporal patterns found in event data, this dissertation examines the interactions of spatial *and* temporal features that are the part of an actor’s decision process for event initiation. This research develops three methods for considering temporal data as part of modeling and forecasting CSS processes. Previous research has shown that temporal patterns exists in criminal data (Rossmo et al., 2005; Groff, 2007), we merge the traditional feature-space model with a hierarchical specification to improve current forecasting capabilities. Conditioning the base feature-space model with a hierarchical component built at a distinct temporal resolution gives law enforcement personnel a method to account for seasonality in the CSS process.

However, just as crime patterns change in conjunction with a certain amount of seasonality, changes in crime patterns are also associated with the temporal proximity to special events (Cohen et al., 2003, 2007). In this dissertation, we use the cross-correlation function (CCF) as a tool to determine the relationship between crime patterns and a time series accounting for special events within the study region. Incorporating the temporal distance to the special event as a hierarchical component for the base feature-space model assists the analyst with understanding how the temporal proximity to a special event impacts on the CSS process. This method also offers significant improvement for law enforcement resource planning during the temporal windows surrounding these special events.

Finally, considering criminal events within a MRF construct presents an opportunity to combine the spatial-temporal influence of a neighborhood structure with the benefits of the spatial-choice model. We investigate two approaches to

account for the impact of spatial and temporal recency on the CSS process. Exploring the hierarchical feature-space model through a Bayesian framework provides a methodology for analyzing event initiations at higher spatial resolutions while reducing uncertainty in the feature preferences. The second approach uses temporal- and neighborhood-indicator functions with the base feature-space model to rapidly examine the CSS process across a continuous temporal horizon. Although both approaches develop accurate predictions that assist law enforcement personnel with adjusting staffing and asset allocation to meet the most likely patterns of criminal activity, the use of indicator functions with the base feature-space model requires less computation time than the Bayesian approach.

Chapter 2 begins with an overview of spatial and temporal analysis. It considers historical and current research regarding pure-spatial and pure-temporal analysis. A review of integrated spatial-temporal analysis in environmental sciences provides an introduction to the potential applicability of hierarchical modeling and further illustrates the complexities and computational challenges involved with spatial-temporal modeling. Finally, the transition to integrated spatial-temporal analysis allows us to review the theoretical foundation of the discrete choice-problem with applications to CSS problems. It also provides a brief review of feature-space event prediction that serves as an introduction to the proposed methodology.

Chapter 3 introduces the three approaches used for this research. It provides an initial description of the hierarchical-modeling framework for including temporal information in the feature-space model. It also discusses the flexibility of the hierarchical framework to account for seasonality in the CSS process through temporal conditioning and to account for the temporal proximity to special events that affect the criminals' spatial-temporal preferences. Additionally, it reviews the construction of an MRF for modeling aggregated data across irregular lattices. Considering the CSS problem within the construct of an MRF allows us to account for the continuous

evolution of an actor’s decision process over time.

Chapters 4 and 5 provide a more detailed description of the two-stage spatial-temporal analysis and the hierarchical-modeling framework used in this dissertation. Chapter 4 considers techniques for identifying spatial-temporal clustering and conditions the feature-space model on temporal intervals of specific temporal resolutions. This hierarchical feature-space model (HFSM) gives analysts a method to account for seasonal patterns in the CSS process. It also introduces a small, agent-based simulation to create an environment in which we can control the CSS process of criminal agents. Using an agent-based model helps us account for the probabilistic nature of crime while introducing known temporal patterns into the criminal agent’s rule set. We analyze data from the simulation using both the base-feature space model and the HFSM. Using the simulated data set with weekly seasonality, the HFSM provides a 13% improvement over the base FSM without temporal information.

Chapter 5 introduces spatial-temporal “pulse events” as events that change the CSS process for a short duration, or window, of time. We adjust the hierarchical component of the feature-space model to account for the temporal proximity to these pulse events. Modifying the agent based simulation from Chapter 4, we introduce a series of pulse events in the criminal agents’ site-selection process. We build a series of training and tests from the simulation and compare the predictive performance of the HFSM against the base FSM. Against the simulated data, the HFSM provides an 18% improvement in predictive performance against the base FSM without temporal information.

Chapter 6 gives further details on the Bayesian construct for examining spatial-temporal intensity rates across areal units. An examination of an irregular lattice structure of neighbors within a limited spatial-temporal study region provides both a theoretical and practical approach for identifying an actor’s spatial-temporal feature preferences. Additionally, an expanded feature-space model (EFSM) that includes

variables from neighborhood and temporal indicator functions is presented as a computational alternative to the Bayesian construct of the feature-space model.

Chapter 7 provides an application of the proposed methodologies on a crime-event dataset. We model the spatial-temporal patterns of various crime types across Charlottesville using both HFSM and a Bayesian hierarchical feature-space model (BHFSM). A non-nested structure for different temporal intervals provides the most detailed understanding of the changes to the hierarchical model across both space and time. Using data aggregated at the census block-group level for a medium temporal resolution, the BHFSM allows us to model an actor's spatial-temporal preferences within a limited temporal period. Incorporating elements of the feature-space methodology into the hierarchical MRF construct allows us to blend the benefits gained from understanding multiple covariates in the actor's spatial-temporal decision process with the basic elements of geographic recency and spatial dependence found in hotspot modeling. The enhanced feature-space model (EFSM) with neighborhood and temporal indicators provides a computationally efficient and sufficiently accurate alternative.

The final chapter summarizes the contributions of this research. It reviews the dissertation's development of three techniques for consideration when modeling and forecasting spatial-temporal events. Using the CCF as a tool to determine the relationship between two time series, we can identify special events to include as spatial-temporal features in the HFSM to improve current forecasting capabilities. The more general application of the HFSM includes non-nested temporal intervals at varied resolution in order to account for daily, weekly, or other seasonal patterns. From a more theoretical perspective, the development of a BHFSM for CSS problems combines the spatial dependence information gained from a neighborhood structure with the feature-space construct to better identify an actor's spatial-temporal preferences. The combination of hierarchical modeling with the MRF



provides a methodology for analyzing event initiations at higher spatial resolutions. For a more computationally efficient method, we use temporal and neighborhood indicator functions within an expanded feature-space model to examine an actor’s spatial-temporal preferences rapidly with acceptable accuracy. The EFSM provides significant improvement over traditional modeling in terms of providing a decision maker with information about events within a space-time continuum. Across a long study horizon, the EFSM provides a 16% improvement in predictive performance over a model without temporal information. Across a smaller study horizon, the BHFSM provides slight improvement in predictive performance, but the increased computational requirements hinder the application of the BHFSM for “just-in-time” modeling. Understanding a criminal’s spatial-temporal preferences offers a strategic benefit with regard to resource allocation decisions. We expect that hierarchical modeling of criminal events will improve resource-allocation strategies for both crime management and response.

# CHAPTER 2

## Literature Review

*Knowledge is of two kinds. We know a subject ourselves, or we know where we can find information on it.*

– Samuel Johnson

What is a spatial-temporal event? Can we predict criminal or other events using only spatial information? Can we model or forecast crimes using only temporal data? What are the limitations to event prediction using one approach without the other? What other research efforts integrate spatial and temporal information into event prediction? Prior to examining new methods of spatial-temporal analysis, this chapter attempts to answer the above questions and provide a literary foundation for the methodology proposed.

This chapter briefly reviews the mathematical constructs of spatial-temporal processes introduced in Chapter 1 and explores some of the analytical methods used to examine crime patterns and criminal site selection. Section 2.2 reviews some of the fundamental considerations and applications of spatial event modeling. Event modeling falls short, however, when we consider only the spatial aspects of crime. Some cases of temporal analysis of crime date back over a 100 years, but recent theoretical and computational advances offer potential for advancing beyond “point in time” crime analysis and considering temporal patterns as part of the criminal’s site-selection process (Townshley and Pease, 1997; Gorr et al., 1999; Groff and LaVigne, 2002; Ratcliffe, 2002). Section 2.3 provides a look into the challenges of temporal analysis and some of the recent advances in crime studies and other actor-initiated events. After examining the spatial and temporal analytical processes separately, we proceed to look at other fields that consider the integration of space and time for modeling processes in geostatistics, disease mapping, and environmental sciences. Why borrow from the fields of environmental sciences and geostatistics? We believe the focus in these disciplines on modeling environmental events as realizations of a binary space-time process across a long temporal horizon offer benefits in adapting the work on feature-space modeling to better model the temporal considerations of the criminal’s site-selection process (CSS). Section 2.4 also provides more details on CSS and feature-space modeling including an introduction to Markov random fields

and their applicability to crime modeling.

## 2.1 Spatial-Temporal Events

As noted in Chapter 1, we assume that a spatial-temporal event is a realization of a stochastic spatial-temporal process  $\mathbf{Z}$  observed at times  $t_1, \dots, t_m \in T$  and at locations  $s_1, \dots, s_n \in S$ . We further define the random variable  $Y(s, t)$  as the observed realization of our spatial-temporal process  $Z(s, t)$  (Liu and Brown, 1998). We represent the spatial-temporal event mathematically as follows:

$$Y(s, t), s \in D \in \mathbb{R}^d, t \in T \quad (2.1)$$

where  $Y$  is a process observed at a location  $s$  at a point in time  $t$ .  $Y(s, t)$  is a random variable that represents the observed value of a process in both time and space (Chen et al., 2006).

For this work on crime modeling, we expand on some of the initial work of feature-space modeling and view Equation 2.1 as a binary random variable (Liu and Brown, 1998; Brown et al., 2004). Given a crime occurs, we model the likelihood of a criminal event at a given location and time:

$$\Pr(Y_{s,t} = 1|X) = \text{logit}^{-1}(\alpha + \beta_1 X_{s,1} + \dots + \beta_k X_{s,k}) \quad (2.2)$$

Building on the above notation, we label the response  $Y_{s,t} = 1$  for instances where a crime occurred at a location  $s$  at a point in time  $t$ .  $\mathbf{X}$  is a matrix of predictor variables of length  $k$  and width  $S$  that represents each location's feature-space description, and  $\boldsymbol{\beta}$  is a vector of coefficients that can be used to estimate the criminal's preference on the selected features.

Prior to modeling crimes as a spatial-temporal process, many analysts conduct a two-stage analysis of the events. Following the two-stage methodology for this dissertation, we first analyze the spatial components of the crime during a single temporal interval. Second, we analyze the temporal process of crime rates at a single spatial location (Schabenberger and Gotway, 2005). The next two sections examine each stage with particular focus on the considerations, challenges, and analysis of criminal events.

## 2.2 Spatial Analysis

Spatial analysis is the application of statistics to data that contain a spatial element. In the study of events, whether crime, disease outbreaks, or traffic accidents, a geographic location is always present even if this location is not always known. This research focuses on events with a known or reported location and considers spatial analysis as the exploration of “relationships, patterns, and processes” of trends or events that have a geographic component (Dangermond, 2003).

Several early works set the stage for the spatial analysis conducted today. This collection introduces a statistical framework for examining spatial data (Cliff and Ord, 1973; Getis and Boots, 1978; Cliff and Ord, 1981), reviews the integration of cartography and quantitative analysis (Unwin, 1981), and stresses the importance of examining the data visually and statistically (Ripley, 1981). Cressie’s work (1993) collected the efforts of a decade earlier to provide a concise treatment of the three types of spatial data: geostatistical, lattice data, and point patterns. Since this dissertation focuses on criminal events, we will primarily limit our discussion to point processes. However, subsequent sections will review challenges associated with data collected or aggregated at the areal - or regular - lattice level .

To expand on the notation presented in Chapter 1, we re-examine the spatial-

temporal process minus the temporal component. We represent a spatial process mathematically as follows:

$$Z(s), s \in D \in \mathbb{R}^d \quad (2.3)$$

where  $s$  is a location within a space  $D$  of dimension  $d$ . For every location  $s$ , we define  $Y(s)$  as an observed random variable and a representation of the physical process being modeled (Christensen, 1991).

## 2.2.1 Spatial Considerations

### Spatial View

Prior to continuing the discussion of spatial analysis, we first consider the concept of space with regard to analysis and modeling. Most research today conceptualizes space in either a field or object view (Haining, 2009). Environmental studies use the field view to conceptualize attributes that are continuous across a surface. This dissertation, however, uses the object view to model the real world as a series of well-defined spaces represented as points, lines, or polygons. We represent the criminal events in this dissertation as points on a geographic space. However, some of the demographic data used to analyze the criminal's preferences for certain spaces represent point collected data aggregated at the areal or polygon level. We add surfaces to the object view in order to provide volumetric representation of these demographic variables (Demers, 2005). In Chapters 4 and 6, we examine two methods for minimizing the loss of information from the aggregation of data collection. We present the calculated incident rates as threat surfaces representing the probability of an actor choosing a specific location and time for initiating a criminal event.

Within the object view of spatial analysis, we need measurement levels for the attributes related to spatial data. We use three basic classification of data: nominal,

ordinal, and interval / ratio (Unwin, 1981; Rogerson, 2006). Table 2.1 provides an overview of the different types of mapped data. This dissertation focuses mainly on nominal point data representing the presence or absence of crime at a certain space. We use both point and line ordinal data to identify an actor’s preference for certain spaces based on distances to the feature. We also include population and other demographic data aggregated at the areal level and recorded as interval/ratio data.

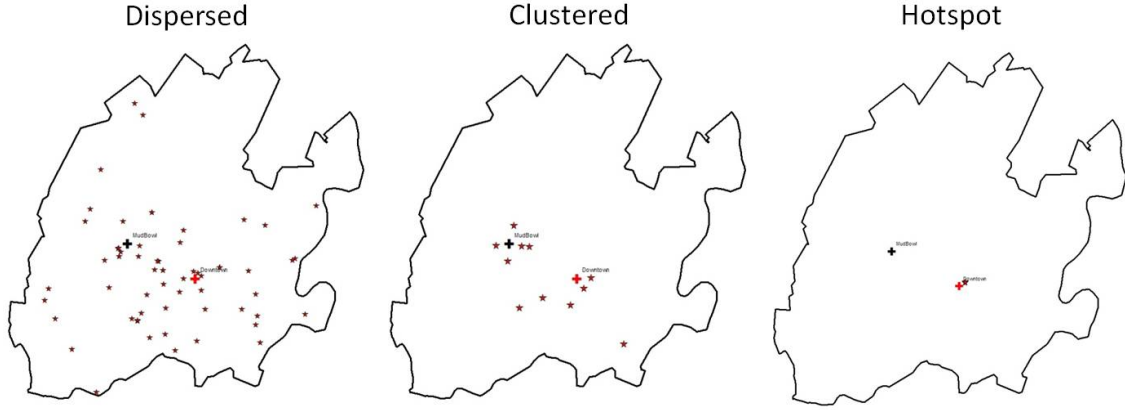
**TABLE 2.1: Types of Mapped Data** (Unwin, 1981)

Measurement Level	Point	Line	Area	Surface
Nominal	Event	Road	Name of Area	Soil Type
Ordinal	Large City	Major Road	County Rankings	Soil Classification
Interval / Ratio	Temperature	Traffic Flow	Per Capita Income	Precipitation in cm

### Spatial Categories of Clustering

Using a given collection of observed crime events, we categorize the spatial distribution using event clustering (Townsend et al., 2000). Event clustering examines the distribution of events within the study area and provides the initial foundation for developing spatial models. Figure 2.1 depicts three spatial categories: dispersed, clustered, and hotspot (Chainey and Ratcliffe, 2005).

A dispersed spatial pattern has no discernable pattern of observed activity at the examined spatial resolution. This does not mean that there are *no* patterns in the examined spatial resolution; rather, it means that the spatial clustering is not significant enough for forecasting and resource allocation. A clustered spatial pattern presents blocks of activity that are more or less focused in specific areas rather than dispersed across the spatial region. A hotspot is an acute clustering of



**FIG. 2.1: Spatial Categories** (Ratcliffe, 2004)

event activity in the vicinity of a specific point within the spatial study region. For example, the number of alcohol-related crimes might peak around a sporting arena or in the vicinity of popular nightclubs. As seen in Figure 2.1, a point mapping of activity within a study region offers a visual depiction of the spatial categories (Chainey and Ratcliffe, 2005).

### 2.2.2 Spatial Challenges

The spatial components of event data answer the most basic question with regard to spatial analysis: “where is it?” (Bivand et al., 2008). However, the spatial components also present some challenges for spatial analysis. Since the incident data and demographic data considered in this dissertation come from various sources, we must consider how variations within the data could affect our efforts for increased knowledge discovery. Dalton (2005) provides a short essay on the challenges of temporal data. We cover this essay in Section 2.3.2. The following sections use Dalton’s construct to consider six challenges of spatial analysis.



## Collection of Spatial Data

Prior to beginning spatial analysis of criminal event data, we ask several questions regarding the spatial data:

- How were the data collected?
- Are the attributes applied to the spatial data continuous or discrete values?
- When were the data recorded?

These questions are important since often users of spatial data may not know the specifics related to the collection of the data. In order to provide spatial data users with a common method for answering these questions and others, the US Federal Geographic Data Committee (FGDC) established the metadata standards for spatial products in the United States. These metadata standards define the characteristics of a given dataset; enable the discovery of spatial libraries for research and analysis; and assist with the evaluation of a dataset's appropriateness for the intended analysis (Goodchild, 1998). Many geospatial communities have not fully adopted the FGDC metadata standards (Dalton, 2005), resulting in challenges for sharing data across different organizations within law enforcement communities. For this dissertation, we use a variety of data sources: the 2000 US Census demographic data, the Topologically Integrated Geographic Encoding and Referencing (TIGER) system, and local law enforcement criminal-incident databases. Using the available metadata, we are able to make the appropriate data transforms to move from different geographic projections and make the appropriate assumptions related to the sensitivity of modifiable areal units (Rogerson, 2006). As mentioned above, some of the demographic data used to analyze the criminal's preferences for certain spaces represent point-collected data aggregated at the areal or polygon level. We will cover techniques related to minimizing the loss of information through dense point

lattices and interpolation of random field neighborhood structures in Section 2.2.2 and later in Chapter 6.

### **Aggregation of Spatial Data**

One of the reasons for the growth of spatial analysis in the last thirty years is the use of geographic information systems (Chainey and Ratcliffe, 2005; Goodchild, 2010). A geographic information system (GIS) allows an analyst to include data from a wide variety of sources. However, each of these sources might record attributes of spatial data at different levels. Although a modern GIS can transform event data from street addresses to grid coordinates based on spatial components, researchers are often left on their own for considering and dealing with the changes of spatial resolution across the numerous datasets. We observe one of the challenges in merging data across datasets when the zoning system used to aggregate the spatial data influences the results of the statistical analysis. Often referred to as the modifiable areal unit problem (MAUP), aggregating point-collected data to a human developed boundary such as city limits or census blocks can lead to false pattern identification (Rogerson, 2006). For this dissertation, we conduct sensitivity analysis by modeling the processes at different aggregations in order to determine whether the relationships hold or if the results are dependent on the areal unit selected for study. Oftentimes, the level of aggregation affects the correlation between the spatial data sets. For instance, at larger areal units, the ability to accurately identify an actor's preferences for certain features is reduced (Cressie, 1993). The other challenge related to data aggregation is commonly called the ecological fallacy which occurs when we make stereotypical decisions from aggregated data (Demers, 2005). Multiple partitioning and indexing schemes are available to assist with facilitate data access, retrieval, and comparison of analysis at multiple spatial resolutions (Yuan, 2009).

## **Retention of Spatial Data**

Spatial data have a complex structure of geographic information, attribute information, and relationship data that is not well suited for relational database management systems. To store this data requires software capable of using specific structures and classes for spatial data and hardware capable of storing the large volume of data that has a spatial component such as satellite and other overhead imagery (National Research Council, 2003). The development of geographic information systems to collect, store, retrieve, transform, and display spatial data equipped researchers and law enforcement personnel with the ability to retain large amounts of spatial data for real-time analysis and long-term studies (Burrough and McDonnell, 1998).

However, these long-term studies produce another challenge for data retention. As spatial information changes in the “real” world, the changes to the spatial representations inside a GIS will probably require a change as well. But how should a GIS retain the original information? The spatial components analyzed in the application portion of this dissertation consist of incident data, demographic data, and feature data. While the incident data occur, are recorded and stored, the possibility exists that when the analyst examines the spatial data, some of the feature data considered by the criminal could be different than that contained in the geographic information system. For example, the street layer and building data collected in 2000 might change over a decade. Zoning conditions for residential and business districts often change based on the needs of the local population. For a law enforcement official conducting long term analysis, the GIS used must be capable of retaining and documenting the changes to specific spatial features such as new construction, street network changes, and zoning area changes (residential versus retail areas). Rollback and transaction recording features of large scale databases are only

recently making appearances inside geographic information systems (Ormsby et al., 2001; Dalton, 2005).<sup>1</sup>

## Visualization of Spatial Data

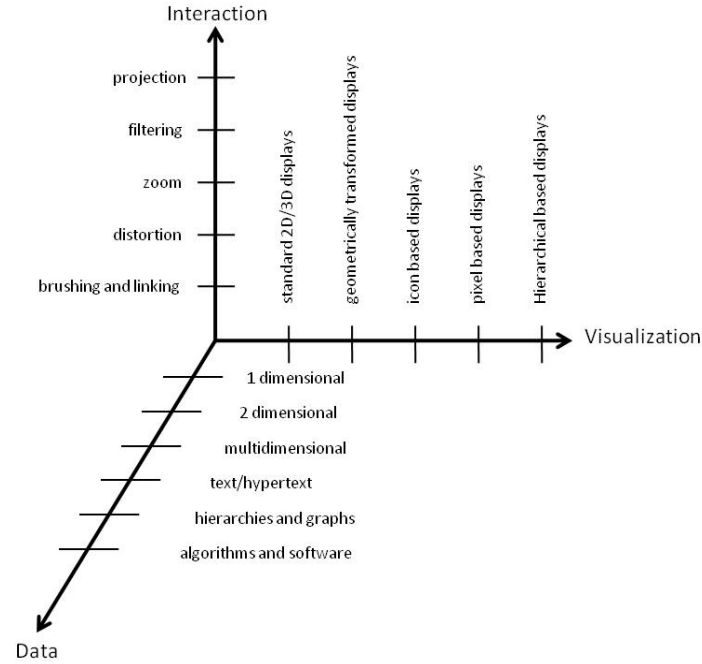
MacEachren et al. (1992) defined visualization as “first and foremost an act of cognition, a human ability to develop mental representations that allow us to identify patterns and create or impose order.” Advances in geographic information systems have improved the analyst’s ability to identify patterns in spatial data. ESRI’s relatively recent (the last 10 years) release of ArcGIS geostatistical analysis offers law enforcement personnel a tool-specific methodology for exploratory spatial data analysis (Johnston et al., 2001). The R open-source statistical software community offers specific spatial classes and packages to provide the well-trained law enforcement analyst a nearly limitless range of possibilities for conducting spatial analysis (Bivand et al., 2008).

When thinking about spatial analysis, we must consider which graphical methods are appropriate for the data. Figure 2.2 depicts a unique plot that shows the potential intersections between data and visualization as a comprehensive classification system of graphical data representations (Keim and Ward, 2003).

The software applications mentioned earlier allow the analyst to produce displays across the three-dimensional space as introduced by Keim and Ward (2003). For initial data investigation, we use a combination of brushing and linking for multidimensional data on standard two-dimensional displays as a visual means for identifying potential criminal event patterns. Final visual displays of predictive surfaces are hierarchical displays projected on a multidimensional view of the study

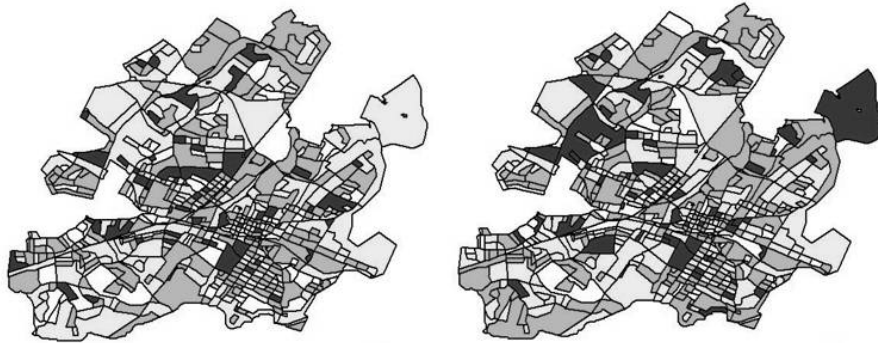
---

<sup>1</sup>Fortunately, even though retaining this large volume of spatial data consumes large amounts of storage, the spatial analysis communities have seen the cost of collecting and storing digital data fall drastically over the last decade (Miller, 2009).



**FIG. 2.2: Three-Dimensional Space of Graphical Displays** (Keim and Ward, 2003)

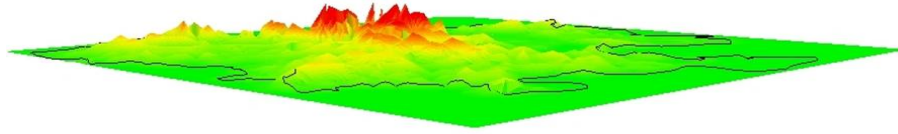
region.



**FIG. 2.3: Sample Results from Brushing and Linking**

The images in Figure 2.3 present a flat two-dimensional view of sample results from brushing and linking. By changing the source of the incident data, we can identify a change in the spatial distribution from the left map to the right map.

The image in Figure 2.4 depicts a three-dimensional view of a criminal-incident predictive surface projected over the study area. Notice the height and color shift



**FIG. 2.4: Sample Predictive Surface**

in the center of the image indicating an increased probability of crime. Section 2.3 provides more details related to visualizing temporal data.

### **Investigation of Spatial Data**

One of the primary benefits in today's geographic information systems is the ability to examine or investigate statistically the spatial changes of event data. One contribution to exploring spatial data is a combination of visual exploration with geocomputation to produce a Self-Organizing Map (SOM). The SOM uses data-mining techniques to compress multidimensional data onto a two-dimensional lattice to produce a series of small planes that enable visual discovery of potential relationships between attributes (Demsar, 2009). A less computationally intensive method would use Tufte's small multiples to compare either varying spatial plots or bivariate matrices across the study space and span (Tufte, 1997). For the initial spatial-data investigation in this dissertation, we use a variation of the methodology introduced by Johnston et al. (2001). Beginning with a visual exploratory data analysis using small multiples to examine spatial patterns over time, we identify areas warranting further investigation using statistical-based analysis. Multiple queries assist in developing an understanding of underlying spatial patterns, correlation, and other interactions in the site-selection process. These queries begin with a simple spatial query and extend to more complex region and relationship queries Yuan (2009). Specific methods related to identifying spatial clustering and exploring

spatial autocorrelation are discussed in Chapters 4 and 5. Section 2.4 discusses the investigation of spatial-temporal data and the challenges inherent to dealing with the more complex spatial-temporal models.

### 2.2.3 Spatial Applications

The literature includes extensive coverage of the use of geographic information systems and statistical methods to study spatial point data. Fotheringham, Brunsdon, and Charlton (2000) describe clustering techniques borrowed from the field of data mining and applied in the context of spatial data. A more general, statistical approach involves testing the hypothesis that spatial event data is clustered against the hypothesis that the data is distributed in a spatially random manner. This clustering of events could reflect aggregation around a point source or along a linear source rather than simply clustering in general. With regard to criminal event data, alcohol-related incidents tend to be clustered within districts zoned for entertainment or drinking establishments. However, clustering of residential break-ins is by definition spatially aligned with residential areas. Gatrell (1994) reviews the lessons learned from early disease-mapping studies that identified the importance of establishing a reference distribution with which to compare the data that are the focus of the study. For instance, patterns of criminal events and disease outbreaks often mirror the spatial pattern of the population density within the study area.

Several techniques are available for determining the presence of clusters compared with the baseline population distribution (Baily and Gatrell, 1995). Roger-son (2006) recommends using proportional sampling, if possible, during the data-collection phase. Lawson (2009) provides for the inclusion of a population variable in the modeling phase. We use Lawson's method in Chapter 6 to account for the criminal's preferences for site selection within spatial areas based on the population

densities. In the context of our problem, this method has the potential to help in determining whether there is an abnormal concentration of attacks — for example, along a particular road or in the vicinity of a particular municipal or religious landmark.

All of these methods are accomplished through the specification of an intensity function of events that is dependent on the distance from the alleged source point. Diggle and Ribeiro (2007) examine spatial statistics focusing on “spatially continuous phenomenon” and Christensen (1991) provides an introduction to the application of linear modeling for spatial data.

### **2.2.4 Spatial Analysis of Crime Incidents**

Of primary concern to this dissertation is the use of spatial analysis to support law enforcement operations. Many law enforcement offices include computational statistics and analysis as part of their community police operations (Townesley et al., 2000; Chainey and Ratcliffe, 2005). At the basic level, analysts predict future events on the basis of the geographic location of past events; to do so, they use a method often referred to as “hotspotting” (Eck et al., 2005; Cliff and Haggett, 1998). The assumption for hotspotting is that the crimes of tomorrow will occur near the crimes of today (Groff and LaVigne, 2002). Early research demonstrates how spatial analysis can identify clusters, or “hotspots”, of automobile theft and assist police with allocating resources into suspected trouble areas (Rengert, 1997). Later research explores the relationship between criminal hotspots with patterns of criminal behavior related to recidivism and spatial limits on criminal behavior (Townesley et al., 2000; Block et al., 2007).

With the increasing availability of geographic information systems and computational capability, more advanced methods use artificial neural networks to predict



the future locations of crime (Olligschlaeger, 1997). Although the predictive performance of these models exceed basic hot spot methods (Olligschlaeger and Gorr, 1997), the use of neural networks fails to provide the analyst with a clear understanding of what spatial features are providing the most impact on the predictions (Groff and LaVigne, 2002). In order to improve on the predictive power of hot spotting while also gaining insight into which spatial features are important, we modify the discrete choice theory from economics (McFadden, 1986) to model the criminal’s site-selection process (Liu and Brown, 1998; Xue and Brown, 2003; Smith and Brown, 2004; Bernasco and Block, 2009).

The commercial sector and academia have contributed several applications related to spatial data analysis in support of law enforcement efforts. The ArcGIS software produced by ESRI is one of the most popular applications for larger organizations (Johnston et al., 2001). While not specific to crime analysis, the S+ SPATIALSTATS module provides law enforcement agencies with a means for conducting exploratory spatial data-analysis and for modeling point patterns as well as areal and geostatistical data (Kaluzny et al., 1997). Similarly, the open-source statistical software R provides researchers and other professionals with capabilities for robust statistical analysis of spatial data (Bivand et al., 2008). Two research initiatives started in the academic community deserve special mention. Developed at the University of Virginia, the Regional Crime Analysis Program (ReCAP) provides local law enforcement agencies with a collection of spatial-analysis and data-mining tools to assist with understanding criminal patterns (Brown, 1998). The National Institute of Justice (NIJ) sponsored research that led to the development of CrimeStat, which allows law enforcement agencies to conduct spatial analysis and “journey to crime” modeling (Levine, 2009). We examined several spatial analysis platforms for this dissertation, but most of the modeling was done using the R and ArcGIS platforms.

As the operational demands for law enforcement agencies increase and resource budgets decrease, decision makers depend more on spatial analysis to support resource allocation and law enforcement policing strategies. Spatial analysis offers benefits to all levels of services in the law enforcement community. By understanding environmental issues related to certain types of crime, law enforcement and community officials have options for changing the physical environment to reduce the preference of criminals for certain areas. Understanding spatial crime patterns allows the law enforcement field to adjust policing patrols both to make certain areas less desirable to criminals and to reduce the local community's fear of certain types of crimes (Chainey and Ratcliffe, 2005). Additional details on these and other approaches with respect to criminal analysis can be found in (Xue and Brown, 2003), (Boba, 2005), (Eck et al., 2005), (Tseloni, 2006), (Gorr and Kurland, 2011), and (Mohler et al., 2011).

## 2.3 Temporal Analysis

The previous section focused on the spatial analysis of events. However, every event happens not only in space but also in time. An alternative analytical approach switches the focus from the spatial components to the temporal patterns that might exist in event series. Ignoring the spatial portion, we can mathematically represent a temporal process as a group of random variables

$$\{Y(t), t \in T\}$$

where  $T$  is a set of time points that can be defined as continuous or discrete. For every time  $t$ ,  $Y(t)$  is an observed value of the process being modeled (Chatfield, 1975).

### 2.3.1 Temporal Considerations

Before discussing temporal analysis, we must first clarify how this dissertation will define *time*. The literature overflows with definitions of time, for example, as a quantity of movement, a dimension, a line without endpoints, and even an illusion. Galton identifies the fundamental components of time as *duration* and *direction*. A segment of duration can be considered an interval. The ordering of these intervals provides a representation of direction. Finally, the consolidation of measured intervals into temporal blocks can represent discrete time (Galton, 1997). This dissertation uses temporal intervals as the initial application of temporal considerations for spatial event prediction.

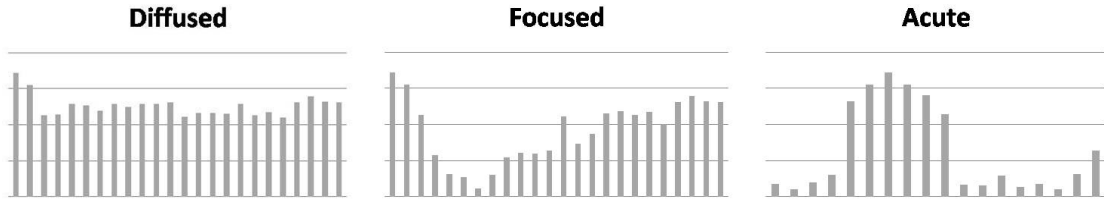
Accepting time as discrete temporal intervals requires additional considerations when integrating space and time in current geographic information systems. Langran’s work stands as one of the core components in the literature regarding time within spatial-analysis tools. One of Langran’s key contributions is the concept of “cartographic time.” Langran defines cartographic time as a method of distilling “the characteristics of time that are essential for representing *spatiotemporality* in the most pragmatic and general fashion” (Langran, 1992). The following table provides a brief overview of the major components of cartographic time.

#### Temporal Categories of Clustering

Much like spatial patterns, temporal patterns can exist in a series of event occurrences. We define three types of temporal patterns: diffused, focused, and acute (Ratcliffe, 2004). A diffused temporal pattern has no discernable pattern to activity at the examined temporal resolution. This does not mean that there are *no* differences in the examined temporal blocks, just that the differences are not significant enough for forecasting and resource allocation. A focused temporal pattern presents

**TABLE 2.2: Components of Cartographic Time** (Langran, 1992)

Component	Synopsis
Fourth cartographic dimension	Integrates with the dimensionality aspects of spatial analysis
Non-interacting time and space	More useful applications for considering time and space being recorded independently
Temporal boundaries	Exist when change occurs
Language of temporal maps	Similar to linguistic concept of tense: present, past, future
Temporal objects and topology	The sequences of states
Facets	World time represents real time while database time represents the time when information was recorded
Constant identify	The challenge of maintaining changing temporal information to the associated entry

**FIG. 2.5: Temporal categories** (Ratcliffe, 2004)

a block of time that is more or less active than the others for event activity. An acute temporal pattern is an increase of event activity within a smaller block of the examined temporal resolution. For example, the number of alcohol-related crimes might peak during sporting events or after the closing times of nightclubs. As seen in Figure 2.5, a histogram of event occurrences for a specific temporal resolution offers a visual depiction of the temporal categories (Chainey and Ratcliffe, 2005).

### 2.3.2 Temporal Challenges

Dalton provides a clear synopsis of the challenges with considering time in the analysis of spatial events. Since most data is not exclusively spatial or temporal, the intersection of these data types provides an area for increased knowledge discovery. The sections below briefly summarize the six challenges of dealing with the temporal component of event analysis (Dalton, 2005).

#### Collection of Temporal Data

How old are the data? When were the data recorded? At what time have the data changed? Questions of time are important since users of GIS often have data without a time attribute. But in reality, all data has a temporal component — either a date of creation or a date of recording. For law enforcement personnel, crime event data must be tagged with a temporal component, even if the tagging of the data is not “timely” (Ackerman, 2006). Much like with the standards for spatial data discussed in Section 2.2.2, the Federal Geographic Data Committee’s metadata standards developed for geographic information systems contain standards for storing temporal data (Dalton, 2005). Metadata offers the opportunity to store the most basic temporal component—date of creation. Additional metadata schemas offer the opportunity to expand upon the temporal aspects of the dataset. However, the recording metadata is not consistent across the global community. In several systems, furthermore, the metadata is not embedded in the spatial dataset which offers the opportunity for data loss when transferring or exchanging among users (Dalton, 2005). Advances in standardized reporting assist in reducing, but not eliminating, errors resulting from varied reporting practices (Mosher et al., 2011).

Part of the challenge with the collection of temporal data is in establishing and maintaining the “temporal consistency” (Guptill, 1995) required for operations

involving multiple agencies and organizations from a broad spectrum of resource providers (Dalton, 2005). As awareness of an event filters through the organizations, each level has an “awareness time” in addition to the event times, observation times, and transaction times defined by Guphill (1995). Guphill’s expansion of Langran’s facets – seen in Table 2.2 – recommends that time consists of three types: event time – when did the event occur?; observation time – when did the collector record the information?; and transaction time – when was the information added or updated in the database? These three times are normally not identical and may be expressed as a time of occurrence or in a relative sense (Guphill, 1995). This dissertation assumes that the time being considered is the event time—when the event occurred. Specific discussion on estimating event time based on observation times can be found in (Ratcliffe and McCullagh, 1998; Ratcliffe, 2000, 2002). For the application portion of this dissertation (Chapter 7), we assume that the temporal value of the criminal incident recorded in police call records is the actual temporal value.

### **Aggregation of Temporal Data**

Another issue stems from the use of data collected from multiple sources. Each source might be recorded at a different point of time. Although modern software applications can transform data based on spatial components, researchers are often left on their own to consider and deal with changes of time across the numerous datasets. Industry efforts for embedding temporal information into spatial layers provide researchers and law enforcement personnel with one mechanism for dealing with this challenge (Dalton, 2005). Aoristic analysis offers another solution for dealing with the challenges of various temporal components associated with multiple series of events recorded at different times. Aoristic analysis allows for the calculation of the probability of event occurrence when the specific time of the event is not known but a time span of when the event occurred is (Ratcliffe, 2000, 2002). As

mentioned above, in this dissertation we assume that the temporal value recorded in police call records is the actual temporal value. We use that temporal value as the basis for aggregating events into specific temporal intervals for hierarchical modeling introduced in Chapter 4. The dynamics of event data require the spatial-temporal researcher to consider scale, quantity, and representation during both the collection and aggregation phases. Accounting for the distribution of event data across the space-time continuum also impacts the storage, or retention, requirements (Goodchild, 2010).

### **Retention of Temporal Data**

As information changes, the changes to the spatial representations inside a GIS will probably require a change as well. However, how should a GIS retain the original information? Dalton stresses several applications in which modification of spatial data conditions without tracking of temporal changes could impact decision making: school choice, police patrol assignments, and legislative representation (Dalton, 2005). Earlier work on temporal databases examined efforts to define appropriate data structures for the time domain. The representational alternatives and data models presented by Snodgrass are varied and some have passed their “time.” The basic discussion of querying, transaction storage, and temporal indexing serve as a cornerstone for today’s applications (Snodgrass, 1992). The “rollback” and transaction-recording features of large scale databases are only recently making appearances inside GIS (Dalton, 2005). However, increased storage requirements for these transaction-recording features complicate the required indexing schemas for efficient searches across the multiple temporal versions of spatial data (Yuan, 2009).

## Visualization of Temporal Data

Even though most of us often consider time in terms of discrete instances, we also recognize that time is continuous. The initial challenge for an analyst is not only to store and retain the attribute of time but also to display it inside the visualization component, whether static or dynamic (Dalton, 2005). After analysis, the remaining challenge is presenting the results in a manner the reader can understand. Several researchers have developed unique approaches to this second challenge: small multiples (Tufte, 1997), movies (Jesse, 2000), and three-dimensional displays (Booth, 2000). Each of these methods offers some benefit to the analyst; however, the ability to incorporate any or all of the above is not native to the spatial-analysis tools commonly available to law enforcement personnel (Dalton, 2005).

An expansion of Tufte's small multiples was the inspiration for a recent effort between the research and intelligence communities that resulted in the development of a GIS "superset." Starlight is a visual-information-analysis system that provides improved visualization not available in every GIS. Starlight can act as an extension for the user's base GIS. Combining data that has textual origins with spatial and temporal data from a GIS, Starlight allows the analyst to examine data validity, qualitative spatial information, and other data within an environment familiar to GIS users (Rex and Rasmussen, 2000). As visually impressive as the Starlight application might be for examining temporal changes across spatial areas, the computational and other resource requirements are beyond the capability of most users.

Much like Starlight, WebTAS is another government-sponsored software tool that not only replicates the basic capabilities for a GIS, but also includes the ability to examine temporal, spatial, and entity information (Jesse, 2000). WebTAS does have a downside for the hardcore GIS user, however; the focus for WebTAS is more on the fused analytical product and not the purely spatial functionality. Current



research has contributed several additional add-on and integrated components to assist analysts with investigating patterns in both space and time (Morgan, 2009; Hadlak et al., 2010; Shaw, 2010; Gorr and Kurland, 2011).

### **Investigation of Temporal Data**

Another component some GIS tools lack is the ability to examine or investigate statistically the temporal changes of the spatial information. Several areas of study (target recognition, traffic monitoring, and wildlife management) require a correlated spatio-temporal model. Studying processes that are jointly dynamic in time and space require not only this correlated spatio-temporal model but also the aforementioned retention challenge (Dalton, 2005). Continuous improvements in the analytical capabilities of geographic information systems, in addition to the data storage and manipulation requirements of temporal information, will benefit research and application. Improvements in spatiotemporal indexing schemes in conjunction with refined temporal data models assist analysts with understanding event dynamics and correlation (Yuan, 2009).

### **2.3.3 Temporal Applications**

There has been some unique coverage in the literature of time series analysis with respect to violent events. Work by Midlarsky et al. introduced the examination of the temporal aspects of terrorism and initiated an increasing focus on data collection and analysis (1980). Mickolus's work and commentary presents considers datasets and analytical techniques on the basis of the selected data (1987). Enders and Sandler have consistently published analysis based on international trends (2000), patterns (2002), and the impact of terrorism (2006). Their work on transnational terrorism serve as a template for this dissertation in examining temporal issues of

criminal activity.

The use of temporal analysis is not limited to the study of violent events, however. The benefit of temporal analysis can be seen in environmental research and economic studies. Combining regression models with time-series models improves multiseason forecasts for hurricane activity. Using Monte Carlo simulations to generate sea surface temperatures, researchers sample from that posterior distribution to forecast hurricane counts across a given temporal resolution (Elsner et al., 2008). The use of Bayesian regression minimizes the disadvantage of the reduced precision in historical data to more accurately forecast hurricane landfalls six months in advance (Elsner et al., 2006).

### **2.3.4 Temporal Analysis of Crime Incidents**

Of primary concern to this dissertation is the integration of temporal considerations with spatial analysis in support of law enforcement operations. A significant amount of research activity focuses on analyzing temporal elements of crime data to provide information on the behavior and patterns of criminals (Townesley and Pease, 1997; Townesley et al., 2000; Ratcliffe, 2000; Clark and Eck, 2005). The routine-activity theory explains that crimes occur at certain times and places that match the criminal's and the victim's daily routines (Rossmo et al., 2005). Additional research examines the impact of seasonal weather patterns on criminal behavioral patterns (Jacob et al., 2007). A motivating factor within a large portion of crime research is improving resource allocation within police departments on the basis of temporal forecasting (Gorr et al., 1999; Groff and LaVigne, 2002).

Creating a temporal study of event data requires the collection of location, attribute, and time data (Langran, 1992). Using time-series methods for forecasting crime rates, researchers have shown improvements over the basic police practice of

“yesterday” forecasting (Deadman and Pyle, 1997; Gorr et al., 2003). Short-term forecasting allows crime analysts to predict more accurately crime rates one month ahead using smoothing algorithms with seasonality. The choice of a smoothing parameter is important in crime-series forecasting because while a small parameter will produce estimates that are too erratic for resource allocation, a larger parameter might miss important temporal pattern shifts (Faraway, 2006). Using a rolling-horizon design (Makridakis et al., 1982), one study examines crime across a 36-month temporal horizon with a variety of forecast methods, most of which use the previous 60 days to predict against the next 30 (Gorr et al., 2003). Although this dissertation does not examine the application of smoothing algorithms, it does incorporate the use of seasonality indexes into the feature-space model in Chapter 3.

To enhance event prediction, other research considers temporal data as a temporal feature. Temporal features are an attempt to capture temporal information for modeling and predicting criminal activity. Temporal features can include time-of-day, event time, and relational time. Kerchner’s examination of temporal features begins with a discussion of temporal resolution. If the goal is to forecast daily operations at a police department, examining annual crime patterns may not provide the most useful information. However, once the level of resolution has been selected, the temporal data can be analyzed as a component of “feature-space.” The feature-space methodology allows the analyst to include temporal information—not just as a component of time series-analysis, but rather as a component of the multidimensional criminal site selection process (Kerchner, 2000). In Chapter 5, we expand the feature-space model to focus on more than a specific temporal interval. Using a hierarchical framework, the analyst can now examine how the temporal distance from a select temporal reference point affects the criminal site selection process. The concept of temporal features allows the analyst to move away from purely spatial or temporal analysis and closer to an integrated analysis.

As a means of examining criminals’ temporal preferences, analysts take “snapshots” of time in order to create temporal criminal hotspots. The snapshots vary in temporal resolution from hours to days to months or seasons. Many researchers use event density maps as temporal snapshots. These temporal surface density maps may then be compared to similar historical snapshots in order to assist with identifying temporal trends or making predictions for upcoming time periods (Eck et al., 2005).

However, since data are not exclusively spatial or temporal, the intersection of these data types provides an area for increased knowledge discovery. The following section reviews several alternative methods for analyzing spatial-temporal data.

## 2.4 Spatial-Temporal Analysis

Spatial-temporal analysis is the application of statistics to data that contain both spatial and temporal elements. In the study of events, the spatial-temporal components are always present even if they are not known. The integration of space and time information into a spatial-temporal data model allows us to pursue the question of “what happens where when?” (Bivand et al., 2008). As noted in Chapter 1, we represent a spatial-temporal event mathematically by adding a temporal index to the spatial process of Equation 2.3 such that

$$Y(s, t), s \in S \in \mathbb{R}^d, t \in T$$

where  $Y$  is a process observed at a location  $s$  at a point in time  $t$ .  $Y(s, t)$  is a random variable that represents the observed value of a process in both time and space (Chen et al., 2006).

### 2.4.1 Spatial-Temporal Considerations

#### Criminal Site Selection

Site-selection is the process by which actors select the time and space to execute events on the basis of their preferences (Porter, 2006; Xue and Brown, 2006). Early research on site-selection defines the actor's site selection process as one of spatial choice. Spatial choice models assume an actor will select a site (e.g., for migration, retail establishment, or criminal event) on the basis of the perceived utility or the worth of that site from a set of alternatives (Ewing, 1976; McFadden, 1986). This view of behavioral geography assumes that spatial-point processes involving actors are based on the actors' mental processes and perceptions (Burnett, 1976).

More recent work in criminology views this spatial choice as "rational criminal theory." Rational criminal theory assumes that individuals have specific reasons for committing a crime at a certain time and a certain location (Clark, 1980). By examining the historical criminal activity data within a spatial region, we can discover patterns that might indicate criminals' preferences for executing crimes at certain locations (Brantingham and Brantingham, 1984). Spatial choice models assume an actor will select a site (e.g., for migration, retail establishment, or criminal event) based on the perceived utility, or worth, of that site from a set of alternatives (Ewing, 1976; McFadden, 1986). The use of spatial choice models nests well within the rational criminal theory since it assumes that spatial point processes involving actors are a result of the actors' mental processes and perceptions (Burnett, 1976). Consider a criminal who wants to steal a car. Will he choose a parking garage at the center of town with restrictive traffic flows or will he choose the mall parking lot near a major freeway on the outskirts of town? Previous work has shown that the car thief will take the car from the mall since features surrounding a location are as critical as the location itself (Rengert, 1997). Brown et al. (2001) showed

that data mining previous criminal events provides insight to what spatial features might be considered by a criminal in selecting a location to commit a crime. We define this set of spatial considerations to be the feature-space. An expansion of this rationality consideration for terrorist site-selection can be found in Brown et al. (2004) and Enders and Sandler (2006).

This dissertation expands on the spatial choice problem to examine the impact of both geographic and temporal features on the criminal’s site-selection process. The preferred location for an individual to execute an event might change from populated to unpopulated areas and could change depending on the time of day or the day of week (Rossmo et al., 2005). Rather than using a latitude and longitude to describe each location in a study region, we use spatial distances to environmental features — such as schools, streets, or stadiums — and spatial representations of social demographics — such as population, percent rental properties, and household income — to examine which locations are preferred by criminals for certain types of crimes (Bannatyne and Edwards, 2003; Liu and Brown, 2003; Huddleston and Brown, 2009).

## **Random Fields**

This dissertation’s third effort is in exploring criminals’ temporal considerations for initiating crimes at specific times and places. Although much theoretical work has been done on the use of Markov random fields for geostatistics, the application within the construct of the CSS problem has been limited. Markov random fields (MRFs) for spatial-temporal events specify that a variable is only dependent on its neighborhoods and its previous temporal state. This dissertation explores the use of the MRFs as a means of modeling event data. A review of neighborhood structure forms the basis for understanding this modeling application for both marked-point and areal processes within a specific spatial-temporal zone. Reviewing the theo-

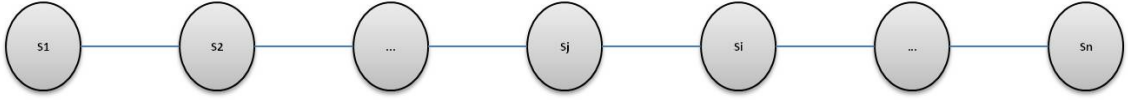
retical and computational work from disease-mapping and environmental studies, we examine how a Bayesian hierarchical model (BHM) might apply for modeling criminal site-selection problems. The BHM reviewed in Chapter 6 includes elements of feature-space information in a logistic regression model with an auto-regression on the state of the neighboring locations across a regular or an irregular lattice at discrete temporal intervals.

### Neighborhood Structure

To examine a spatial random field, we specify a neighborhood system within a lattice structure. Following the work of (Besag, 1974; Rue and Held, 2005; Marin and Robert, 2007), we define a random field as a set of conditional densities across a neighborhood system. For the purpose of this dissertation, the lattice consists of irregular regions based on administrative boundaries with discrete variables measured across specific temporal blocks. Given a lattice  $I$  of sites, we can consider a site  $i \in I$ . We use  $\sim$  to denote the neighbor relationship, and we force the relationship to be symmetric such that if  $i \sim j$  then  $j \sim i$  and  $i$  is not a neighbor of itself (Marin and Robert, 2007). We define a site  $s_i$  to be a neighbor of  $s_j$  if the conditional distribution of  $Y(s_i)$  depends on  $Y(s_j)$  for  $j \neq i$  (Cressie, 1993). The neighborhood set of site  $i$  is now defined as

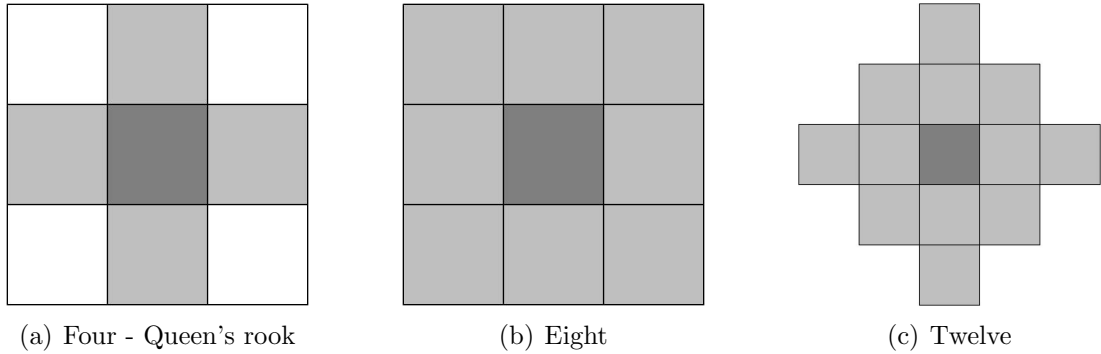
$$N_i \equiv \{j : j \text{ is a neighbor of } i\} \quad (2.4)$$

and a clique is the set of sites that are neighbors of each other. At the most basic level of a Markov field—a one dimensional Markov chain—it is easy to show the neighborhood relationships such that the  $\Pr(y_i|N_i)$  is dependent upon the value at  $y_j$  if  $i \sim j$  (Besag, 1974). Figure 2.6 depicts this one-dimensional structure. A collection of sites with an established neighborhood specification generates a class



**FIG. 2.6: First-Order Markov Chain** (Besag, 1974)

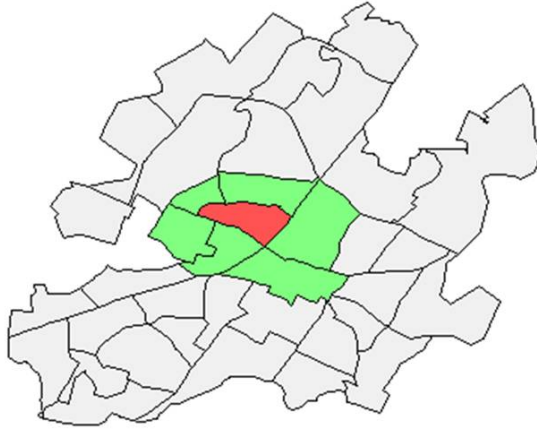
of stochastic schemes called Markov random fields (Besag, 1974). In order to take advantage of the Hammersley-Clifford theorem concerning local and global Markov properties and define the joint distribution, we need only to identify the probability structure for each of the sites along with these neighborhood relationships (Besag, 1974). For spatial statistics, the use of a regular lattice structure matches well with the accepted practices related to spatial sampling. Figure 2.7 depicts several common neighborhood structures used in imagery analysis and geostatistics for a regular lattice structure (Higdon, 2007). For this dissertation, we focus on using



**FIG. 2.7: Neighborhood Structures**

an irregular lattice such that the second-order neighbor structure is based on edge or vertex adjacency. We know that most administrative boundaries create a region with shared borders (Rue and Held, 2005), so we concentrate our MRF exploration at the United States Census block-group level. Figure 2.8 presents a study region using block-group consolidation with the second-order neighbors highlighted. Once we have an established neighborhood structure, the next step in preparing to model spatial-temporal events using an MRF methodology is creating an undirected graph



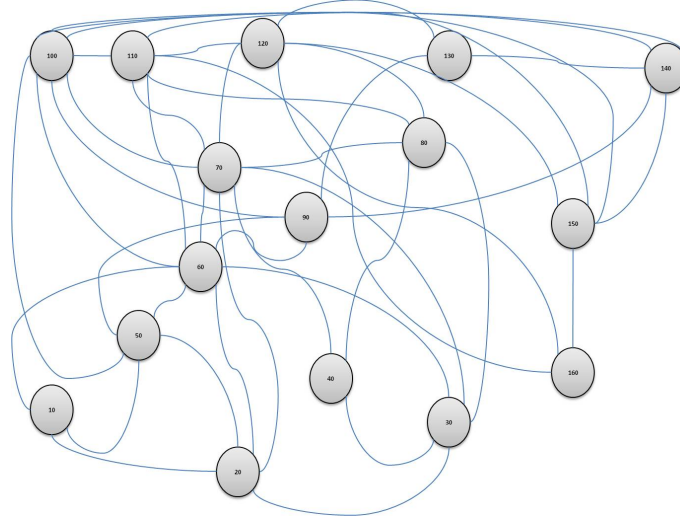


**FIG. 2.8: Irregular Lattice with Second-Order Neighbors.** An example of the study region used in Chapter 7 depicting the lattice structure based on census block-groups. In this figure, the red block-group is the focus area with the green block-groups depicting the second-order neighbors.

(Rue and Held, 2005). The undirected graph with nodes from an irregular lattice allows us to represent the conditional independence structure required for a Gaussian MRF. Following the notation of Rue and Held, we introduce an undirected graph  $\Gamma$  as a tuple with  $\Gamma = (V, E)$  where  $V$  is the set of nodes and  $E$  is the set of edges  $\{i, j\}$  where  $i, j \in V$  and  $i \neq j$ . Figure 2.9 depicts an undirected labeled graph of the study region used in the simulation portion of this dissertation. Constructing this undirected graph with a defined neighborhood structure sets the conditions for examining the formal definition of an MRF in Chapter 6.

### 2.4.2 Spatial-Temporal Challenges

Combining spatially referenced explanatory variables with the MRF probability structure can yield poor estimates of the regression coefficients and produces computational challenges related to multicollinearity (Lawson, 2009). We can remove some of the explanatory variables to reduce the impact of the intercorrelation, especially if the variables demonstrate spatial correlation (Kutner et al., 2004). This



**FIG. 2.9: Undirected Labeled Graph for Simulation Region.** The graph does not represent the spatial layout of the study region but rather shows 16 nodes and the relationships built upon shared borders for a total of 84 neighbors.

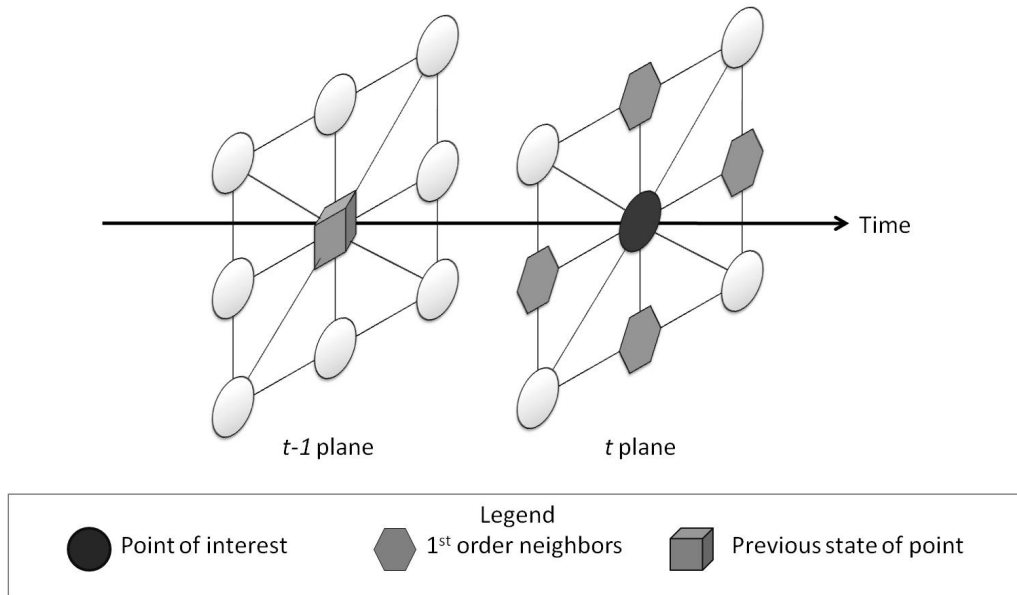
dissertation does not fully explore feature selection but rather refers the reader to more intensive reviews found in (Riese, 2001; Smith and Brown, 2004; Huddleston and Brown, 2009).

### 2.4.3 Spatial-Temporal Applications

As noted earlier, the use of temporal analysis is not limited to the study of crime events; similarly, the use of spatial-temporal modeling extends beyond crime analysis. A large portion of the work on spatial-temporal analysis is seen across the environmental sciences, from weather patterns to disease outbreaks. Several works on modeling weather patterns start with the two-stage analysis methodology proposed by Schabenberger and Gotway (2005). Using the two-stage approach for temperature prediction, we first examine temporal patterns at one spatial point to identify patterns of trend or seasonality. Including the temporal patterns into a spatial Gaussian random field model shows performance benefits similar to other purely spatial correlation models but also includes the benefit of modeling season-

ality (Benth et al., 2005).

The benefit of spatial-temporal modeling is also seen in using Poisson space-time models for modeling hurricane activity in the North Atlantic. Using hurricane counts aggregated across a distinct regular lattice, the truncated Poisson space-time autoregressive (TPSTAR) model incorporates spatial and temporal information to improve annual forecasting along known hurricane tracks (Jagger et al., 2002). For the crime study in Chapter 7, we use the space-time model constructs of the TPSTAR model related to neighbor associations and first order autoregressive selection but substitute a binary function to represent the results of the criminal's decision to initiate or not initiate a criminal event at a certain time and place. Figure 2.10 depicts the space-time model construct for modeling event activity as a function of previous state and neighborhood activity. Several works on disease mapping use an



**FIG. 2.10: Space-Time Neighborhood Framework for Modeling Event Activity** (Jagger et al., 2002)

MRF specification to model disease outbreaks and spread in space and time. Using logistic regression across areas on a regular lattice framework, researchers added a regression and temporal component to one of Besag's models to accurately model

outbreaks of pine beetle infestations in North Carolina (Zhu et al., 2006). In Section 7.1, we mimic the study’s use of data aggregation with the application of thresholds for determining a binary response in administrative areas.

Another study using random fields models spatial distributions of hickory trees in both continuous and discrete settings. Using a true Poisson cluster process as the basis of the random field model fails to consider other potential variables that affect the distribution, however, we see the benefit of using the discrete case for dealing with aggregated data such as crimes within census blocks (Wolpert and Ickstadt, 1998). We may also use an MRF construct within an autologistic regression model. The linear regression model provides a method for incorporating potential explanatory variables while the autologistic portion accounts for the spatial dependence. Incorporating a temporal component within the autologistic portion assists in modeling disease outbreaks and spatial distributions of ecological species (Huffer and Wu, 1998; Zhu et al., 2006). The use of Akaike’s information criterion (AIC) for model selection assists with finding the right balance of explanatory variables without over fitting (Zhu et al., 2008).

Hierarchical modeling supports efforts in integrating the multilevel feature-space modeling across the irregular space of the study area (Rue and Held, 2005). Rue and Follestad’s work (2003) reviewed the basics of Gaussian MRFs and presented several case studies across multiple disciplines. Space-time fields using a Bayesian approach offer potential solutions for reducing the computational challenges that come with modeling large spatial-temporal datasets (Banerjee et al., 2004; Le and Zidek, 2006). Of three approaches to inference and modeling of binary data across a lattice, pseudolikelihood provides the easiest implementation and fastest performance on medium-sized lattices (  $n < 1000$  ). Bayesian and traditional maximum-likelihood approaches perform well on smaller spatial regions with little spatial dependence (Hughes et al., 2011).

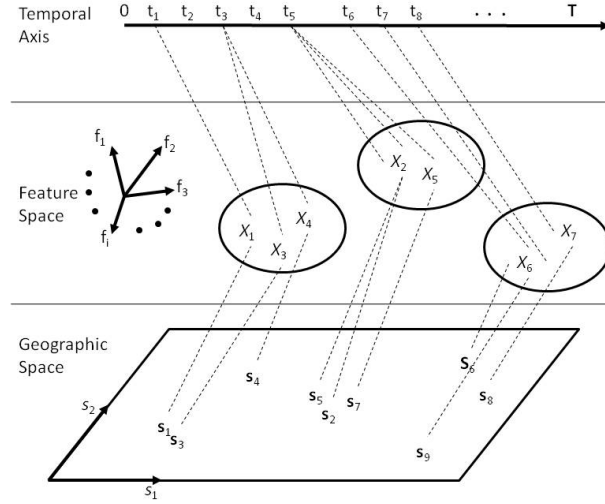
The feature-space model introduced in Section 2.4.4 has application outside crime-event modeling as well. Extracting feature-space information from over 500 historical events, researchers used spatial-preference, logistic-regression, and merged approaches to adapt the feature-space model for predicting the likelihood of terrorist events. Given available spatial and demographic data, analysts are able to broaden their understanding of the site-selection process used by terrorists (Brown et al., 2004). For this dissertation, we improve our ability to model criminal events with temporal-patterns by adapting the spatial logistic-regression model to include temporal features such as temporal intervals and temporal distances to specific events.

## 2.4.4 Spatial-Temporal Analysis of Crime Incidents

### Feature-Space Modeling

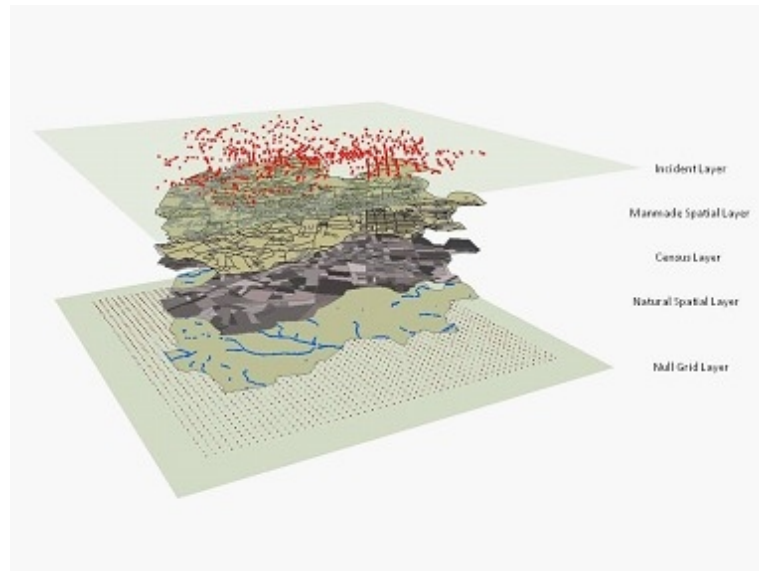
This dissertation continues the work related to feature-space modeling of criminal events. Similar to the work on economic spatial choice (McFadden, 1986), it assumes that criminals choose to execute crimes at specific points and specific times and that their choices are based on a spatial-temporal analytical process (Brantingham and Brantingham, 1984). The feature-space approach includes cultural information obtained through census data with distances to significant geographic areas as predictor variables for modeling event likelihood. Significant geographic features include man-made features—such as buildings, streets, police stations, or banks—and natural features—such as rivers or mountains. These variables, or “features”, help model an actor’s target preferences and enable analysts to see a wide variety of *possible* event locations, not just past locations. This feature-space model is then used to derive spatial clustering, or hotspot, information that is based on past events and augmented with the spatial and cultural features obtained from the empirical analysis. Figure 2.11 depicts the original feature-space model developed by Liu and

Brown (1998). Each event occurs within geographic space is related to a set of spatial or cultural features, and each occurs along a time axis.



**FIG. 2.11: Original Feature-Space Model** (Liu and Brown, 1998)

While Figure 2.11 depicts the theoretical construct, the advances in geographic information systems enable the practical application of feature-space information to modeling crime data (Brown and Liu, 1999). Since our initial focus is on criminal events as a point process, we create a null grid layer across the study region to represent the lack of events within the space (Chawla et al., 2001; Liu and Brown, 2003; Brown et al., 2004). We measure spatial distances from both the null grid and incident layer to various features within the study region. Demographic values are assigned to points on the basis of the containing or nearest polygon's values. Figure 2.12 provides an expanded view of the GIS layers used to model feature-space. Using the feature-space methodology, we estimate crimes at the census-block level in two-week intervals with mixed models. By reducing the temporal horizon to two weeks for model building and forecasting, we capture changes in the temporal transition patterns of the criminal site-selection process (Brown and Liu,



**FIG. 2.12: Geographic Information System View of Feature-Space**

1999; Liu and Brown, 2003). To identify distinct spatial choices among an unknown number of criminals, we use cluster analysis to identify areas of preferences for certain individuals or groups of criminals. Within a two-month temporal window, a Gaussian mixed model performs as well or better than traditional kernel models (Brown et al., 2001; Xue and Brown, 2003, 2006). Hierarchical correlation clustering also benefits analysts focused on exploring linkages between known or unknown groups and observed crimes. This methodology groups similar variables and marks the variables noted as “interesting” to represent each data cluster. Recent work indicates that this method is more robust and has better predictive capabilities than traditional hotspotting models (Smith and Brown, 2004).

Multinomial-choice modeling helps identify a narrowed data set chosen by the decision maker to increase the utility of the developed model (Smith and Brown, 2004). Reducing the modeling scope can make significant insurgent attacks (those causing large numbers of casualties) stand out. In addition, by identifying choice-modeling patterns, an analyst can identify events linked to a specific group (Hudleston and Brown, 2009). The analyst can also use predictor variables—such as

incident relation to specific features or temporal events—to help build the model. With this information and additional background knowledge of the target groups involved, predicting insurgency attacks or criminal activity can be simplified.

### **Other Spatial-Temporal Crime Modeling**

A significant amount of research focuses on analyzing crime data in space and time to improve understanding of criminal patterns and allocation of law enforcement resources. Since resource allocation is a concern for law enforcement personnel, several methods have been developed to analyze crime data at the precinct or patrol level. Using aggregate crime counts within each spatial unit, researchers condition their crime forecast using traditional time series methods with crime data aggregated either by precinct or census blocks or across a regular grid cell draped across the spatial region (Gorr et al., 1999; Cohen et al., 2003; Gorr et al., 2003). In Chapter 4, we expand the feature-space model with a hierarchical framework based on temporal conditioning. Adding the hierarchical framework to the feature-space model enables analysts to better understand the impact of temporal intervals on the criminal site-selection process.

For other studies at the precinct or patrol level, analysts have correlated specific weather events and cultural events with crime rates in efforts to enhance forecasts of criminal activity (Cohen et al., 2007). Chapter 5 introduces a methodology for identifying the impact of special events on the criminal site selection process. Applying the hierarchical framework to the feature-space model on the basis of temporal distances to special events assists us not only with correlating special events to crime patterns but also in understanding the temporal shifts in criminal site-selection with regard to these cultural events.

Building on the theory of repeat victimization, researchers model the time courses—or temporal distance between incidents—for spatial hotspots of crime ac-



tivity to identify *near repeats*. The term *near repeats* describes the occurrence of similar crimes happening within a short temporal window near a previous crime location (Townsley et al., 2000). Chapter 6 expands the base feature-space model to include indicator functions as one method for accounting for near repeats within the criminal site selection process.

# CHAPTER 3

## Methodology

*To make sense of an observation, everybody needs a model ... whether he or she knows it or not. It is difficult to imagine another method that so effectively fosters clear thinking about a system than the use of a model written in the language of algebra.*

– Marc Kery

Building upon previous work in spatial and temporal crime analysis and recent advances in geostatistics, this chapter introduces this dissertation’s three primary approaches for exploring temporal considerations in the analysis of criminal site-selection (CSS) problems. First, we combine elements of hierarchical modeling with the feature-space model to examine temporal patterns in the CSS process. Next, we modify the hierarchical framework by replacing temporal intervals with temporal distances to special events in order to account for the impact of these special events on the CSS process. Finally, we borrow from studies on disease-mapping and incorporate elements of the feature-space model into a Bayesian hierarchical model to examine the CSS process across a continuous temporal horizon. For the continuous temporal horizon analysis, we also borrow from studies on leading indicators and neighborhood patterns to add indicator functions to the base feature-space model. We conclude the chapter with a discussion of measuring model accuracy using three performance measures: the surveillance plot, area under the curve (AUC), and the performance-limit ratio (PLR). We use each measure to examine how including temporal information to account for temporal considerations in the CSS process improves the feature-spaces model’s predictive performance across the study region and study horizon.

### 3.1 Overview of Methodology

Just as models can help us make sense of observations (Kery, 2010), maps can help us make sense of our journeys. Figure 3.1 provides a pictorial map of the methodology used in this dissertation. Beginning with the model formulation of event occurrence at a certain time and place, we apply traditional spatial and temporal analysis to the event space horizon in order to identify patterns of spatial clustering that occur within focused temporal periods. We then explore the temporal horizon

to identify temporal intervals with spatial clustering. Identifying the appropriate temporal resolution for intervals with spatial clustering sets the condition for modeling the CSS process with temporal conditioning. Temporal conditioning places the basic feature-space generalized linear model (GLM) into a hierarchical framework in order to account for seasonality or other temporal patterns that might be part of a criminal actor’s site selection process. For acute temporal intervals that accompany spatial shifts in event hotspots, we introduce the idea of temporal “pulse events.” The temporal pulse event methodology continues to use the hierarchical framework around the base feature-space model but rather than examining patterns of seasonality, we examine how temporal distances to or from special events impacts on the criminal’s propensity for executing events at certain times and locations.

The final approach considers the challenge of including temporal information as part of the CSS process across a continuous temporal horizon with a diffused temporal pattern and a dispersed spatial patterns. For the final methodology we transition from the traditional point process approach used for feature-space analysis and model criminal events within spatial regions. By adapting theoretical and computational work from disease mapping and environmental studies, we develop a Bayesian hierarchical feature-space model for the criminal event prediction problem in order to examine both parameter estimation and predictive inference. For all three methodologies, we compare model performance by examining the performance of the hierarchical models against the base feature-space model using the surveillance plot, the AUC, and the PLR.

## 3.2 Spatial and Temporal Analysis

Schabenberger and Gotway (2005) introduced three basic lines of analytical approach for dealing with spatial-temporal processes. The first two approaches are

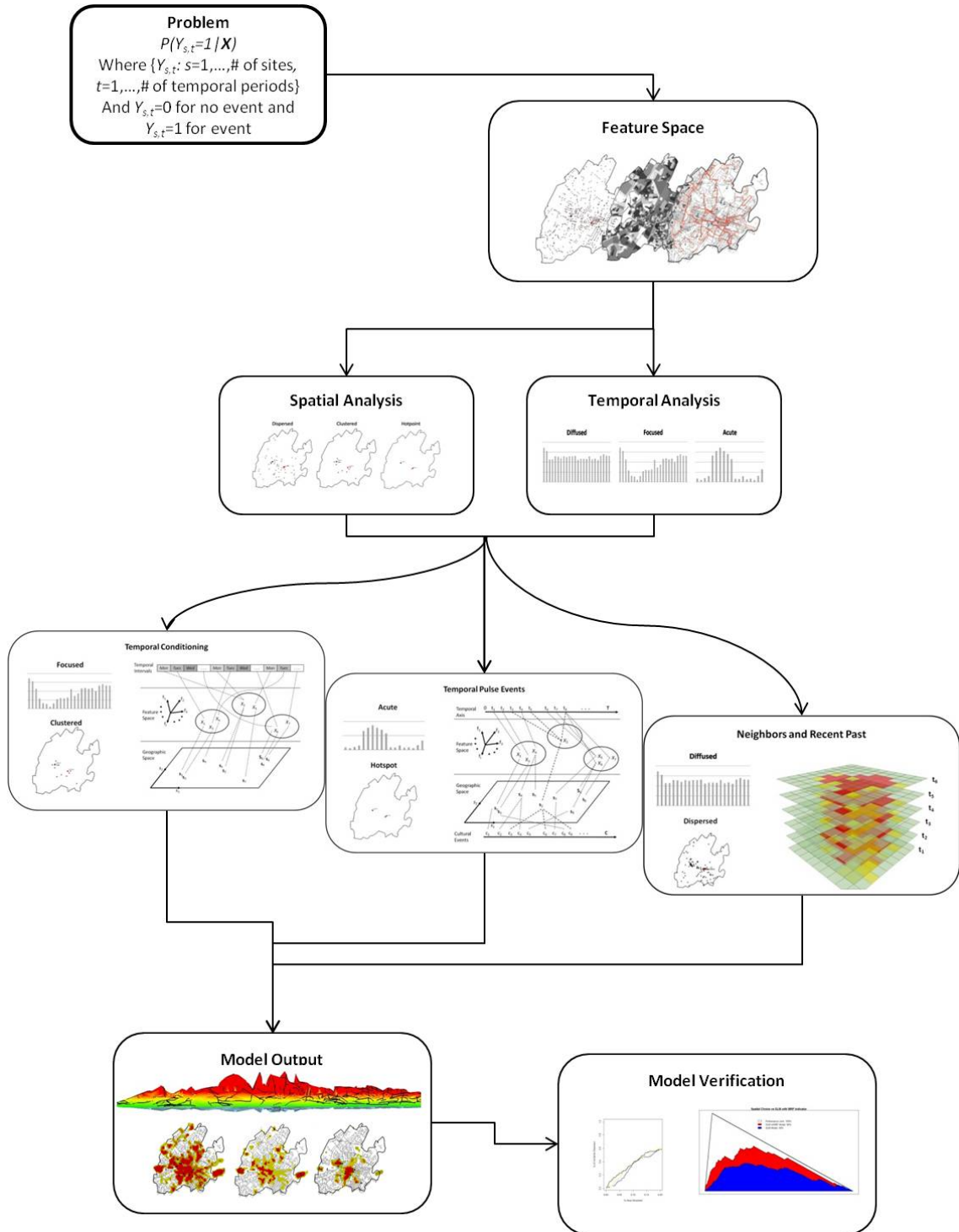
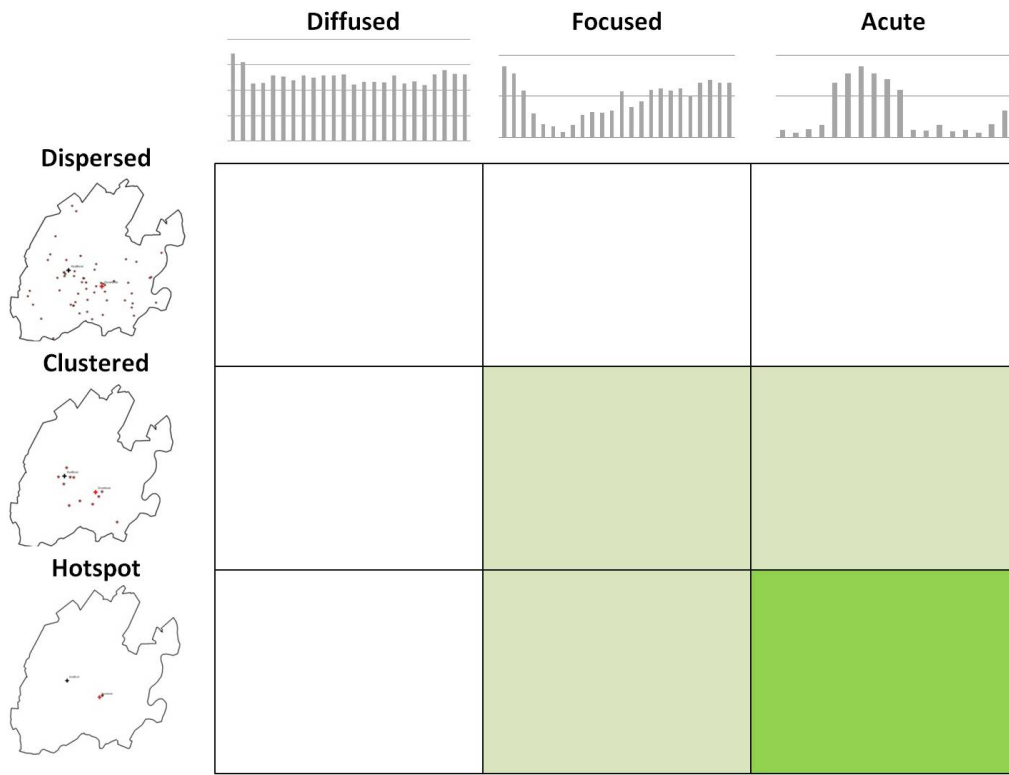


FIG. 3.1: Overview of Dissertation Methodology

conditional in that they condition the spatial analysis for each time interval or condition the temporal analysis for a distinct spatial location. These two conditional approaches serve as the baseline for the first two methodologies of this dissertation. Analyzing the data spatially and temporally allows us to identify both spatial and temporal patterns for further investigation and modeling. As seen in Figure 3.2, the intersection of acute temporal patterns with spatial clustering or hotspots offers an expected area for exploration (Ratcliffe, 2004).



**FIG. 3.2: Appropriateness of Hierarchical Modeling** (Ratcliffe, 2004)

When focusing on predictive models, the expectation is that hierarchical modeling offers the most benefit when both spatial *and* temporal patterns exist. A series of events with a diffused temporal pattern does not offer enough distinction between the temporal interval rates to cause the estimate between intervals to differ from a spatial analysis without temporal information. Similarly, a series of events with dispersed spatial patterns might not offer enough deviation to differ from a temporal

analysis without spatial information.

To identify temporal features that might be significant to an actor's decision process, we first search for temporal patterns in the event data. A visual analysis of the histograms of event occurrences for a specific temporal resolution offers a method for examining the variation between temporal intervals (Chainey and Ratcliffe, 2005). Visual analysis alone is not enough, however, to test the hypothesis of significant temporal patterns within the event data. In Chapter 4, we discuss the autocorrelation and partial autocorrelation functions as statistical methods to identify the appropriate resolution of temporal intervals to examine diffused patterns in the CSS process. In Chapter 5, we use the cross-correlation function to identify potential relationships between special events and windows of acute temporal activity. As mentioned in Chapter 2, several research efforts have focused on the impact of special events to crime rates in cities (Rengert, 1997; Gorr et al., 2003; Cohen et al., 2003). Section 3.5 expands on the study of these special events as temporal pulse events.

After identifying an appropriate temporal resolution for modeling, we then examine these temporal intervals within a spatial context. In Chapter 4, we use both the nearest-neighbor statistic and a comparison of the geographic mean center within the examined temporal intervals to identify potential shifts in the spatial preferences of the criminal site selection process.

### 3.3 Hierarchical Feature-Space Modeling

This dissertation considers that the event data are observed across an irregular lattice at sites  $S_1, \dots, S_n$ . We define a spatial-temporal event as a realization of a process in space and time. Following the work of Liu and Brown (1998) and Chen et al. (2006), we assume that the spatial-temporal event is a realization of

a stochastic spatial-temporal process  $\mathbf{Z}$  that is observed at times  $t_1, \dots, t_m \in T$  at locations  $s_1, \dots, s_n \in D$ . We can further define the random variable  $Y(s, t)$  as the observed realization of our spatial-temporal process  $Z(s, t)$ . As discussed in the previous chapter, the works of Liu and Brown (1998), Xue and Brown (2003), and Brown et al. (2004) provide a method for calculating the likelihood of future criminal events across a spatial region given a dataset of previous criminal events.

To calculate the likelihood and to estimate the criminal's spatial preferences, we mathematically represent the criminal's site-selection process as a binary random variable  $Y_{s,t} \in 0, 1$  where  $\Pr(Y_{s,t} = 1 | \mathbf{X})$  is the probability that a crime occurred at a location  $s$  at a point in time  $t$  given a set of features  $\mathbf{X}$ . We now model the likelihood of a criminal event at a given location and time as a function of the criminal's preferences for each set of features:

$$\begin{aligned} \Pr(Y_{s,t} = 1 | X) &= \text{logit}^{-1}(\beta_0 + \beta_1 X_{s1} + \dots + \beta_k X_{sk}), \\ &\text{for } s = 1, \dots, S \text{ and} \\ &\text{for } t = 1, \dots, T. \end{aligned} \tag{3.1}$$

Equation 3.1 uses a set of features  $\mathbf{X}$  as a vector of length  $k$  for each location  $s$  in order to estimate  $\beta$ , a vector of coefficients that represents the criminal's preference for executing a crime at a location with the selected spatial features. For this article, we assume that the feature vector  $\mathbf{X}$  remains constant over time. Rather than using a latitude and longitude to describe each location in a study region, we use spatial distances to environmental features — such as schools, streets, or stadiums — and spatial representations of social demographics — such as population, percent rental properties, and household income — to examine which locations are preferred by criminals for certain types of crimes (Bannatyne and Edwards, 2003; Liu and Brown, 2003; Huddleston and Brown, 2009). Analyzing previous criminal event data enables



us to model the criminal’s preferences for selecting a certain location in feature-space for executing a crime (Brown et al., 2001). However, just as the preferred location for an individual to execute an event might change from populated to un-populated areas, the CSS process can also change depending on the time of day or the day of week (Rossmo et al., 2005). Although Equation 3.1 assumes that the preferences represented by  $\beta$  remain constant over time, after reviewing temporal patterns in the next section, we will expand Equation 3.1 to account for temporal patterns within the CSS process using hierarchical modeling.

In practice most crime analysis has assumed a constant temporal density during spatial analysis or a constant spatial density for temporal analysis. But many of the questions asked by law enforcement personnel depend on examining criminal events both spatially and temporally:

- Within a specific spatial region, are incident rates higher during specific seasons or months?
- Do daily patterns exist in event likelihood for different spatial areas with similar attributes?
- If temporal patterns exist in the event data, can we assume that changes in temporal patterns match changes in the CSS process?
- Do other event patterns cause shifts to the criminal’s “normal” spatial-temporal decision process?

A normal method for examining changes in spatial event patterns is to build discrete models of events at discrete temporal intervals. Hierarchical modeling provides a method for examining event processes on the basis of varying intercept and varying slope models (Gelman and Hill, 2007). Using hierarchical modeling, we can examine how the impact of seasonality across temporal intervals might affect the

CSS process. Hierarchical modeling also provides a method for examining how the temporal proximity to special events affects the criminal spatial-choice process for short durations.

In order to identify periods of study that might not hold to the constant temporal density assumption, we first analyze the incident set temporally to identify cyclic changes at different temporal resolutions (Chatfield, 1975). After identifying these temporal intervals, we use hierarchical modeling to expand the feature-space model with the inclusion of distinct temporal resolution intervals:

$$\begin{aligned} \Pr(Y_{s,j} = 1|X) &= \text{logit}^{-1}(\alpha_j + \beta_1 X_{s1} + \dots + \beta_K X_{sK}), \\ &\text{for } s = 1, \dots, S \text{ and} \\ \alpha_j &\sim N(\mu_\alpha + \gamma_j, \sigma_{\text{temporal interval}}^2), \\ &\text{for } j = \text{range of temporal intervals} \end{aligned} \tag{3.2}$$

This simple model provides a varying intercept to examine the changing base likelihood of criminal activity during a specific temporal interval. We slightly modify the notation from Equation 3.1 and label the response  $Y_{s,j} = 1$  for instances in which an event occurred at a location  $s$  during temporal interval  $j$ . We have replaced the temporal index  $t = 1 \dots T$  with a range of temporal intervals such as hour of day or day of week. The varying intercept,  $\alpha_j$ , represents the intercept at each temporal unit  $j$  where  $j$  is equal to a specified temporal resolution.  $\mathbf{X}$  remains a vector of predictor variables of length  $k$  as used in the feature space modeling and  $\boldsymbol{\beta}$  is a vector of coefficients that can be used to estimate the criminal's preference for the selected features. We are now modeling how the criminal agent's site selection process changes based on the specific temporal interval during which the crime occurred. This hierarchical feature-space model (HFSM) uses the temporal interval as the hierarchical component. The hierarchical modeling approach allows the analyst

to explore spatial and temporal information in a conditioned approach in order to improve both the understanding of the criminal site selection process and the prediction of crime patterns in time and space (Schabenberger and Gotway, 2005). In Chapter 4, we will expand Equation 3.2 to model the dynamic (changing in time) preferences for certain spatial features on the basis of the temporal interval.

### **3.4 Of Seasonality and Patterns – Temporal Conditioning**

As mentioned in Chapter 2, a significant amount of research focuses on analyzing crime data in space and time in order to better understand criminal patterns and allocate law enforcement resources appropriately. Two primary concerns of resource allocation are daily patrol assignments and seasonal manning. Researchers traditionally apply time series methods to forecast crime rates at either precinct- or patrol-block levels (Gorr et al., 1999; Cohen et al., 2003; Gorr et al., 2003).

Using the autocorrelation and partial autocorrelation functions, we identify temporal intervals with focused patterns of seasonality. Applying the hierarchical framework to the feature-space model gives us a means to understand the criminal’s spatial considerations at a refined spatial resolution while including potential temporal considerations of the CSS process. As seen in Equation 3.2, we are “conditioning” the base feature-space model by incorporating a dynamic intercept that accounts for changes in the CSS process across a range of discrete temporal intervals (Schabenberger and Gotway, 2005). Chapter 4 explores the two-stage analysis as a method for identifying appropriate temporal intervals and applies the hierarchical feature-space model to data simulated with distinct temporal patterns.

## 3.5 Of Indicators and Interactions – Temporal Pulse Events

Given the research that has focused on the impact of special events to crime rates within cities (Rengert, 1997; Gorr et al., 2003; Cohen et al., 2003), how can we model the impact of these increased temporal intervals of opportunity on the CSS? An expansion of the conditioning approach further modifies the feature-space methodology to include temporal features as measures of temporal distances from known special events.

Spatial-temporal pulse events are special events that are correlated with changes in the crime rates within a spatial region for a limited temporal interval. Pulse events may be known one-time events such as elections, school openings, or major concerts. Additionally, pulse events may also include recurring events such as sporting events or weekly prayer meetings. Whether a recurring or one-time event, a pulse event is correlated with a shift in the spatial distribution of crimes across the region for a brief temporal interval. Analyzing the impact of special events on the CSS process proceeds from the work of Enders and Sandler (2006) in defining the impact of intervention to patterns of terrorism.

Since the special events are temporary in duration, we would expect the impact, or pulse, of the event would be temporary as well. However, the impact of the pulse event on the spatial distribution and rates of crime might be seen prior to the occurrence for expected events and after the occurrence of unexpected events. Chapter 5 expands the two-stage analysis to identify periods of acute temporal patterns and applies the hierarchical feature-space model to data simulated with spatial-temporal pulse events that affect the criminal site-selection process.

## 3.6 Of Neighbors and Recent Past

While much theoretical and practical work has been done on the use of Bayesian hierarchical modeling for geostatistics and disease clustering, the application within the construct of the criminal site selection problem has been limited. The last methodology for this dissertation focuses on modeling spatial-temporal events with the inclusion of a neighborhood structure and a temporal component. We compare two approaches for including information about the recent past in the feature-space models for the CSS process. First, we merge the feature-space model of Liu and Brown (1998) with the Markov random field construct of Zhu et al. (2006) to model the criminal's preference for initiating a crime within a specific spatial-temporal zone. By adapting theoretical and computational work from disease mapping and environmental studies, we develop a Bayesian hierarchical feature-space model (BHFSM) for the criminal event prediction problem in order to examine both parameter estimation and predictive inference. The second approach modifies the base feature-space model used for analyzing CSS problems to include temporal and neighborhood indicator functions. The following sections provide a quick review of neighborhood structures, Markov random fields, and generalized linear models. Chapter 6 provides more details on the Bayesian approach used with the hierarchical model and defines the temporal and neighborhood indicator functions.

### 3.6.1 Accounting for Neighbors

Looking at the criminal's decision methodology as a continuous process, however, adds both theoretical and computational challenges. Building on the work of Zhu, Huang, and Wu (2006), we include a neighborhood structure such that

$$N_i \equiv \{j : j \text{ is a neighbor of } i\} \quad (3.3)$$

Chapter 6 expands on the formulation of neighborhood structures used in this dissertation. For now we consider our criminal incident modeled as a random variable  $Y_{i,t}$ , which is a binary indicator on a lattice where  $Y_{i,t} \equiv Y(s_i, t) = 0$  or  $1$ , and  $i$  is an index for sites on the lattice from  $1, 2, \dots, n$  and time is indexed as  $t$  from  $1, 2, \dots, T$ . Zhu, Huang and Wu (2006) provide a full conditional distribution

$$\begin{aligned} & \Pr(Y_{i,t} = 1 | Y_{i,t-1} : (j, t) \in N_{i,t}) \\ &= \text{logit}^{-1} \left( \sum_{k=1}^p \beta_k X_{k,i,t} Y_{i,t} + \sum_{j \sim i} \beta_{p+1} Y_{i,t} (2Y_{j,t} - 1) + \beta_{p+2} Y_{i,t} (2Y_{j,t-1} - 1) \right) \end{aligned} \quad (3.4)$$

where  $\beta_{p+1}$  is the spatial autogression coefficient and  $\beta_{p+2}$  represents the temporal coefficient. The spatial-temporal model shown in Equation 3.4 builds upon the construct of an MRF since the response observed at site  $s_i$  at time  $t$  is conditional on the state of the neighbors at time  $t$  and the state of  $s_i$  at time  $t - 1$ . Chapter 6 provides details on the assumptions of the MRF and an alternative computational construct used to estimate the coefficients and provide insight to the impact of certain variables that follow a temporal pattern as part of the CSS process.

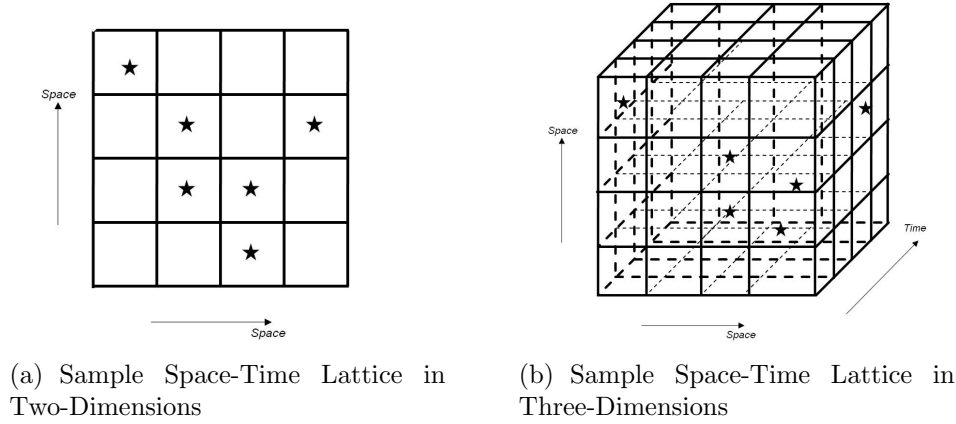
### 3.6.2 Accounting for the Past

Generalized linear models (GLMs) provide an extension to the basic linear model that is the core of regression analysis. Introduced by Nelder and Wedderburn (1972), GLMs provide a means for analyzing binary data. Extensive discussions of GLMs are found in McCulloch and Nelder (1989), Faraway (2006), and Gelman and Hill (2007). The two basic components of GLMs are a response from the exponential family and a link function that provides the relationship between the response and the predictors (Faraway, 2006). For this dissertation, we use binary data, a zero or one, to represent the criminal's decision to initiate or not initiate a criminal event

at a certain time and place. The feature-space model reviewed in Section 2.4.4 is an example of a GLM applied to criminal event data. Chapter 6 provides additional details on the temporal indicator function extensions to the base feature-space GLM.

### 3.7 Significance Testing

How do we measure the performance of the methods reviewed above? We return to our problem formulation and consider that our event data remain a collection of observations process across a regular lattice where  $Y_{i,t} \in \{0, 1\}$  with  $i = 1, \dots, N$  for  $N$  points on the lattice and  $t = 1, \dots, T$  for  $T$  time steps in the study horizon. Figure 3.3(a) depicts a simple  $5 \times 5$  lattice in 2 dimensions.

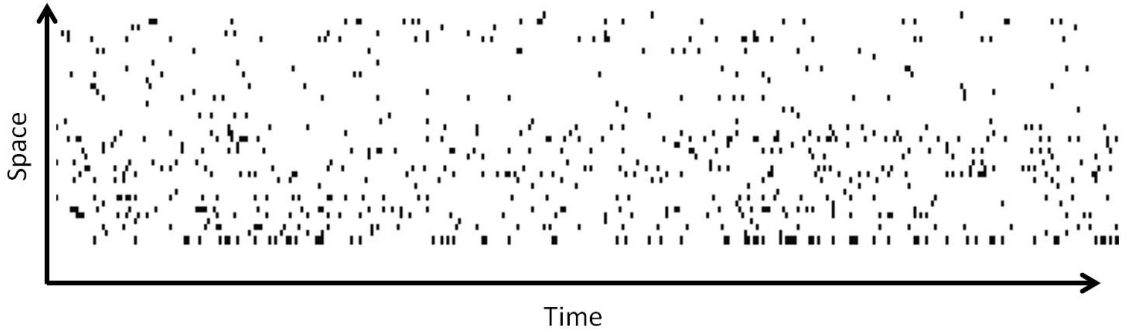


**FIG. 3.3: Sample Space-Time Lattice Structures**

This graphic of a simple lattice, however, fails to account for time. The lattice in Figure 3.3(a) is one realization of the lattice across a continuous time horizon. If we plot event points on a strict two-dimensional lattice, we can not distinguish how including temporal information into a model impacts predictive performance. Figure 3.3(b) provides a slightly better view of our expanded lattice through four

temporal intervals.

If the lattice in Figure 3.3(b) represents a space-time decision domain, a criminal has 64 possible space-time blocks to consider for event initiation. For our dataset in Chapter 7, we consider a space with approximately 2,500 points on the space lattice across 1,000 time points. Figure 3.4 depicts the flattened space-time domain for our application data. The dark points indicate the occurrence of a crime.



**FIG. 3.4: Flattened Space-Time Lattice for Study Domain.** The arrangement of spatial regions along the  $y$  axis might falsely identify spatial clusters, however, the temporal horizon along the  $x$  axis does allow for visual identification of temporal clusters for assaults within the study region. When using the surveillance plot for measuring model performance, we iteratively evaluate all spatial locations within each temporal interval.

### 3.7.1 Receiver Operating Characteristic Curves

One of the challenges for spatial-temporal data is selecting an appropriate statistical measure for examining model performance. Originally used to assess radar performance in World War II, the receiver operating characteristic (ROC) curve is particularly useful for evaluating the ability of a model to predict the occurrence of an event accurately while minimizing the number of false positive predictions (Bradley, 1997; Swets et al., 2000).



**TABLE 3.1: Contingency Table**

	True Condition		
<b>Test Result</b>	Positive	Negative	Measures
Positive	TP	FP	TP + FP
Negative	FN	TN	FN + TN
<b>Measures</b>	TP + FN	FP + TN	

### 3.7.2 Surveillance Plots

Similar to the ROC curve, the surveillance plot provides a method for evaluating model performance in spatial-temporal classification problems. The surveillance plot gives the analyst a method for monitoring the amount of area within the study region that needed to be observed in order to identify the highest percentage of crimes (Huddleston and Brown, 2009; Kewley and Evangelista, 2007). Using a contingency table, or decision matrix, similar to Table 3.1, we record the possible outcomes of prediction estimated with the model being considered against the true conditions observed in the test set. We build the surveillance plot by plotting the rate of accurate crime predictions against the rate of crime incidents predicted when crimes did not occur. The surveillance plot depicts the percentage of area observed compared with the percentage of total incidents observed. In practical terms, the surveillance plot assists law enforcement by showing which predictive model is more accurate over the smallest percentage of the observed area. Using the predictive model with the steepest plot within the first 20 % of area observed allows law enforcement personnel to focus patrols and surveillance resources on the smallest spatial footprint to observe the largest number of potential crimes.

### 3.7.3 Area Under the Curve

Although the surveillance plot provides an appropriate measure for comparing model performance visually, translating the surveillance plot into a numerical measure provides a method for comparing the performance of multiple models against a common test set. Returning to the ROC curve, we can calculate the area under the curve (AUC) to provide a simple measure of each model's accuracy. A model with high accuracy — predicting all the crime locations perfectly — would have a ratio of all true positives versus zero false positives while a model with an equal ratio of true positives and false positives is basically guessing (Bradley, 1997; Swets et al., 2000).

### 3.7.4 Performance Limit Ratio

The performance limit ratio (PLR) measures the model's trade off in accuracy and precision by focusing on the model's better-than-chance ratio (Gorr, 2009b) of correctly predicting crimes out of a test set of size  $n$ . A model that is more accurate in predicting crimes across the space-time surface will have a higher PLR.

$$PLR = \frac{n \times TP}{(TP + FN) \times (TP + FP)} \quad (3.5)$$

Using the formulation suggested by Gorr (2009b), we build Equation 3.5 using the results from our contingency table. Rather than focusing on the entire area under the curve, we reduce the focus to the first 20% of the space-time surface observed while discounting the area under the curve that accounts for random guessing.

# CHAPTER 4

## Method 1 – Of Seasonality and Patterns

*Sometimes the best map will not guide you*

Jimmy Buffet, “Pacing the Cage”

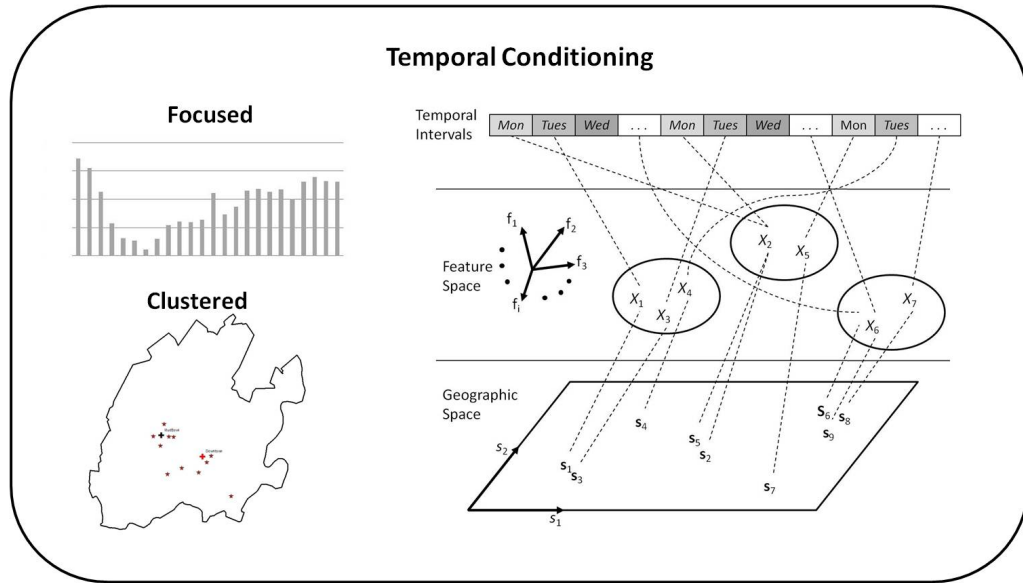
*...but I don't care. It's five o'clock somewhere.*

Jimmy Buffet and Alan Jackson, “It's Five O'Clock Somewhere”

The initial query of this dissertation is how to identify temporal features that might be significant to a criminal’s decision process. And once a method is identified, how can we combine the temporal features with spatial information to enhance an analyst’s understanding of event patterns and improve the predictive performance of current criminal site-selection (CSS) event models? The following sections review the base model, expand on the methodology for identifying relevant temporal patterns, and develop the hierarchical structure used for incorporating temporal intervals in the feature-space model. The discussion in this chapter expands on the two-stage analysis for spatial-temporal event prediction discussed in Schabenberger and Gotway (2005). We use the two-stage analysis to identify the appropriate temporal resolution for inclusion into the hierarchical feature-space model (HFSM) and to identify spatial shifts in the distribution of crimes across discretely defined temporal intervals. The HFSM provides a method to account for seasonality in the criminal actor’s site selection process. After introducing an agent-based crime simulation, we conclude the chapter with an examination of the methodology against a simulated data set.

## 4.1 Formulation

Hierarchical modeling offers a method for incorporating the temporal information into the feature-space model by accounting for temporal shifts in the CSS process at a specific temporal resolution. Hierarchical modeling is a regression technique that enables modeling across groups or hierarchies. Traditional methods for dealing with hierarchical data include building separate models for each distinct group—and risk missing common factors among groups—or build one model for all groups—and risk missing any variations among the groups. Hierarchical modeling bridges these two techniques by pooling the factors and variations from all groups (Gelman and



**FIG. 4.1: Method 1 – Temporal Conditioning.** Beginning with a series of observed crimes, we analyze the data to identify spatial clustering during temporal intervals with focused rates of criminal activity. Using both the base feature-space model and the hierarchical framework, we model historical data and develop predictive threat surfaces for future events. Finally, we compare the models using the surveillance plot and the AUC.

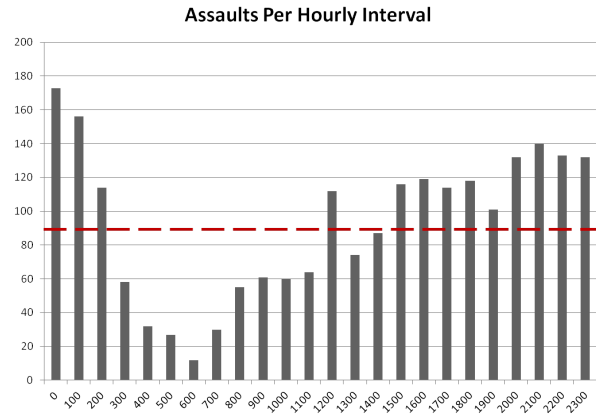
Hill, 2007; Huddleston and Brown, 2009). For this dissertation, we use temporal intervals as the hierarchical component. Adapting the feature-space model to a hierarchical framework provides a method to account for temporal variations in the CSS process. Previous research using hierarchical modeling includes health sciences (Banerjee et al., 2004), disease mapping (Lawson, 2009), voting analysis (Gelman and Hill, 2007), and crime studies involving gang activity (Huddleston and Brown, 2009). However, just as the preferred location for an individual to execute an event might change from populated to un-populated areas, the CSS process can also change depending on the time of day or the day of week (Rossmo et al., 2005). After reviewing temporal patterns in the next section, we expand Equation 3.1 to account for temporal patterns within the CSS process using hierarchical modeling.

## 4.2 Methodology

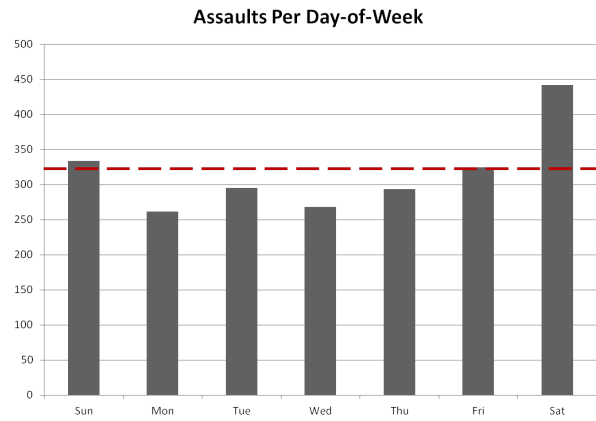
### 4.2.1 Two-stage Analysis

Section 3.4 introduced the formulation for conditioning the feature-space model on the basis of discrete temporal intervals. Prior to applying the hierarchical framework to a feature-space model, we conduct a preliminary analysis of the event data to identify temporal intervals with shifts in the CSS process. The preliminary analysis gives us an understanding of the spatial and temporal patterns that might exist in the data. Recall from Chapter 2, the underlying assumption of feature-space modeling is that there are patterns to the historical criminal activity data within a region. If criminal activity is truly random, predictive modeling is impossible (Kerchner, 2000). Sequentially analyzing the data across time and space allows us to identify both spatial *and* temporal patterns in the CSS process that we can account for using a hierarchical feature-space model.

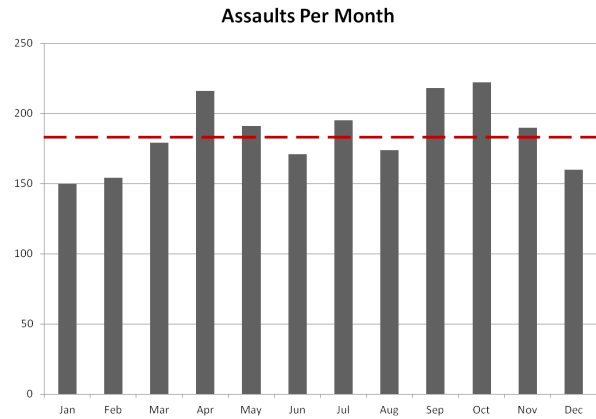
Figure 3.2 provides our introductory look at the intersection of spatial and temporal categories. For this dissertation, we begin with a visual analysis of the histograms of event occurrences at distinct temporal resolutions. The visual analysis highlights variations between temporal intervals at a given resolution (Chainey and Ratcliffe, 2005). As introduced by Ratcliffe (2004), crime series can have several types of temporal patterns. A *diffused* temporal pattern has no discernable pattern of activity at the examined temporal resolution. This does not mean that there are not differences in the examined temporal intervals. Instead the differences are simply not significant enough for forecasting and resource allocation. A *focused* temporal pattern presents an interval of time that is more or less active than the other time intervals for event activity. An *acute* temporal pattern is an increase of event activity within a smaller interval of the examined temporal resolution. For example, the number of alcohol related crimes might peak around sporting events or closing times of nightclubs. But an acute temporal pattern alone is not enough for determining the application of hierarchical modeling. The intersection of focused temporal patterns with spatial clustering offers our first expected area for exploration (Ratcliffe, 2004). In order to identify temporal patterns within a spatial region, we examine a series of events across several temporal resolutions to identify an appropriate temporal interval to use in constructing a hierarchical model. As an example of this approach, Figure 4.2 presents histograms from the set of crimes analyzed in our sample data study. We examine the crime counts across three separate temporal resolutions; hourly, daily, and monthly. The horizontal line in each plot represents the mean value for number of crimes per temporal interval across the observed study period. Given the construct of a hierarchical model, we are looking to identify parameter estimates that are unique to specified temporal intervals rather than using the same parameter estimates across the entire temporal range. Selecting temporal intervals with greater variation, either acute



(a) acute - focused



(b) focused - diffused

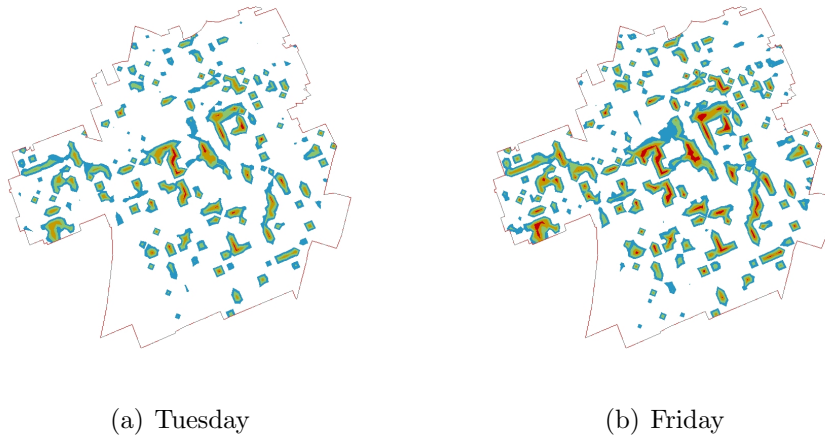


(c) diffused

**FIG. 4.2: Graphical Depiction of Events at Different Temporal Resolutions.** A visual analysis of the histograms across several temporal resolutions aids in identifying potential temporal intervals to use in constructing the hierarchical model (Chainey and Ratcliffe, 2005).



temporal patterns or focused temporal patterns, offers the best potential for model development. In Figure 4.2(a), the hourly incident plots offer an acute pattern within the late evening hours. The day-of-week crime plot in Figure 4.2(b) depicts a focused temporal pattern with clear distinction during the weekend interval. In Figure 4.2, the hourly incident plots offer an almost acute pattern in the late evening hours. For computational simplicity, however, we can build the basic hierarchical model using the daily variation between temporal intervals to capture the varying intercept. Since the most significant difference in daily incident rates is evident between Tuesday and Friday, we hypothesize that the predictive surface between those two time periods would offer the most distinction. Figure 4.3 highlights how the likelihood surface can change in specific areas by conditioning on the temporal intervals.



**FIG. 4.3: Sample Daily Prediction Surfaces**

As introduced in Chapter 3, visual analysis alone does not provide the analyst an understanding of spatial distributions within specified temporal intervals. Using the notation provided by Schabenberger and Gotway (2005), we utilize a nearest-neighbor test statistic such that given  $h_i$  as a distance from an event at one point

to the nearest other event and an indicator function  $I(h_i \leq h)$  that returns 1 when evaluated as true, the function

$$\hat{G}(h) = \frac{1}{n} \sum_{i=1}^n I(h_i \leq h) \quad (4.1)$$

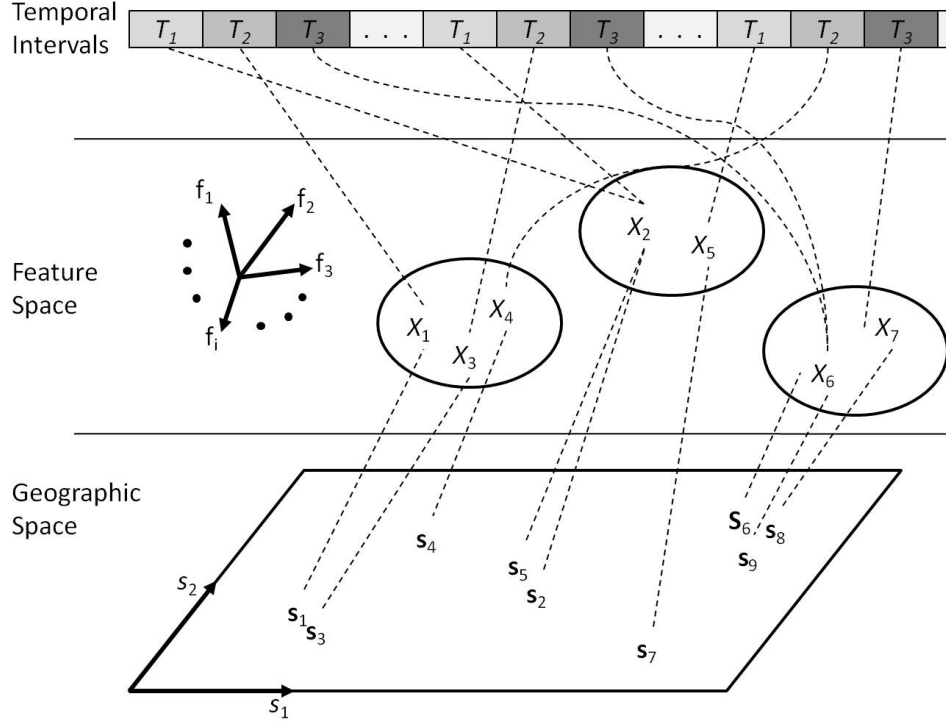
is now an estimate of the nearest-neighbor distribution function. Using the Average Nearest Neighbor index, we calculate the ratio of the observed mean distance between crime events and an expected mean distance from a pattern of events that are completely spatially random. If the ratio is close to one, then we expect clustering of events to be unlikely. If the ratio is significantly less than one, we have strong support of clustering within the examined temporal interval (Gorr and Kurland, 2011).

Although identification of clustering is important for more traditional crime analysis, we are not interested in the clusters themselves but rather in examining if the temporal shifts in clusters are a result of temporal shifts in the CSS process. We use the geographic mean center to compare the distribution of the crime clusters across the identified temporal intervals. Temporally aligned shifts in the geographic mean center further supports the hypothesis that temporal patterns exist in the CSS process. We estimate the mean center of criminal events by calculating the mean  $x$  and  $y$  using the spatial coordinates of the event data (Unwin, 1981). If the spatial distribution of criminal activity shifts across temporal intervals in relation to the overall geographic mean center, we hypothesize that a temporal shift is present in the CSS process.

### 4.2.2 Hierarchical Modeling

Graphically, we adjust the original Liu and Brown feature-space model to reflect an environment in which a criminal is more likely to select a spatial location with a

specific set of features within certain temporal intervals (Liu and Brown, 1998):



**FIG. 4.4: Feature-Space Model of Temporal Intervals.** Based on original graphic developed in Liu and Brown (1998). We replace the continuous time horizon with a set of discrete temporal intervals. For each temporal interval depicted, the preferences in the CSS process shift from features  $X_1, X_3$ , and  $X_4$  on Monday to features  $X_2$  and  $X_5$  on Tuesday. During all other temporal intervals, the criminal prefers locations with features  $X_6$  and  $X_7$ .

As seen in Figure 4.4, each temporal interval has a series of events that occurs in distinct feature-space clusters. In Equation 3.2, we wrap the straight logistic regression of Equation 3.1 inside a hierarchical framework to model the variations in the base line CSS process across temporal intervals at a specific temporal resolution. Equation 3.2 does not account for dynamic (changing in time) preferences for certain spatial features based on the temporal interval  $j$ . Following the work of Gelman and Hill (2007), we further expand the base feature-space model to include both an intercept and a series of coefficients that varies according to a specific temporal interval. Assuming that the regression parameters have independent Gaussian

priors, we can specify the hierarchical structure as follows:

$$\begin{aligned} \Pr(Y_{s,j} = 1|X) &= \text{logit}^{-1}(\alpha_j + \beta_{j[k]} \mathbf{X}_s), \text{ for } s = 1, \dots, S \\ \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} &\sim N \left( \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix} \right), \text{ for } j = 1, \dots, \text{ temporal interval} \end{aligned} \quad (4.2)$$

We label the response  $Y_{s,j} = 1$  for instances in which an event occurred at a location  $s$  and during time interval  $j$  and  $Y_{s,j} = 0$  if no event occurred.  $\mathbf{X}$  is a vector of length  $k$  of predictor features such as the distance to entertainment districts or the distance to police stations from the location  $s$ .  $\beta$  is a vector of length  $K$  that models the dynamic (changing in time) preferences for certain spatial features on the basis of the temporal interval  $j$ . We are now modeling how the criminal agent's site selection process changes during discrete temporal intervals. By examining the resulting coefficients, we identify which features might attract or repel crime during certain temporal intervals. Features with negative-value coefficients can be seen as criminal attractors (areas that make crime more attractive for criminals). Features with positive values are areas less desirable to a criminal. Changes in a coefficient value across temporal intervals might indicate that certain areas become more or less desirable to a criminal at certain times. The hierarchical modeling approach allows the analyst to explore spatial and temporal information in a "conditioned" approach in order to improve both the understanding of the criminal site selection process and the prediction of crime patterns in space and time (Schabenberger and Gotway, 2005).

## 4.3 Simulated Data

Previous work has shown the applicability of using simulations to model criminal behavior (Gunderson and Brown, 2000; Liang, 2001; Liu et al., 2005). For this dissertation, we use a simulation to generate data to test both the two-stage analysis methodology and the applicability of hierarchical feature-space modeling for analyzing crime data where definitive spatial-temporal patterns exist in the CSS process.

### 4.3.1 Introduction of Crime Simulation

Beginning with an agent-based model used to examine relationships between citizens, violent actors, and protectors (Huddleston et al., 2008), we create a small spatial region using the simulation software NetLogo to simulate temporal patterns within a criminal’s CSS process. Within the spatial region, we identify six variables  $x_1, \dots, x_6$  as distances to spatial features of interest. We populate the spatial region with only two breeds of agents: criminals and citizens. We base the ratio of criminals to citizens on empirical results from previous criminal studies with an adjustment in scale to match the limited spatial region of the simulation (Meeker et al., 2002).

#### Agents

We begin each simulation run with the citizen agents randomly distributed across the region with assigned home, work, and entertainment locations. Throughout the simulation, each citizen agent moves from their home location to their work, or entertainment, location with varied temporal patterns. The criminal agents begin each simulation in four distinct regions. Each criminal agent has a specified length of time for moving amongst the population. Within each daily temporal interval, the interaction between a criminal and citizen agent triggers the criminal agent’s

site selection process. If the criminal agent is sharing a location with more than one citizen, no crime will occur. Multiple citizen agents collocated at one location serve as defacto guardians. However, if a criminal agent shares a location with only one citizen, the criminal agent is more likely to initiate a crime if the location matches his preferences.

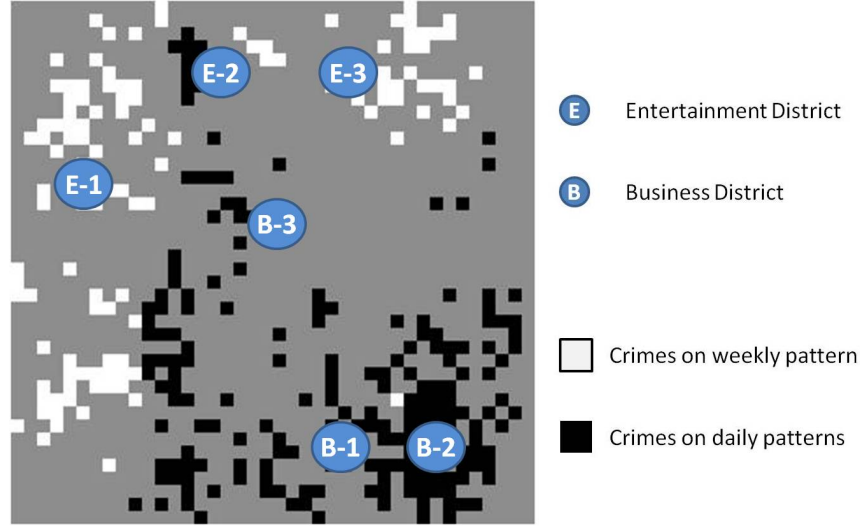
### **Simulation Scenarios**

Although the criminals have separate spatial regions, they have a similar site selection process related to event initiation. During the week, the business districts serve as attractors to criminal activity. Given a convergence event between a criminal agent and a solo citizen, the criminal agents are more likely to initiate a crime if they are closer to the business districts. During the weekend interval, the business districts serve as repellers to criminal activity and the entertainment districts serve as attractors. In practical sense, if criminal agents encounter solo citizens on a Saturday in the simulation, they are more likely to commit a crime if they are closer to the entertainment districts.

Figure 4.5 offers a visual depiction of one simulation run including the preference for a spatial-temporal shift to the entertainment districts during the weekend intervals of the simulation. The white squares indicate crime events occurring on the seventh day of the weekly pattern. The black squares indicate criminal events that occur during the remainder of the weekly interval.

### **4.3.2 Two-Stage Analysis**

We aggregate the simulated crime event data and plot conditioned upon the daily temporal intervals. Although Figure 4.6 provides a clear depiction of the acute



**FIG. 4.5: Sample Simulation Run with Weekly Crime Patterns**

temporal pattern within the simulated crime dataset, we know that visual analysis alone does not always provide an accurate identification of the temporal patterns. Using the autocorrelation (ACF) and partial-autocorrelation function (PACF), we examine the simulated crime data for correlation structure (Crawley, 2007). Since we ran the simulation with a weekly temporal pattern for the criminal's CSS process, we would expect to see regular cycles at the seven day intervals on the ACF. Figure 4.7(a) provides clear evidence of a seven day cycle within the simulated crime data. The partial autocorrelation plot in Figure 4.7(b) also shows distinct lags at the seventh and fourteenth day, further confirming the existence of a weekly pattern within the simulated event data.

Using the Average Nearest Neighbor index, we calculate the ratio of the observed mean distance between crime events and an expected mean distance from a pattern of events that are completely spatially random. If the ratio is close to one, then we expect clustering of events to be unlikely. If the ratio is significantly less than one, we have strong support of clustering (Gorr and Kurland, 2011). For the simulated crime data, the nearest neighbor distance for daily crimes are significantly smaller than

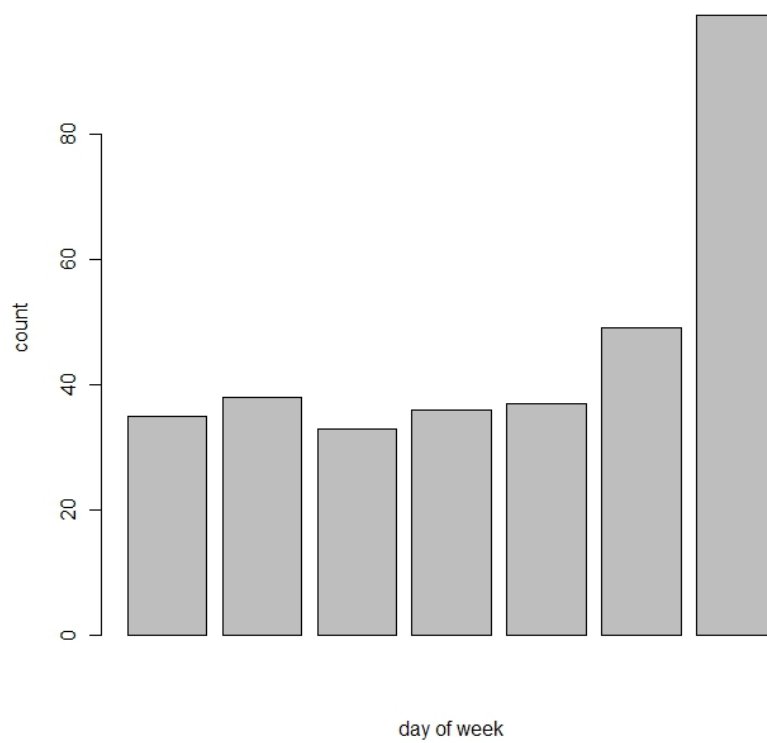
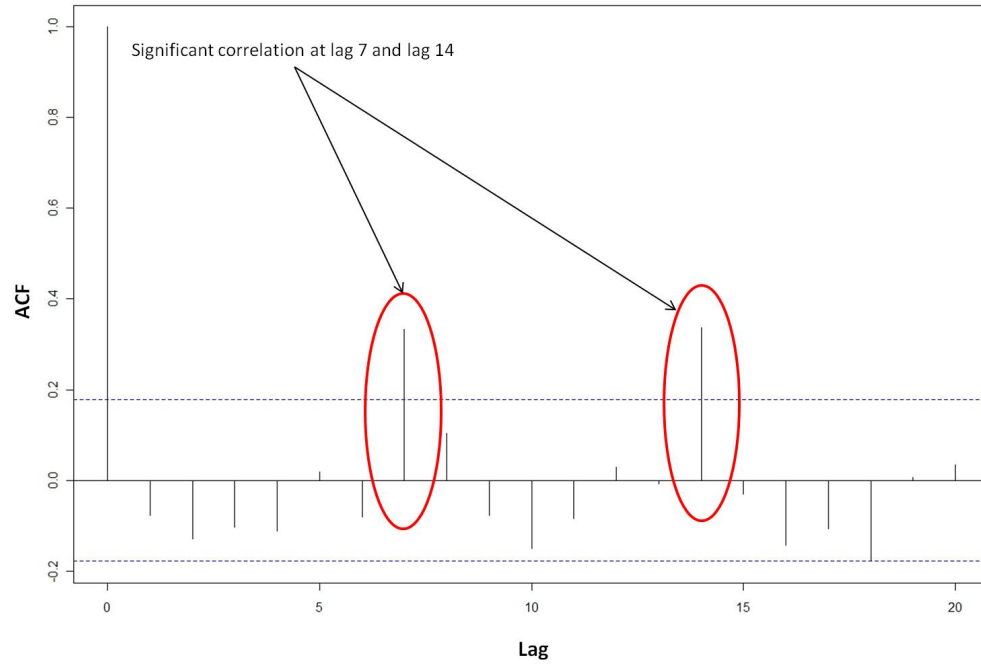
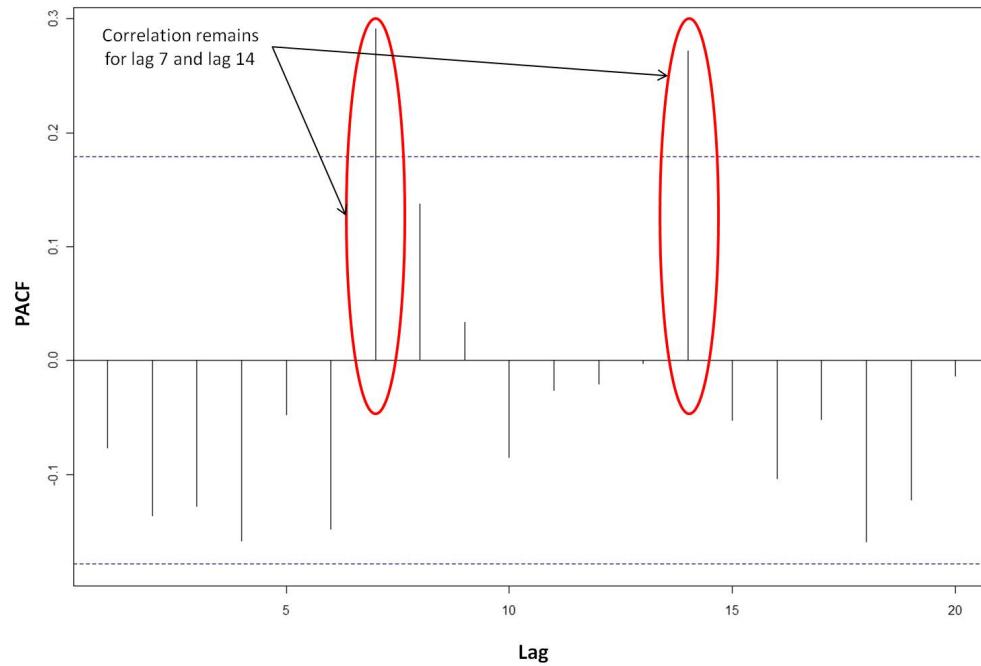


FIG. 4.6: Count of Crimes Per Day of Week in Simulation





(a) ACF



(b) Partial ACF

**FIG. 4.7:** Autocorrelation Plots – Simulated Crimes

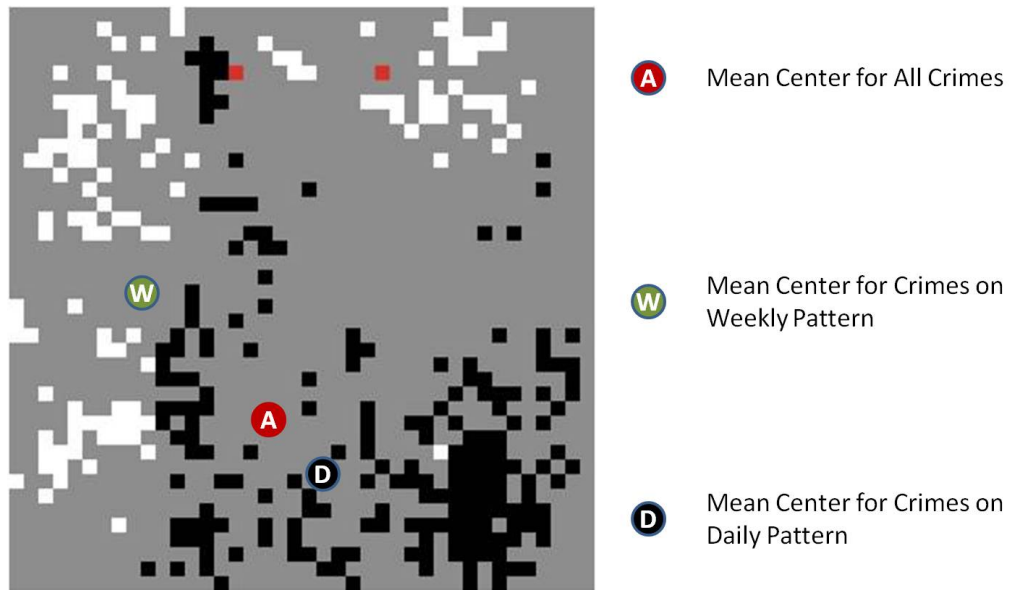
Autocorrelation plots of crime counts from simulation. The ACF and PACF plots of crime data aggregated at a daily temporal resolution identify strong correlation at a seven day interval. Since we ran the simulation with a weekly temporal pattern for the criminal's CSS process, we would expect to see regular cycles at the seven day intervals on the ACF.

the expected mean distance indicating more clustering. Of course, this increased clustering is expected since the business districts serve as attractors for criminal agents and business district two and three are themselves clustered together.

Although Figure 4.5 provides visual evidence of spatial-temporal patterns, an additional test using a daily geographic mean center calculation further supports the hypothesis that spatial-temporal clustering is present in the simulated event data. We use the geographic mean center to compare the distribution of the crime clusters across the identified temporal intervals (Unwin, 1981). We estimate the mean center of criminal events by calculating the mean  $x$  and  $y$  using the spatial coordinates of the event data. If the spatial distribution of criminal activity shifts across temporal intervals in relation to the overall geographic mean center, we hypothesize that a temporal shift is present in the CSS process. Figure 4.8 depicts the three distinct mean centers for crime clusters from the simulation based on the temporal patterns of weekly periodicity, daily activity, and the overall mean center. The initial analysis identifies potential space and time patterns in the criminal agents' site selection process. Taking the temporal patterns found with the correlation functions and the spatial shifts identified with the nearest neighbor analysis and mean center calculations, we adjust the hierarchical framework around the feature-space model to account for the impact of daily temporal patterns in the CSS process.

### 4.3.3 Hierarchical Modeling

The two-stage analysis assists us in identifying temporal intervals with clustered spatial patterns. Since the autocorrelation and partial autocorrelation plots depicted such a strong weekly seasonality and the geographic mean center comparison showed clear shifts in the spatial distribution of crime across temporal intervals, we can now



**FIG. 4.8: Geographic Mean Centers of Temporal Intervals from Crime Simulation.** The shift in the geographic mean center for the two distinct temporal intervals indicates a shift in the criminal agent's site-selection process on the basis of a temporal pattern. Since the criminal agent rule set contains a shift in preference every seven days in the simulation, the shift in geographic mean center is expected.

refine Equation 4.2 to model the CSS process at a daily temporal interval level.

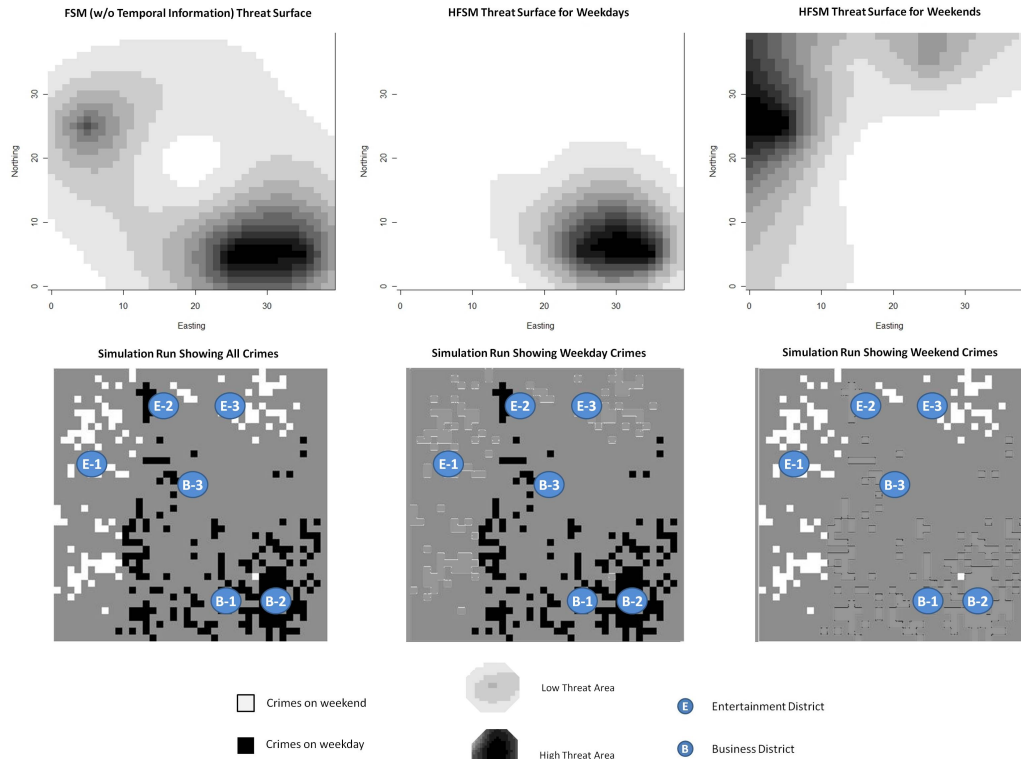
$$\begin{aligned} \Pr(Y_{s,j} = 1|\mathbf{X}) &= \text{logit}^{-1}(\alpha_{[j]} + \boldsymbol{\beta}_{k[j]}\mathbf{X}_s), \text{ for } s = 1, \dots, n \\ \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} &\sim N\left(\begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix}\right), \text{ for } j = 1, \dots, 7 \end{aligned} \quad (4.3)$$

The varying-slope and varying-intercept model shown in Equation 4.3 continues to include the spatial feature-space components of Liu and Brown (2003) and uses the hierarchical structure from Gelman and Hill (2007) to model the impact of temporal patterns in the CSS process. Within the spatial region, we identify six variables  $x_1, \dots, x_6$  as distances to spatial features of interest. Table 4.1 provides an overview of the variables contained within the simulated dataset.

**TABLE 4.1: Predictor Variables for Simulation Dataset**

Variable Name	Description
<i>DisEnt1</i>	Distance from point to entertainment district one
<i>DisEnt2</i>	Distance from point to entertainment district two
<i>DisEnt3</i>	Distance from point to entertainment district three
<i>DisBus1</i>	Distance from point to business district one
<i>DisBus2</i>	Distance from point to business district two
<i>DisTownCenter</i>	Distance from point to center of spatial region
<i>Response</i>	Binary – did a crime occur? 0 – no crime 1 – crime

The predictive surfaces in Figure 4.9 represent the two distinct temporal patterns in the CSS process. Although the base feature-space model (FSM) accurately depicts the two distinct areas of crime hotspots, the HFSM, which is conditioned on the seven-day temporal pattern, better accounts for the temporal shift in the criminal agent’s spatial preferences.



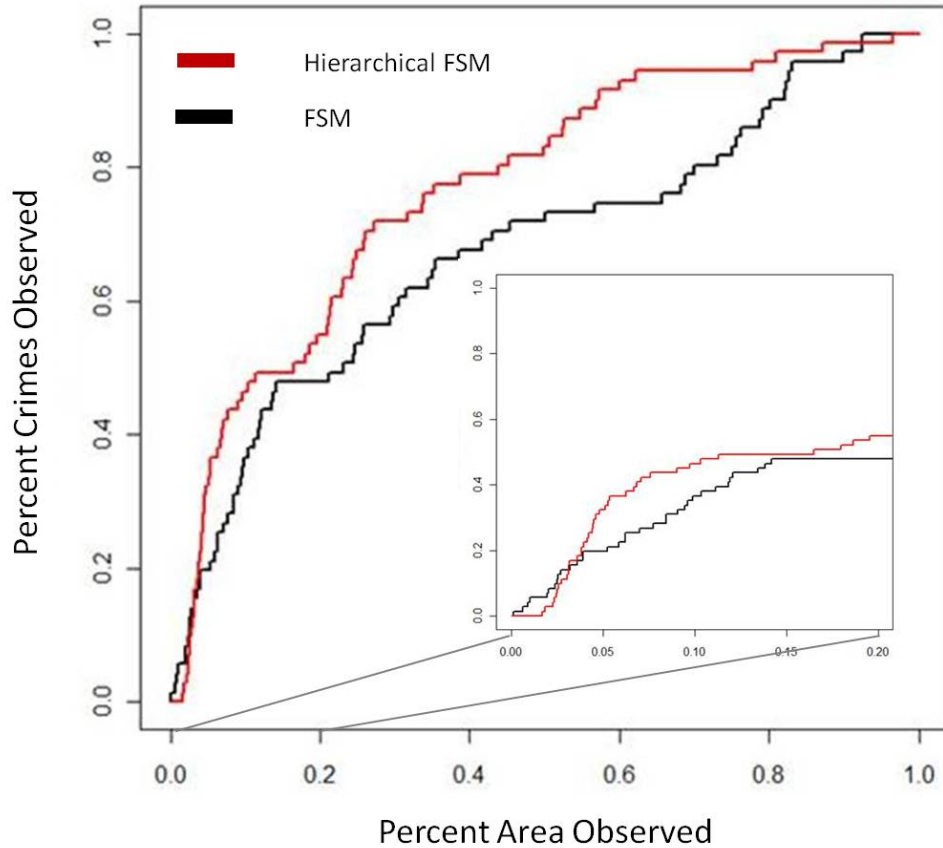
**FIG. 4.9: Predictive Surfaces, aka Heatmaps, for Simulation Using Base and Hierarchical Models.** Areas with higher probability of crimes are shown with darker colors. The HFSM accounts for temporal patterns in the criminal site selection process and can be used to develop threat surfaces for specific temporal intervals. A more focused threat surface assists law enforcement personnel with patrol planning and resource allocation.

### 4.3.4 Significance Testing

The surveillance plot serves as a measure of the model’s performance. Figure 4.10 depicts a space-time surveillance plot measuring the percentage of the possible space-time blocks covered within the study along the X-axis and recording the percentage of incidents observed along the Y-axis. After using the first ninety days of simulated data to train the model, we evaluate against a test set containing the final thirty days of the simulated dataset. The test set represents a possible 48,000 possible space-time blocks for a criminal agent to consider as part of their CSS process. The HFSM is in red and clearly dominates the base FSM in black, especially within the first 20% of the space-time area. We quantify the difference in predictive performance by calculating the area under the curve for the two surveillance plots (Sing et al., 2005). Overall, the HFSM offers a 13% improvement over the base FSM without temporal information.

## 4.4 Discussion

This chapter provides the first methodology for combining temporal features with spatial information to enhance an analyst’s understanding of event patterns and improve the predictive performance of feature-space models. Using the two-stage analysis process proposed by (Schabenberger and Gotway, 2005), we examine the event data for spatial areas of clustering aligned with focused temporal patterns of activity. The ACF and PACF assist in identifying potential seasonality within the crime event data and determining the right temporal resolution for examining shifts in the spatial distribution of crimes. The geographic mean center conditioned on the identified temporal intervals aids us in confirming the appropriate temporal resolution for further investigation and inclusion in the hierarchial feature-space



**FIG. 4.10: Surveillance Plot for Simulated Crime Data – HFSM vs FSM** The surveillance plot depicts the percentage of area observed compared with the percentage of total incidents observed. In practical terms, the surveillance plot assists law enforcement by showing which predictive model is more accurate over the smallest percentage of the observed area. Using the predictive model with the steepest plot within the first 20 % of area observed allows law enforcement personnel to focus patrols and surveillance resources on the smallest spatial footprint to observe the largest number of potential crimes.

model.

We evaluate the methodology against a simulated data set. The agent-based crime simulation replicates a small spatial region with criminal agents who have a propensity to include a weekly temporal pattern in their intelligent site selection process. The simulation closely approximates patterns found in other routine activity studies (Groff, 2007). More importantly, the simulation provides a method for improving our understanding of how routine activities impact the population dynamics and affect the complex interactions that make up the site-selection process. We compare the performance of the basic feature-space model against the HFSM. Initial results indicate that the HFSM provides the best predictive performance. Using the simulated data set, the HFSM provides a 13% improvement in predictive performance over the base feature-space model without temporal information.



# CHAPTER 5

## Method 2 – Of Indicators and Interactions

*“Is there any other point to which you would wish to draw my attention?”*

*“To the curious incident of the dog in the night-time.”*

*“The dog did nothing in the night-time.”*

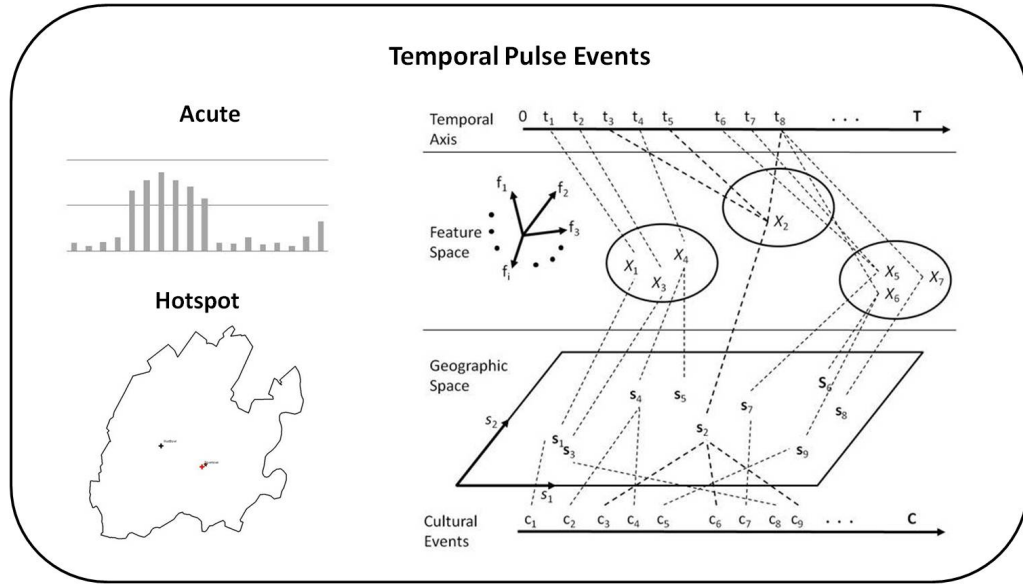
*“That was the curious incident.”*

– Colonel Ross and Sherlock Holmes in *The Adventure of Silver Blaze*

In addition to temporal intervals, what other temporal features can we identify that might be significant to an actor's decision process? And once identified, can we combine the temporal features with spatial information to further enhance an analyst's understanding of event patterns and improve the predictive performance of current event models? Figure 5.1 provides the pictorial map for this chapter. The initial temporal and spatial analysis provides the basis for identifying pulse events within the study horizon. Since pulse events are normally seen in periods of acute temporal patterns, we use the cross-correlation function to compare time series data of crime rates with a time series of potential special events. In order to identify the potential spatial impact of these pulse events, we calculate the geographic mean centers of events during temporal windows preceding, during, and after the proposed temporal pulse events. After identifying potential temporal pulse events, we modify the hierarchical structure of Equation 4.2 and replace the temporal interval with a temporal distance to the identified pulse event. After modifying the rule set for the agent-based crime simulation introduced in Chapter 4, we conclude with an examination of the methodology against a simulated data set.

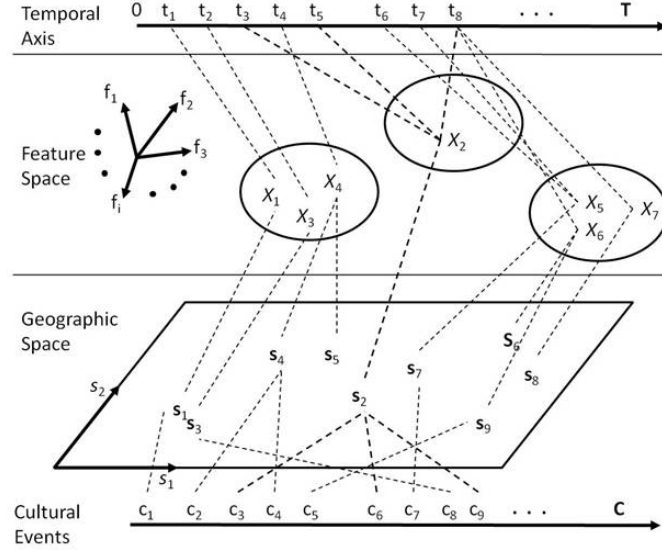
## 5.1 Formulation

In Chapter 4, we used the hierarchical framework with the base feature-space model to include temporal intervals at specific temporal resolutions. Although temporal intervals normally match the seasonal patterns of the criminal site-selection process, how can we account for a criminal's selection of a site during temporal intervals that include specific special events associated with changes in a region's normal activity patterns? We again modify the original Liu and Brown feature-space graphic (seen in Figure 2.11) to include a new input of special events that happen in distinct spatial locations at distinct times.



**FIG. 5.1: Method 2 – Temporal Pulse Events.** Beginning with a series of observed crimes, we analyze the data to identify spatial hotspots during temporal intervals with acute rates of criminal activity. We compare time series of special events with temporal windows of acute rates of criminal activity to identify “pulse events.” We adjust the HFSM to include a temporal distance to the pulse events as the hierarchical component. Finally, we compare the models using the surveillance plot and the AUC.

As mentioned in Chapter 3, the special events could include baseball games, political rallies, or large-scale events such as the Cherry Blossom Festival in Washington, DC. Events of these types have both a temporal and spatial component. Some large-scale events occur annually, while others, such as sporting events, are held with some degree of temporal regularity. Regardless of their temporal occurrence patterns, special events also have a spatial component that contributes to changes in a region's normal activity patterns. Special events might be centered on downtown locations, such as stadiums or arenas, or they might be located outside of city limits, such as a fairground or raceway. A special event becomes a temporal pulse event when the normal patterns of the criminal site-selection preferences are shifted for the duration of the event. Figure 5.2 illustrates one such event. Along the special event timeline, three distinct special events ( $c_3$ ,  $c_6$ , and  $c_9$ ) all occur at one spatial location ( $s_2$ ). Although the three special events do not occur with any degree of seasonality, the criminal's feature-space preference shifts to  $X_2$  for those select temporal intervals. Identifying changes in criminal feature-space preferences assists law enforcement personnel with developing resource allocation during special events. In order to identify pulse events, we select several *posterior* and *a priori* special events during the study horizon. These events occur within the local region of the study area but perhaps not within the specific boundaries of the study area. The posterior events are events that occurred without prior knowledge but that might be assumed to have had an effect on criminal activity. An example might be storms that impact electrical power availability or social-media-inspired gatherings (Cohen et al., 2003). When considering the impact of environmental conditions, we would measure the number of days from a posterior event, such as a storm that caused power outages. The *a priori* events are events that had a scheduled occurrence and that authorities believed might cause an increase in criminal activity. Examples include football games or other large-scale special events, such as the aforementioned



**FIG. 5.2: Feature-Space Model with Cultural Events**

Cherry Blossom Festival. Each of these events would also have a spatial component that could be included within the model. By including potential pulse events from the periphery of the study area, we extend the hypothesis such that pulse events in the greater metropolitan areas might impact the criminal activity in the smaller cities.

## 5.2 Methodology

### 5.2.1 Two-stage Analysis

As in Chapter 4, we use Schabenberger and Gotway's (2005) conditional approaches for the first steps of spatial-temporal analysis. We are still looking for temporal patterns, however, we want to see whether any acute temporal patterns are correlated with temporal patterns of cultural events and whether these acute patterns deviate from the regularly observed spatial patterns of the CSS process.

For temporal intervals, we used the autocorrelation functions calculated at

different temporal resolutions to identify potential temporal patterns in the CSS process. To identify potential temporal pulse events, we use the cross correlation function (CCF) seen in Equation 5.1. The CCF quantifies the strength of the relationship these special events might have with the CSS process. By comparing the criminal-event time series with a time series built upon knowledge of special events, we identify potential spatial-temporal shifts in the site selection process that reflect increased opportunities for the criminal. The CCF is the statistical measure of the relationship between these two time series. We mathematically define the cross-correlation using the notation provided by Chatfield (1975):

$$\rho_{wy}(\tau) = \frac{\gamma_{wy}(\tau)}{\sqrt{\gamma_{ww}(0)\gamma_{yy}(0)}} \quad (5.1)$$

This function measures the correlation between our incident time series  $Y_t$  and our potential pulse event time series  $W(t + \tau)$ .

But the CCF does not identify the spatial impact of a pulse event on the CSS process. The second step in our methodology uses the geographic mean center to compare the distribution of the crime clusters within the pulse event temporal interval to previous temporal intervals (Unwin, 1981). We estimate the mean center of criminal events by calculating the mean  $x$  and  $y$  using the spatial coordinates of the event data. If during the temporal interval of the special event, the spatial distribution of criminal activity shifts significantly from the normal geographic mean center, we hypothesize that the presence of the special event causes a shift in the CSS process.

### 5.2.2 Hierarchical Modeling

Given a positive (or negative) correlation between a criminal event time series and a potential pulse event time series, we again use the hierarchical framework around

the base feature-space model to account for temporal shifts in the CSS process. In this case, we set the hierarchical component to be the distance to, or from, an event at the resolution of the potential set. For instance, when considering sporting events as a set of potential pulse events, we measure the number of days to an a priori event such as a football game or racing event. After identifying these potential pulse events, we modify the HFSM to account for the temporal distance to these events:

$$\begin{aligned}
 \Pr(Y_{s,j} = 1 | \mathbf{X}) &= \text{logit}^{-1}(\alpha_j + \boldsymbol{\beta}_{j[k]} \mathbf{X}_s), \\
 \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} &\sim N \left( \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix} \right), \\
 &\text{for } s = 1, \dots, S \\
 &\text{for } k = 1, \dots, K \\
 &\text{for } j = 0, \dots, \text{max temporal distance} - J
 \end{aligned} \tag{5.2}$$

We continue to label the response  $Y_{s,j} = 1$  for instances in which an event occurred at a location  $s$  and within a temporal distance  $j$  to the actual time of the pulse event. A crime event that occurs on the same day as the cultural event has a temporal distance equal to 0.  $\mathbf{X}$  is a vector of length  $k$  of predictor features such as the distance to business districts or the distance to places of worship from the location  $s$ .  $\boldsymbol{\beta}$  is a vector of coefficients that models the dynamic (changing in time) preferences for certain spatial features on the basis of the temporal distance  $j$  to the special event. We are now modeling how the CSS process changes on the basis of the how close in time the criminal is to the special event. By examining the resulting coefficients, we identify which features might attract or repel crime on the basis of the temporal proximity to the special event.

## 5.3 Simulated Data

### 5.3.1 Modification of Crime Simulation

Using the base simulation from Chapter 4, we introduce a series of special events that shift the preference for certain locations away from the regular patterns of the criminal agents. We retain the six variables in the spatial region  $x_1, \dots, x_6$  as distances to spatial features. Additionally, we retain the weekly seasonality used in Section 4.3.1.

#### Agents

We begin each simulation run with the citizen agents randomly distributed across the region with assigned home, work, and entertainment locations. Throughout the simulation, each citizen agent moves from their home location to their work, or entertainment, location with varied temporal patterns. The criminal agents begin each simulation in four distinct regions. Each criminal agent has a specified length of time for moving amongst the population. Within each daily temporal interval, the interaction between a criminal and citizen agent triggers the criminal agent's site selection process. If the criminal agent is sharing a location with more than one citizen, no crime will occur. Multiple citizen agents collocated at one location serve as defacto guardians. However, if a criminal agent shares a location with only one citizen, the criminal agent is more likely to initiate a crime if the location matches his preferences.

#### Simulation Scenarios

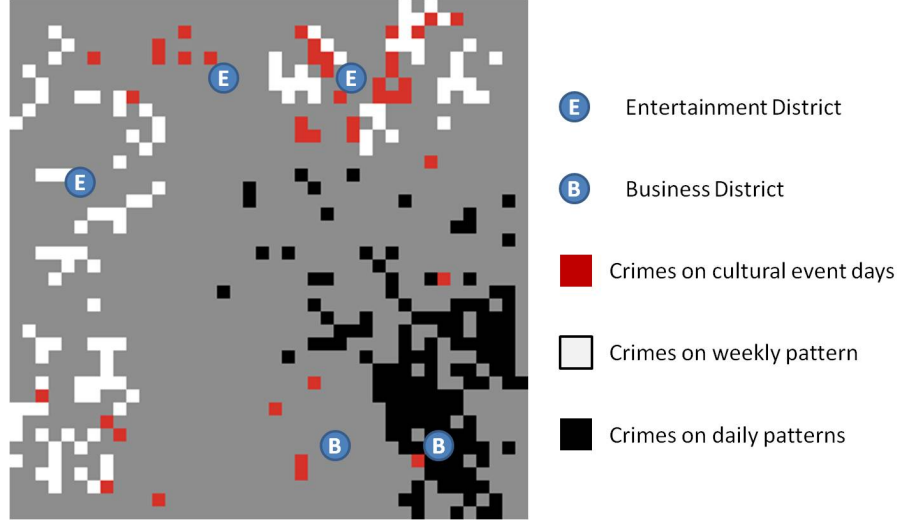
As in Section 4.3.1, the criminals still have separate spatial regions, and they retain a similar site-selection process related to event initiation. During the week, the business districts serve as attractors to criminal activity. Given a convergence event



between a criminal agent and a solo citizen, the criminal agents are more likely to initiate a crime if they are closer to the business districts. During the weekend interval, the business districts serve as repellers to criminal activity and the entertainment districts serve as attractors. For each simulation run, we introduce a series of special events that serve as spatial-temporal pulse events to the criminal agent's site selection process. During the temporal intervals of the special events, entertainment district three becomes the primary attractor for criminal activity. As criminal agents encounter citizen agents without a guardian, they are more likely to commit a crime if they are closer in distance to entertainment district three.

Figure 5.3 offers a visual depiction of one simulation run including the preference for a spatial-temporal shift to entertainment district three during temporal intervals with special events. The white squares indicate crime events occurring on the 7th day of the weekly pattern. The black squares indicate criminal events that occur during the remainder of the weekly interval. The red squares indicate the crimes that occur on the day of the simulated cultural events. Although the majority of the events associated with the simulated special event happen in the upper right quadrant vicinity entertainment district three, criminal agents still have a probability of executing a crime in areas that match their site selection preference.

We establish three training and testing scenarios. For each simulation run, we introduce four possible spatial-temporal pulse events across a simulated temporal horizon of 120 days. We isolate the criminal events on each special event day to include the preceding seven days. We use the eight day interval around the first special event as the first training set. For each simulation run, we record a vector  $Y_{s,t}$  such that a value of 1 indicates the presence of a crime at a location  $s$  at a point in time  $t$  while the value of 0 indicates the absence of a crime. Using the hierarchical feature-space model from Equation 5.2, we estimate the criminal



**FIG. 5.3: Sample Simulation Run with Weekly Crime Patterns and a Spatially Adjusted Crime Pattern During Temporal Intervals with Special Events**

agent's preference for a series of predictor features. For this simulation, we continue to use the six predictor features that include the distance to each entertainment and business district and the distance to the center of the simulation surface. Table 4.1 provides an overview of the variables used with the simulated dataset. We then evaluate each model against a test set using the eight day interval surrounding the second special event. The second training set includes the simulated data from both the first and second special events and predicts against the eight day period that includes the third special event. We build the final training set using the eight day windows preceding the first three special events and test against the final special event. Each test set represents a possible 12,800 possible space-time blocks for a criminal agent to consider as part of their site selection process in an eight day period that includes the programmed special event. Much as the feature-space model accounts for spatial preferences of the criminal agents by using distances to specific features, we hypothesize that modifying the feature-space model to include temporal distances from special events will account for the changes in criminal site

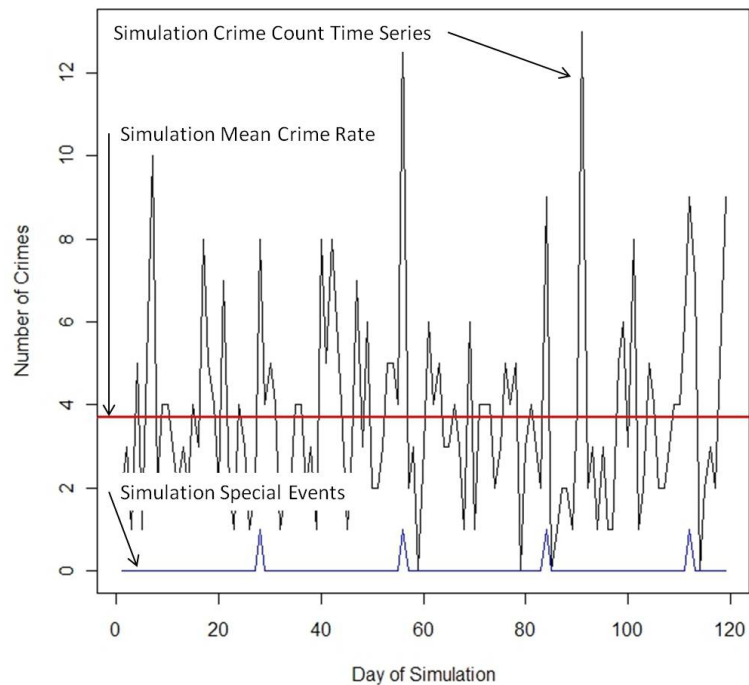
selection motivated by the changes in activity surrounding the special event.

### 5.3.2 Two-stage Analysis

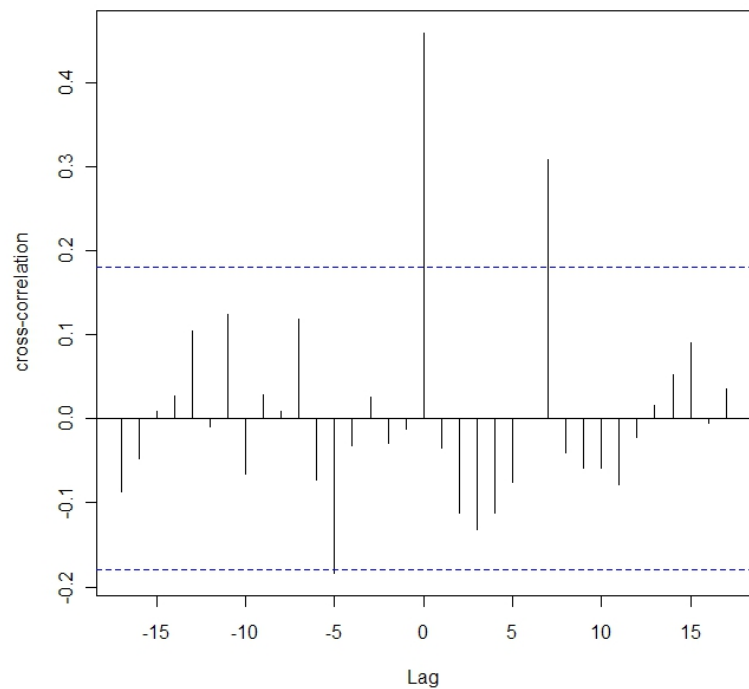
Aggregating the simulated crime event data at the daily level allows us to examine the relationships between the presence of a-priori special events and the potential spatial-temporal preferences within the criminal agent's site-selection process. Figure 5.4(a) depicts a daily count of crimes from one run of the simulation with four days acting as pulse events. The red line indicates the daily crime average over the entire run of the simulation. The blue line shows the days of the simulation with the special event near entertainment district three. Since we know the criminal agent rule set contains a shift in spatial preference within the temporal window of the special events, we hypothesize that the CCF will identify the relationship.

Using the CCF, we examine the relationship between our daily time series data with the simulated special event time series data. Figure 5.4(b) depicts the visual result for the CCF exploration of the simulated crimes with the pulse event time series. The distinct spike in the CCF strongly supports the hypothesis of a relationship between the two time series. Although the CCF helps identify relationships between potential pulse events and temporal shifts in the criminal's site-selection preference, we use the geographic mean center as a basic measure of central tendency to calculate changes in the spatial distributions over the temporal intervals. Figure 5.5 depicts the three distinct mean centers for crime clusters from the simulation data. Each geographic mean center is calculated on the basis of the corresponding temporal window to include weekly periodicity, daily activity, and the spatial-temporal pulse event.

The initial analysis identifies potential space and time patterns in the criminal agents' site selection process. Taking the temporal patterns found with the CCF and

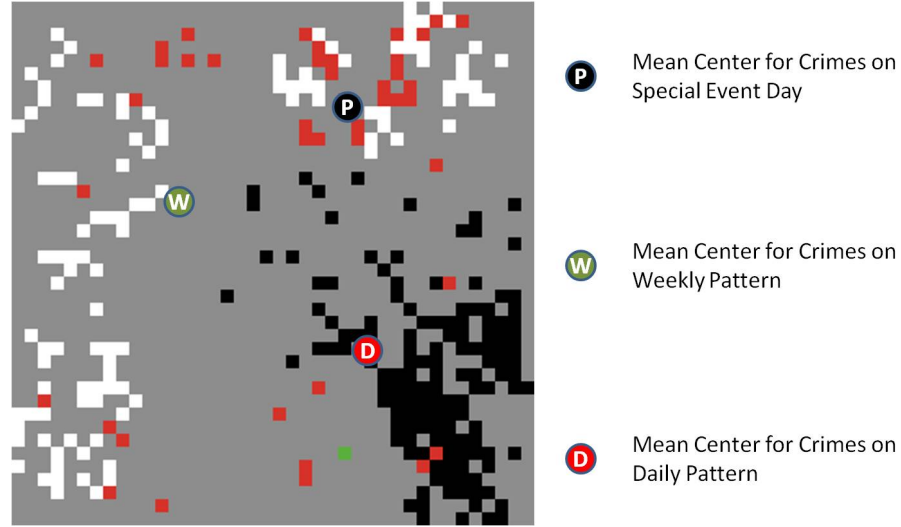


(a) Time Series for One Simulation Run



(b) Cross Correlation for Simulated Crimes with Pulse Events

**FIG. 5.4: Time Series and Cross Correlation Plots**



**FIG. 5.5: Geographic Mean Centers of Simulated Crime Clusters Depicting Impact of Pulse Events on the CSS Process**

the spatial component identified with the mean center calculations, we adjust the hierarchical framework around the feature-space model to account for the impact of the spatial-temporal pulse event.

### 5.3.3 Hierarchical Modeling

Returning to Equation 5.2, we set  $J = 8$ . The hierarchical feature space model now uses the temporal distance in days as the hierarchical component. We are now modeling how the criminal agent's site selection process changes according to the decreasing temporal distance to the pulse event. We continue to account for the spatial component of the special event since we still have the distance to

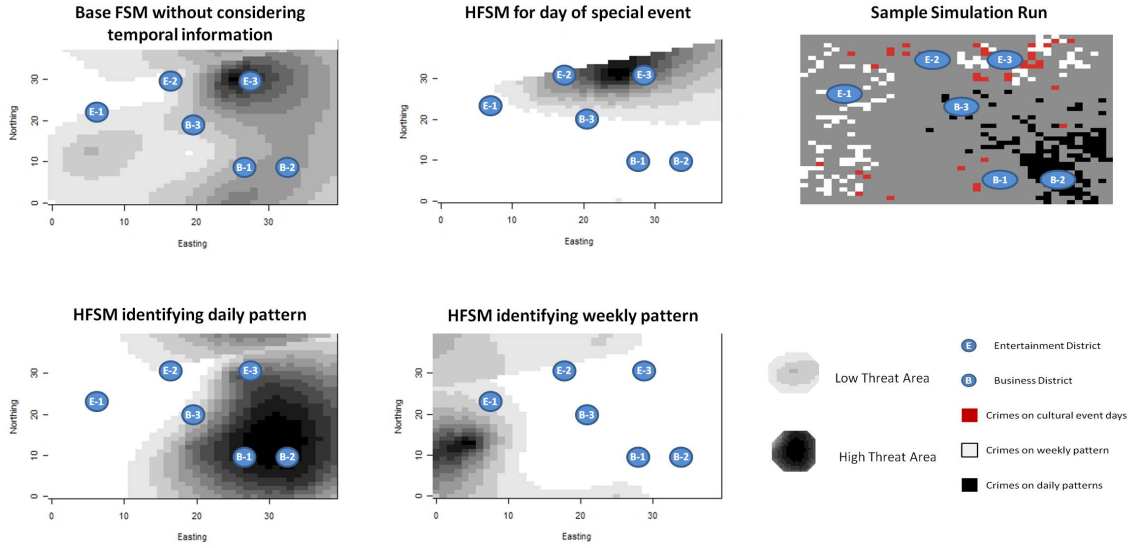
entertainment district three included as one of the six features:

$$\begin{aligned} \Pr(Y_{s,j} = 1) &= \text{logit}^{-1}(\alpha_j + \beta_{j[s]}\mathbf{X}_k), \\ \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} &\sim N\left(\begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix}\right), \\ &\text{for } s = 1, \dots, S \\ &\text{for } k = 1, \dots, K \\ &\text{for } j = 0, \dots, 8 \end{aligned} \tag{5.3}$$

Using the HFSM, we generate predictive threat surfaces that offer promise for improving resource allocation, especially for the day on which the pulse events occur. Significant change is seen in the shift and area reduction of the high-threat space for the HFSM for the day of the pulse event. Figure 5.6 provides a clear view of how the HFSM adapts to the temporal patterns in the criminal agents' site-selection process. Although the base feature-space model (FSM) accurately identifies three primary areas of high crime probability, the HFSM accounts for the temporal changes of the criminal agent's site selection process. The HFSM threat surface accurately identifies the northwest quadrant vicinity entertainment district three as the most likely area for crime on the day of the special event. Of special note is how the HFSM also accounts for the periodicity patterns in the criminal agents' site selection process by identifying both the daily and weekly patterns.

### 5.3.4 Significance Testing

Figure 5.7 depicts a space-time surveillance plot and measures the percentage of the possible space-time blocks covered within the study region across the study horizon. The HFSM is in red and clearly dominates the base FSM (in black), especially within the first 20% of the space-time area. We quantify the difference in predictive



**FIG. 5.6: Comparison of Threat Surfaces – HFSM and the Base FSM without Temporal Features.** Areas with higher probability of crime are shown with darker colors. The HFSM provides law enforcement personnel with a more focused threat surface to assist with resource allocation in high threat areas for specific temporal windows associated with special events.

performance by calculating the area under the curve for the two surveillance plots (Sing et al., 2005). Overall, the HFSM offers an 18% improvement in accuracy over the base FSM without temporal information.

Using the surveillance plot, we further examine model performance by focusing on how each model accurately identifies potential crimes in time and space. Using the partial area under the curve measure for the first 20% of the space-time area, we provide the law enforcement analyst with a measure that aids in resource allocation. Similar to the false positive point of interest introduced by (Kewley and Evangelista, 2007), a high partial area under the curve measure of the surveillance plot identifies a model which is more accurate in the higher threat areas. Given this information, law enforcement personnel can select the model with the lowest threat surface for the temporal interval surrounding the spatial-temporal pulse event.

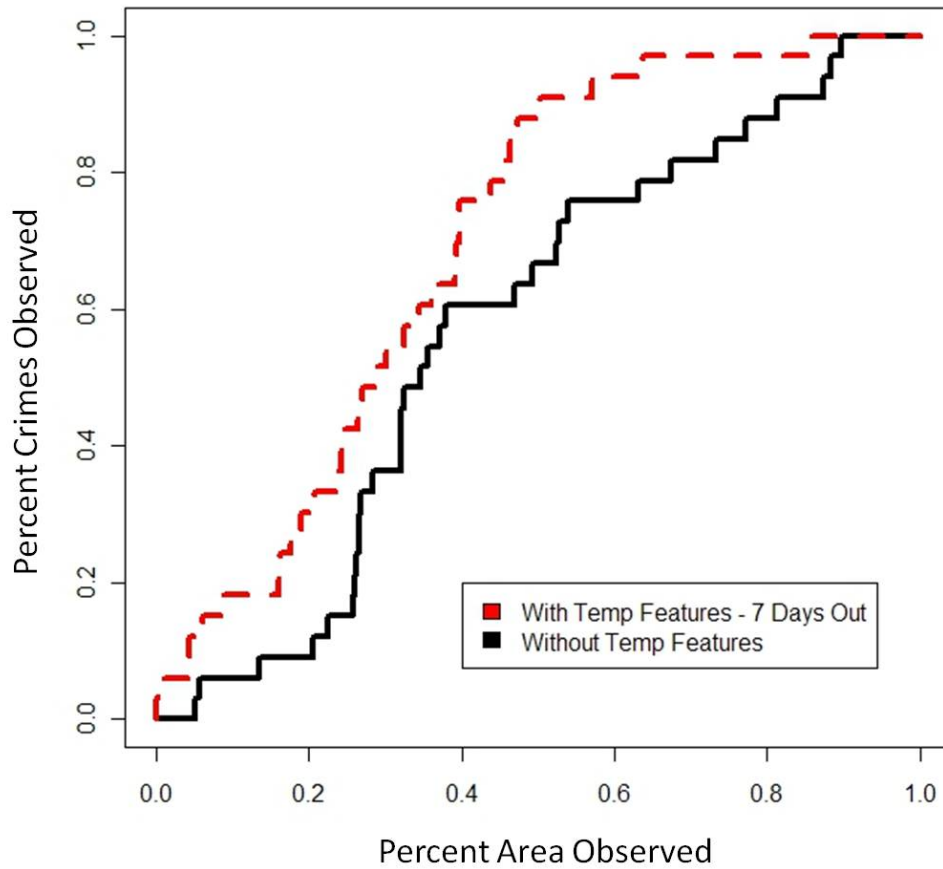


FIG. 5.7: Surveillance Plot – HFSM versus Base FSM without Temporal Information

TABLE 5.1: Partial Area Under the Curve for Surveillance Plots

Training Events	Base FSM	HFSM
1 Event	0.034	0.041
2 Events	0.0163	0.0261
3 Events	0.0108	0.0326



Table 5.1 depicts the accuracy of the HFSM across all three scenarios for a sample simulation run. Regardless of the size of the training set, the HFSM offers better predictive performance on each test set when compared to the base feature-space without temporal information. And while the partial area under the curve measure seems higher for the first test set, the greatest improvement in performance is seen when using information from the three initial pulse events to predict against the fourth pulse event window. On the final test set, the HFSM offers significant improvement over the model built without considering temporal information. The performance of the HFSM over the base FSM holds across multiple simulation runs with varied special event locations and temporal patterns.

## 5.4 Discussion

This chapter expands the second methodology for combining temporal features with spatial information in order to enhance an analysts understanding of the CSS process. We continue to use the two-stage analytical process suggested by Schabenberger and Gotway (2005), however, we use the cross-correlation function to identify potential relationships between our crime time series and a time series of special events. Using the geographic mean center conditioned on the temporal windows surrounding the special event helps us in identifying changes in the spatial distribution of criminal activity. We retain the basic hierarchical framework of the feature-space model from Chapter 4 but integrate the temporal distance to the potential pulse event as the hierarchical component of the model.

We modify the agent based simulation to include a series of special events that serve as spatial-temporal pulse events on the regular patterns of the criminal agents site selection process. Using a series of three training and test sets developed over multiple simulations, we compare the performance of the base FSM with the HFSM.

We review the performance of the model across three pulse-event scenarios: a pulse event occurring every four weeks, a pulse event occurring every twenty-five days, and a pulse event every thirty days. Although the HFSM outperforms the base model across all three scenarios, the most significant gain comes with using training data aggregated across multiple pulse-event windows to develop the model. Against the simulated data, the HFSM provides an 18% improvement in predictive performance against the base FSM without temporal information.

## CHAPTER 6

### Method 3 – Of Neighbors and Recent Past

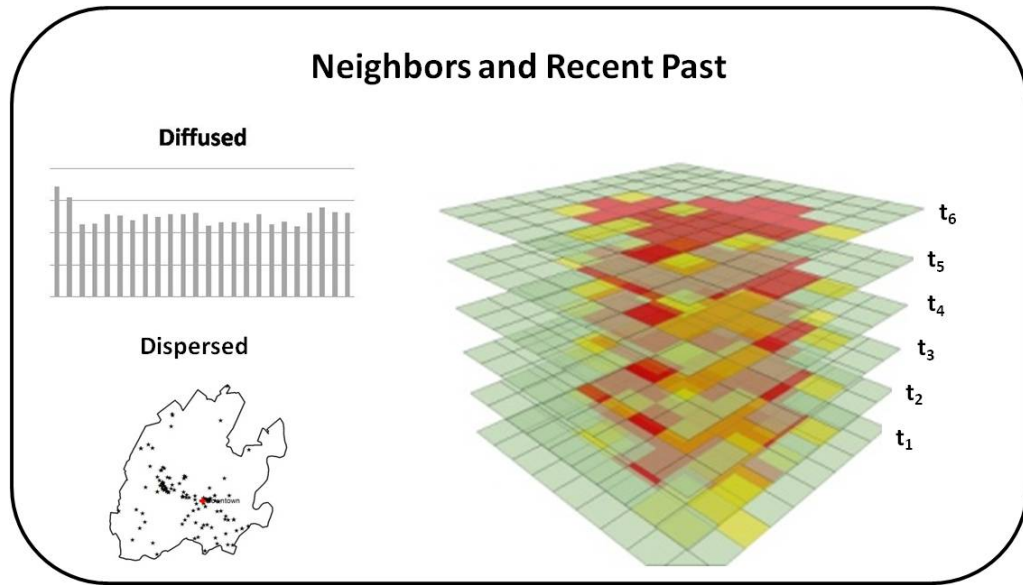
*It is better to ask some of the questions than to know all the answers.*

- James Thurber

The final question of this dissertation is how to model the criminal site-selection process across a continuous time horizon. While the two previous methods used hierarchical modeling to account for temporal intervals and temporal distances, this method merges the feature-space model of Liu and Brown (1998) with the Markov random field construct of Zhu et al. (2006) to model the criminal's preference for initiating a crime across a continuous temporal horizon. By adapting theoretical and computational work from disease mapping and environmental studies, we develop a Bayesian hierarchical feature-space model (BHFSM) for the criminal event prediction problem in order to examine both parameter estimation and predictive inference. We also expand the base feature-space model with indicator functions to answer the two primary questions for this chapter: How can we model the impact of neighborhood patterns within the feature-space model to account for local activity on the criminal's spatial-temporal decision process? And can we model the impact of the recent past events within the feature-space model to account for local temporal patterns on the criminal site-selection process? The following sections review a Markov random field (MRF) treatment of the feature-space model, modify the base feature-space model with temporal and neighborhood indicator functions, and expand on the application of the surveillance plot and the performance-limit ratio (PLR) for measuring predictive performance across the space-time domain. After discussing modifications to the agent based crime simulation, we conclude the chapter with an examination of the two models against a simulated data set.

## 6.1 Formulation

Although much theoretical and practical work has been done on the use of Bayesian hierarchical modeling for geostatistics and disease clustering, the application within the construct of the criminal site selection problem has been limited. Random



**FIG. 6.1: Method 3 – Feature-Space Modeling Using Neighbors and Recent Past.** Beginning with a series of observed crimes with a diffused rate of intensity across the temporal horizon, we aggregate data at the census-block level and model the CSS process using a Bayesian framework. We compare the Bayesian method against an expanded feature-space model using temporal and neighborhood indicator functions. We model historical data and develop predictive threat surfaces for future events. Finally, we compare the models using the surveillance plot, AUC, and PLR.

space-time fields have been used to represent environmental processes such as oil field development, environmental-hazard analysis, and disease outbreaks. Besag's work (1974) sets the baseline for studies of spatial situations. Besag establishes seven illustrations of spatial-point processes and regions. The example of irregular regions with discrete and continuous variables describes the study area for incident activity used in this dissertation. While Besag's point lattice lays the groundwork for the spatial-temporal modeling in Chapters 4 and 5, the discussion of irregular lattice systems more accurately portrays the study regions and efforts for this methodology. Event data for which spatial occurrences can be arranged as a regular lattice in  $\mathbb{R}^d$  is a close match to a time series of events observed at discrete intervals (Cressie, 1993).

Recent work on point processes depicts random fields as a secondary structure that would result from an aggregation process of point counts or marked sums. For count data, we construct the random field along a regular or irregular lattice structure using the construct provided by Illian et al. (2008).

$$Y(x_t) = N(B + x_t)$$

The value of  $Y(x_t)$  is simply the count of events within the defined spatial structure during temporal interval  $t$ . The spatial structure is a sphere  $B$  at location  $x$  with a defined radius (Illian et al., 2008). For this dissertation, we use both a regular lattice draped across the study region and an irregular lattice defined by political and cultural boundaries.

Adapting theoretical and computational work from disease-mapping and environmental studies, we use a BHFSM on the criminal event prediction problem for examining both parameter estimation and predictive inference. The BHFSM is a limited feature-space information logistic regression model with an autoregression on

the state of the neighboring locations across an irregular lattice at discrete temporal intervals. As a means of comparison, we provide an alternative method for including neighborhood and temporal information with the base feature-space model similar to the proposal made by Diggle et al. (1998).

## 6.2 Methodology

Hierarchical models allow us to deconstruct complex problems into a series of smaller tractable problems. Using the methodology developed by Wickle (2003), we formulate the criminal site selection problem into three basic stages: a data model, a process model, and a parameter model. Our data model accounts for our knowledge of the spatial-temporal patterns of crime within the study region. The process model provides insight to the criminal site selection process while accounting for spatial and temporal effects. Finally, our parameter model accounts for the uncertainty in both the data and process models (Wickle, 2003).

We use a Markov random field construct to account for spatial-temporal effects by assuming that the likelihood of a crime at a specific location is dependent only on its neighbors and its previous temporal state (Zhu et al., 2006). Recent work on point processes uses Markov random fields as a secondary structure that results from an aggregation process of event counts. For our crime data, we construct the Markov random field along an irregular lattice structure defined by political and cultural boundaries using the construct provided by Illian et al. (2008).

### 6.2.1 Bayesian Hierarchical Modeling

The primary goal of this chapter is to develop a Bayesian hierarchical model that uses the feature-space methodology to accurately predict crime events across an irregular lattice while providing insight into the criminal site-selection process. To estimate

the criminal's spatial preferences, our **data model** represents the criminal's site-selection process as a binary random variable where  $Y_{s,t} \in 0, 1$  is the observation of the presences, or absence, of crime at location  $s$  at time  $t$  given a set of features  $\mathbf{X}$ .

$$Y_{s,t}|\mathbf{X} \sim \text{Bern}(\mu_{s,t}) \quad (6.1)$$

For our least complex model, we assume that the probability  $\mu_{s,t}$  is a function of the criminal's preferences for certain features and a random effects term. Mathematically, we represent the **process model** as:

$$\begin{aligned} \mu_{s,t} &= \text{logit}^{-1}(\beta_0 + \beta_1 X_{s1} + \dots + \beta_k X_{sk} + \theta_{s,t}), \\ &\text{for } s = 1, \dots, S \text{ and} \\ &\text{for } t = 1, \dots, T. \end{aligned} \quad (6.2)$$

Equation 6.2 uses a set of features  $\mathbf{X}$  as a vector of length  $k$  for each location  $s$  combined with the estimated  $\beta$  values from the parameter model to estimate the probability  $\mu_{s,t}$ . For this dissertation, we use a set of demographic variables as a surrogate for potential features that might be considered by the criminal in their site-selection process. Analyzing previous criminal event data gives us a method to account for the criminal's site-selection process. By modeling the relationship between the features and the probability of crime, we estimate the preferences criminals have for locations with a specific set of features. However, just as the criminal's preferences for certain locations might change depending on proximity to freeways or vacant houses, the criminal site-selection process can also change depending on the time of day or other seasonal events (Rossmo et al., 2005; Gorr, 2009a). The variable  $\theta_{s,t}$  provides a method for including other random effects.

The first random effect considered is the temporal component. We consider a temporal effect  $g_t \sim N(g_{t-1}, \tau_g)$  where the value of  $g_t$  is normally distributed from



the previous state with a standard deviation  $\tau_g$  estimated in the parameter model using a non-informative prior. Based on previous research, we believe that criminal activity often precedes criminal activity (Gorr et al., 2003; Eck et al., 2005). Using this temporal component allows us to account for periods of criminal activity that match the routine activities and population dynamics of the study region. We will discuss the initial conditions for the variance estimates in the parameter model.

We use a Markov random field (MRF) construct as the second random effect by assuming that the likelihood of a crime at a specific location is dependent only on its neighbors and its previous temporal state (Zhu et al., 2006). Recent work on point processes uses MRFs as a secondary structure that results from an aggregation process of event counts. For our crime data, we construct the MRF along an irregular lattice structure defined by political and cultural boundaries using the construct provided by Illian et al. (2008). We consider a MRF effect that accounts for the past value at the location  $s$  and the second-order neighbors such that  $\omega_s \sim N(\omega_{j-1}, \tau_o)$ . The index  $j$  accounts for the second-order neighbors of location  $s$ . The inclusion of the neighborhood spatial effects gives us a method to include criminal repeat information into the feature-space model. Studies on criminal repeats have shown that for short temporal intervals, locations that have experienced crime have an increased likelihood for repeat victimization (Townsend et al., 2000).

The third random effect considered for this article is an interaction term. We consider an interaction term  $\psi_{s,t} \sim N(0, \tau_p)$ . The interaction term is uncorrelated but can identify potential spatial-temporal interactions within the data that are not accounted for in the base feature-space model (Lawson, 2009). The final random effect is an uncorrelated error term  $v_s \sim N(0, \tau_v)$  that accounts for any uncorrelated spatial components of the criminal site-selection process. The research design section outlines the four primary models considered for this article using different combinations of these random effects.

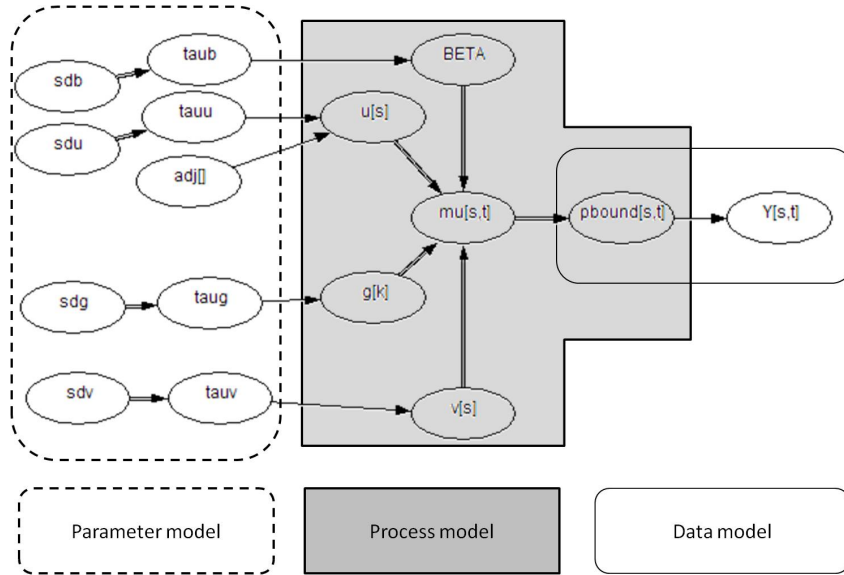
Finally, we specify the *parameter models* by establishing the initial distributions for the parameters. As seen in Figure 6.2, the  $\beta$  vector appears in the process model. However, we provide initial estimates for the individual  $\beta$ s within the parameter model. Estimating the  $\beta$  values increases the complexity of the parameter model, since for both the long term and short term data study, we initially estimate each  $\beta$  for each feature during the model fitting phase. In order to reduce the computational requirements, we substitute a feature-space prior calculated from linear model regression (Lunn et al., 2000). The initial assumptions for the parameter model follow:

$$\begin{aligned}
 \beta &\sim N(\hat{\beta}, \tau_b) \\
 \tau_b &\sim N(0, svb), svb \sim U(0, 10) \\
 \tau_u &\sim N(0, svu), svu \sim U(0, 10) \\
 \tau_g &\sim N(0, svg), svg \sim U(0, 10) \\
 \tau_o &\sim N(0, svo), svo \sim U(0, 10) \\
 \tau_p &\sim N(0, svp), svp \sim U(0, 10)
 \end{aligned} \tag{6.3}$$

The parameter model sets the initial conditions for the simulation methods used to estimate the process and data model and completes the model hierarchy (Wickle, 2003). More details on the simulation methods can be found in Lawson (2009) and Kery (2010).

We provide additional details on the random effects considered for this dissertation below. Figure 6.2 provides a graphical representation, developed using WinBUGS, of one model considered for this dissertation.

Bayesian methods provide a means to calculate the posterior distribution from our three stage-hierarchical model. Using the example from Wickle (2003), our posterior distribution is proportional to our data model conditioned upon the process and



**FIG. 6.2: Directed Acyclic Graph for Bayesian Hierarchical Feature-Space Model.** The  $\beta$  node provides the most complexity for the model. For both the long term and short term data study, we initially estimate each  $\beta$  for each feature during the model fitting phase. In order to reduce the computational requirements, we substitute a single prior to represent the feature-space preferences for each location on the basis of a GLM regression (Lunn et al., 2000).

parameter models times the process model conditioned upon the parameters:

$$\begin{aligned}
 [process, parameters|data] &\propto \\
 [data|process, parameters] &\times \\
 [process|parameters][parameters] &
 \end{aligned}
 \tag{6.4}$$

Since our goals in modeling criminal site-selection problems include both predictive inference and parameter understanding, we desire to solve for the left hand side of Equation 6.4. However, the complexity of the posterior distribution makes obtaining a closed form solution difficult. Using simulation methods, we obtain samples that provide estimates of our target variables (Lawson, 2009). We initiate our prior distributions using either empirical knowledge from the data or expert knowledge from the study domain. When expert knowledge is not available, we use flat or improper priors. For example, we initiate our prior distributions for the  $\beta$  coefficients using the estimates obtained from the generalized linear model regression. However, we initiate our priors for the standard deviations of the parameter models with non-informative uniform distributions.

Following work from disease mapping and geostatistics, we examine four models of random effects for our variable  $\theta_{s,t}$ . The models considered provide several methods for including other random effects (Lawson, 2009). The four models considered for random effects include:

- A time-varying trend  $g_t$  plus an uncorrelated error  $v_s$
- A Markov random field  $\omega_s$  accounting for the sum of the neighboring effects at a previous time plus  $v_s$
- $g_t$  plus  $\omega_s$  plus  $v_s$
- $g_t$  plus  $\omega_s$  plus  $v_s$  and an interaction term  $\psi_{s,t}$

Figure 6.2 displays a graphical representation of the third model developed for this dissertation without an interaction term.

### 6.2.2 Expanded Feature-Space Model

We use  $Y_{s,t} = 1$  for the presence of a crime during a given temporal interval and  $Y_{s,t} = 0$  as the absence of crime at the site  $s$  within the specified temporal interval  $t$ . We model the criminal events within the study area and horizon as a binomial distribution similar to Liu and Brown (2003) and Gelman and Hill (2007) such that:

$$\begin{aligned} Y_{s,t} &\sim \text{Binomial}(n_{it}, \theta_{st}) \\ \theta_{st} &= \text{logit}^{-1}(\alpha + X_{s,t}\beta) \end{aligned} \tag{6.5}$$

We expand the basic binomial model to replicate the feature-space model with temporal and neighborhood indicators as proposed by Diggle et al. (1998) and developed by Christakos (1992):

$$\begin{aligned} Y_{s,t} &\sim \text{Binomial}(n_{st}, \theta_{st}) \\ \theta_{st} &= \text{logit}^{-1}(\alpha + \beta_{k-2}X_{k-2,st} + \beta_{k-1}X_{k-1,st} + \beta_kX_{k,st}). \end{aligned} \tag{6.6}$$

where  $\beta_{k-1}$  is the spatial autoregressive coefficient and  $\beta_k$  represents the temporal coefficient. The indicator function for  $X_{s,t,k-1}$  can be constructed as an MRF such that

$$X_{s,t,k-1} = \begin{cases} 1 & \text{if } \sum_{j \sim s} Y_{s,t-1}Y_{j,t-1} + (1 - Y_{s,t-1})(1 - Y_{j,t-1}) \geq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{6.7}$$

Equation 6.7 provides us a positive response if the second-order neighbors of  $Y_{s,t}$  have observed an incident during the previous temporal period. A variation of the above indicator function could include the weighted sum of the clique to further increase the impact of the neighbor structure on the model.

The indicator function for  $X_{s,t,k}$  is a representation of the temporal relationship and can be written as a first-order autoregressive indicator functions such that:

$$X_{i,t,k} = \begin{cases} 1 & \text{if } Y_{i,t-1} \geq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (6.8)$$

This equation provides a positive response if the previous state of  $Y_{s,t}$  was an observed incident. Variations of the above indicator function could include second-order autoregressive or moving-average functions to provide increased fidelity to the temporal relationships in the CSS processes. Adding the temporal regression variable gives us an additional avenue to explore temporal relationships and dependencies in the CSS process. This method offers fast sampling from the dataset to reduce the computational time while still accounting for the state of the neighborhoods and the previous state of the studied locations. This new expanded feature-space model (EFSM) builds upon the construct of an MRF because the response observed at site  $s_i$  at time  $t$  is conditional on the state of the neighbors at time  $t - 1$  and the state of  $s_i$  at time  $t - 1$ .

### 6.2.3 Model Comparison

For this dissertation, we compare each model's predictive performance against a test set from the dataset. For the long term study, we use a 365 day temporal window for model fitting and then evaluate against a ninety day test. For the short term study, we use a thirty day temporal window surrounding special events in Charlottesville for model fitting and then evaluate against the thirty day temporal window surrounding the same special event in the following year.

Prior to comparing predictive performance, we use a goodness of fit measure to evaluate each model. Borrowing from conventional generalized linear modeling,

we use deviance as a measure of how well the model fits the data. In the software used for this dissertation, we can expect the deviance to decrease by one for each predictor added to the model (Gelman and Hill, 2007).

As an additional method for comparing goodness of fit, we use the mean squared predictive error (MSPE). Given our known spatial-temporal dataset from the test period,  $\mathbf{Y}$ , our estimated spatial-temporal dataset,  $\hat{\mathbf{Y}}$ , and a number of observations  $m$  from a simulation sample of  $G$ , we use Lawson's (2009) formulation such that:

$$MSPE = \frac{|\mathbf{Y} - \hat{\mathbf{Y}}|^2}{(G \times m)} \quad (6.9)$$

## 6.3 Simulated Data

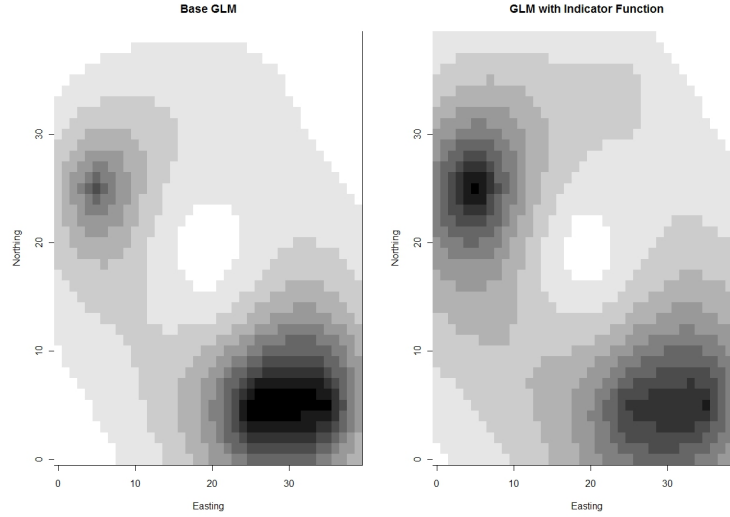
### 6.3.1 Modifications of Crime Simulation

Returning to the original simulation from Chapter 4, we adjust the criminal agents' rule set back to a simple weekly seasonality. The probability of a criminal agent's decision shifts from the lower right portion of the simulation to the upper left quadrant with a seasonal temporal pattern. Each run of the simulation still produces 120 days of activity.

### 6.3.2 Two-Stage Analysis

We aggregate the simulated crime data and model using both the base feature-space model (FSM) presented in Chapter 3 and the EFSM from Equation 6.6. The predictive surfaces shown in Figure 6.3 represent the high probability areas of crimes calculated using the resulting models. Although the FSM identifies the two areas of crime, the EFSM more accurately models the shift in preference along a continuous time analysis from the lower right corner to the upper left corner and

back in accordance with the criminal agent’s rule set. The performance of the models against the test data from the simulation indicates that the EFSM provides a 13% increase in predictive performance across the space-time surface when compared to the FSM without indicator functions.



**FIG. 6.3: Predictive Threat Surfaces Based on FSM and EFSM for Simulated Data.** Areas with higher probability of crimes are shown with darker colors.

## 6.4 Discussion

For strategic-level crime monitoring, the EFSM offers a methodology for modeling criminal activity within spatial regions using continuous time horizons. For this research, we expanded the base feature-space model to include indicator functions that provide insight to both spatial and temporal patterns within the criminal site-selection process. We applied the methodology to a simulated data set with specific temporal patterns in the CSS process. The EFSM provides significant lift when compared against a test set. The use of the EFSM construct provides a method that is computationally faster than the Bayesian methodology while providing improve-



ments in predictive performance when compared to the base feature-space model. The inclusion of temporal and neighborhood indicator functions allows law enforcement personnel to account for recent past activity at a given location and within the surrounding neighborhoods.

# CHAPTER 7

## Crime Event Modeling

*The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a body of data.*

- John Tukey

This chapter explores the results of applying the methodologies described in this dissertation to the challenges of spatial-temporal event prediction of crime within a medium size city. For this chapter, we consider crime events as a realization of a criminal’s site-selection (CSS) process based upon spatial and temporal features. Using a hierarchical feature-space model (HFSM), we account for shifts in the CSS process across discrete temporal intervals and in temporal proximity to special events. We also consider the contribution of using a Bayesian hierarchical feature-space model (BHFSM) for examining criminal events across a continuous temporal horizon. We use receiver operating characteristic (ROC) curves, the area under the curve (AUC), and the performance limit ratio (PLR) as means of comparison across all methodologies. Although the hierarchical and Bayesian methodologies provide improvements over the basic feature-space model when temporal patterns exist, the expanded feature-space model with neighborhood and temporal indicator functions (EFSM) provides an extremely fast and sufficiently accurate alternative to the computationally demanding HFSM and BHFSM.

## 7.1 Introduction of Dataset

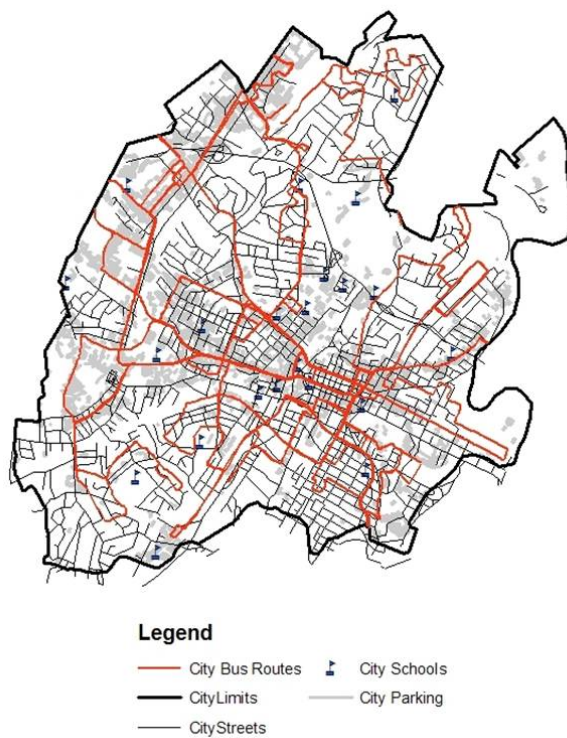
The primary source of data for this research is an incident database for the city of Charlottesville, Virginia police department. The database is maintained by the Department of Systems and Information Engineering at the University of Virginia. We sample the complete dataset to develop a subset that contains a time horizon of three years and over 5,000 incidents. Crime types within the dataset include burglaries, assaults, and thefts from automobiles. The crime dataset from Charlottesville is suitable for modeling criminal activity since it contains multiple actors within a limited temporal horizon spread across a distinct spatial region with a variety of incident types.

We use a geographic information system (GIS) to geocode each incident. Geocoding allows us to convert the street address of the criminal events to a set of coordinates useful for plotting in our GIS. The final stage of preprocessing the event dataset involves combining the crime events with a uniform null grid of marked points across the study region (Chawla et al., 2001; Liu and Brown, 2003; Brown et al., 2004). We measure spatial distances from both the null grid and the incident layer to various features within the study region. For demographic variables obtained from census data, we assign values to our uniform null grid and incidents on the basis of the containing or nearest polygon's values. We use spatial data from the University of Virginia Geostat Library to establish the base feature-space layer of distances to natural and man-made features. For the demographic data, we use block-level data from the 2000 United States Census. Table 7.1 provides an overview of the variables contained within the crime dataset.

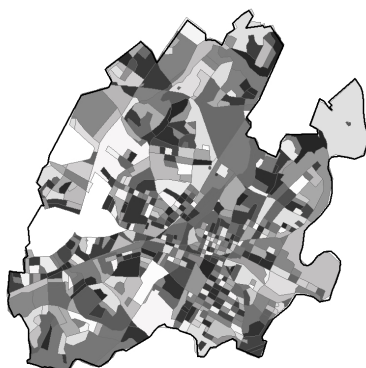
Figure 7.1(c) depicts the resulting plot of crime events and the null grid. The null grid serves as the base layer for locations in space and time where the response  $Y_{s,j}$  from Equation 4.2 is equal to 0. Calculating the feature-space distances for the null grid provides law enforcement personnel with a means of producing threat surfaces that depict the probability of crime in areas that have not experienced criminal activity, but contain similar features to high crime areas. Figure 7.1(b) depicts the spatial distribution and census block-level resolution of the cultural features contained within the census data. Figure 7.1(a) depicts a base map of Charlottesville with some of the man made spatial features used for this dissertation.

**TABLE 7.1: Predictor Variables for Crime Dataset**

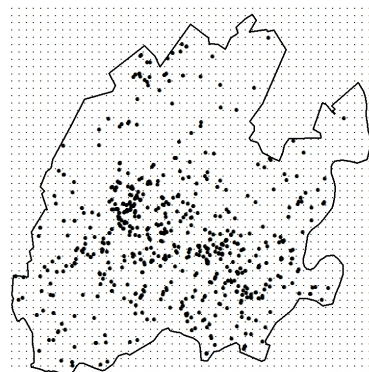
<b>Variable Name</b>	<b>Description</b>
<i>Disstreet</i>	Distance from point to nearest street
<i>Disrail</i>	Distance from point to nearest railway
<i>Disbus</i>	Distance from point to nearest bus route
<i>Dischurch</i>	Distance from point to nearest church
<i>Disschool</i>	Distance from point to nearest school
<i>Income</i>	Average household income for the nearest census tract vicinity point
<i>population</i>	Total population for the nearest census tract vicinity point
<i>PerBlack</i>	Percent of population vicinity point – Black
<i>PerAsian</i>	Percent of population vicinity point – Asian
<i>Housing</i>	Total number of houses for the nearest census tract vicinity point
<i>PerRental</i>	Percent of housing classified as rental in the nearest census tract
<i>Rent</i>	Average of median rental rates for census blocks consolidated at tract level
<i>Housevalue</i>	Average of median housing values for census blocks consolidated at tract level
<i>PerVacant</i>	Percent of houses not occupied during time of census
<i>Response</i>	Binary – did a crime occur? 0 – no crime 1 – crime



(a) Sample of Base Layer Spatial Features



(b) Census Blocks for Demographic Data



(c) Incidents with Null Grid

**FIG. 7.1:** Mapping Steps for Feature-Space Analysis

## 7.2 Two-stage Analysis

As indicated by the dissertation map shown in Figure 3.1, analyzing the data spatially and temporally allows us to identify both spatial and temporal patterns for further investigation and modeling. The two-stage analysis also aids in answering many of the questions asked by law enforcement analysts:

- What is the appropriate resolution of temporal intervals for study?
- Within the given spatial environment, are incident rates higher during specific temporal intervals?
- Do daily patterns exist in event likelihood for the different spatial areas with similar attributes?

The initial temporal analysis of each type of criminal incident results in a series of conditioned plots that provide insight into the variations across the different temporal intervals. Figure 7.2 depicts the monthly, daily, and hourly incident rates for assaults in Charlottesville. At first glance, Figure 7.2 appears to depict some variance across each level of temporal resolution. Figure 7.2(c) depicts an acute temporal pattern, while Figure 7.2(b) is a focused temporal pattern. However, the variance in the monthly rates is deceptive. Conditioning the months for the number of days produces a more balanced result, as seen in Figure 7.3. Since visual analysis alone does not identify all the patterns that might exist within the dataset, we use the autocorrelation functions to find patterns of seasonality and determine the appropriate temporal resolution for the hierarchical model. The autocorrelation plot and partial autocorrelation plots depict strong weekly seasonality when used to analyze the daily crime rates in Charlottesville. Again, we are looking for the intersection of focused temporal patterns with spatial clustering as seen in Figure 3.2. Now that we have an understanding of the patterns at different temporal resolutions, we



FIG. 7.2: Distribution of Crimes Across Different Temporal Resolutions

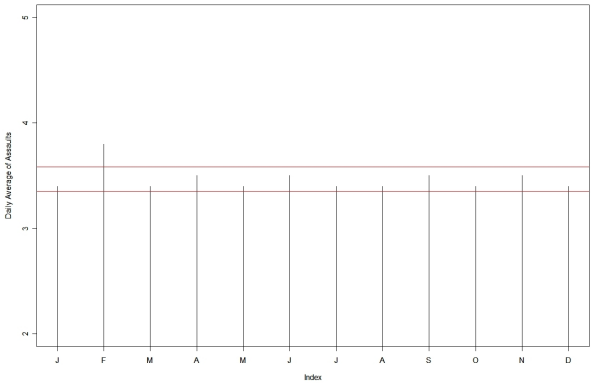
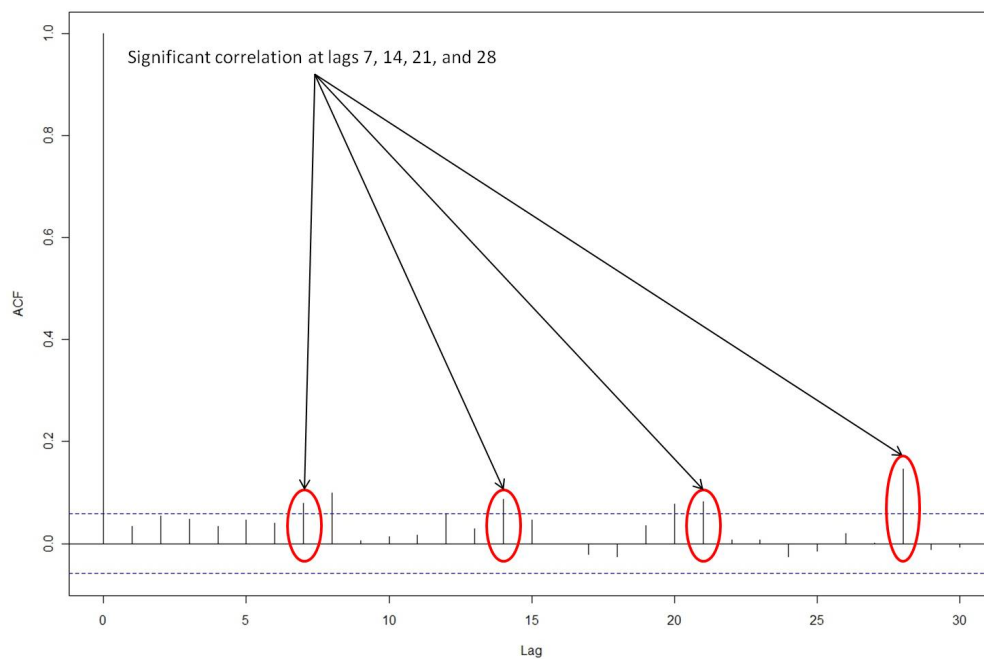
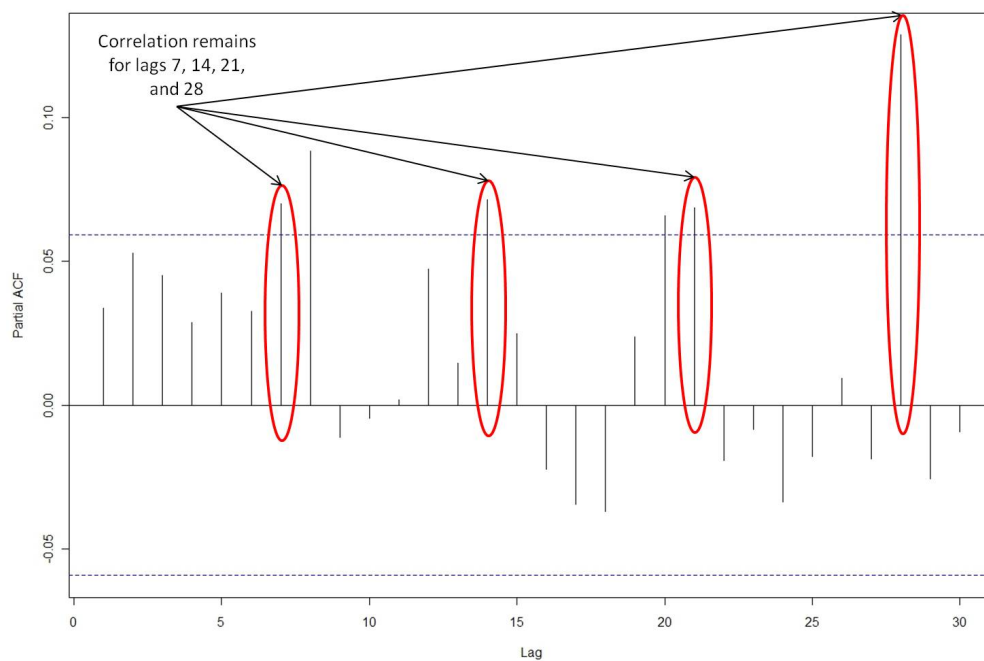


FIG. 7.3: Monthly Average of Assaults Conditioned upon Number of Days





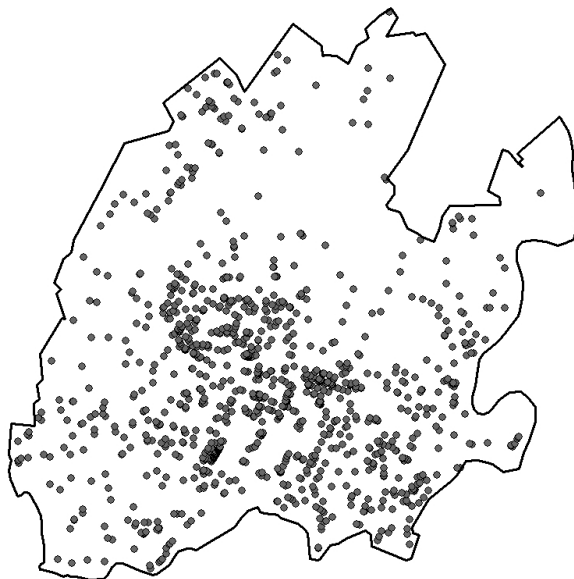
(a) Autocorrelation for Assaults



(b) Partial Autocorrelation for Assaults

**FIG. 7.4: Correlation Plots for Assaults Viewed at Daily Temporal Resolution**

can examine the spatial distribution of the events within the appropriate temporal intervals. A straight plot of all events in the study region shows distinct clustering in the vicinity of the University of Virginia (UVA) and downtown Charlottesville. However, since this plot includes all events aggregated across the temporal horizon, it fails to depict the possible temporal components of the CSS process. Each tem-



**FIG. 7.5: Point Plot of All Assaults within Study Region and Horizon**

poral interval offers some variance across space and time. Using the assaults for Charlottesville, we first look at the plots for the geographic mean center measure of events in our three primary temporal intervals. Although each temporal interval's geographic mean center is clustered near the University, the dispersion at the hourly and the daily levels indicate that temporal patterns might exist in the CSS process. We confirm evidence of clustering using the nearest neighbor analysis.

For examining the likelihood that the events are clustered during a specific temporal interval, we examine the nearest neighbor distances and compare to the expected nearest neighbor distance for a completely spatially random distribution

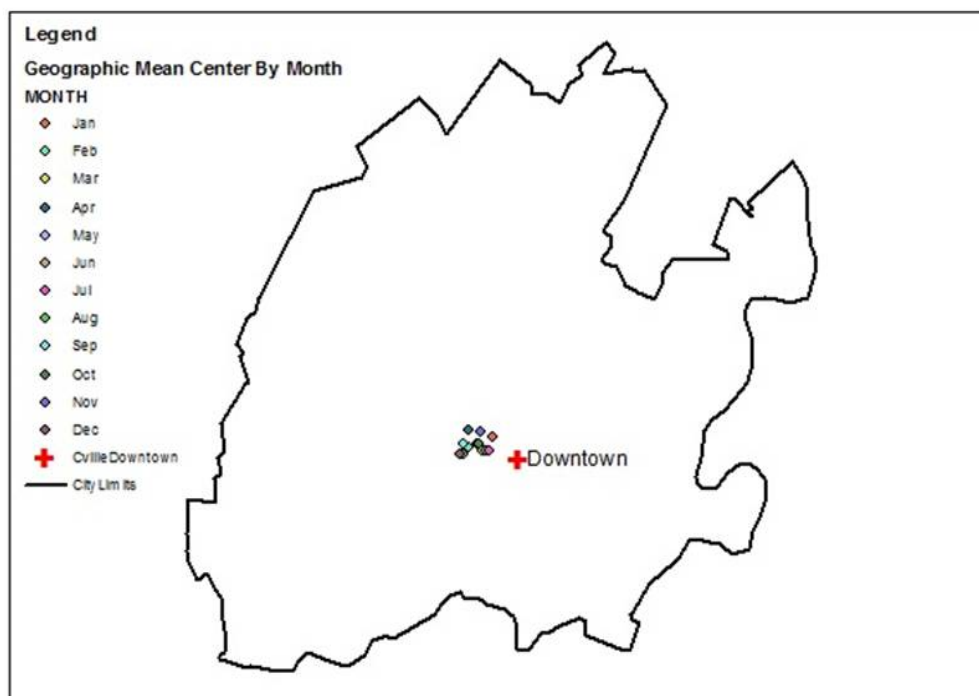


FIG. 7.6: Monthly Distribution of Geographic Mean Centers

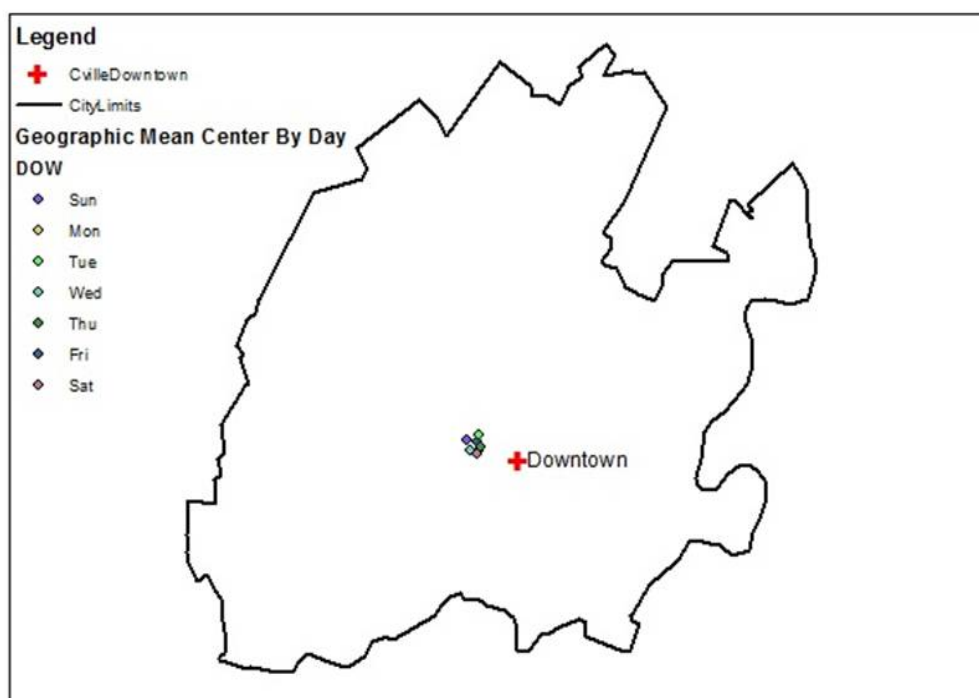
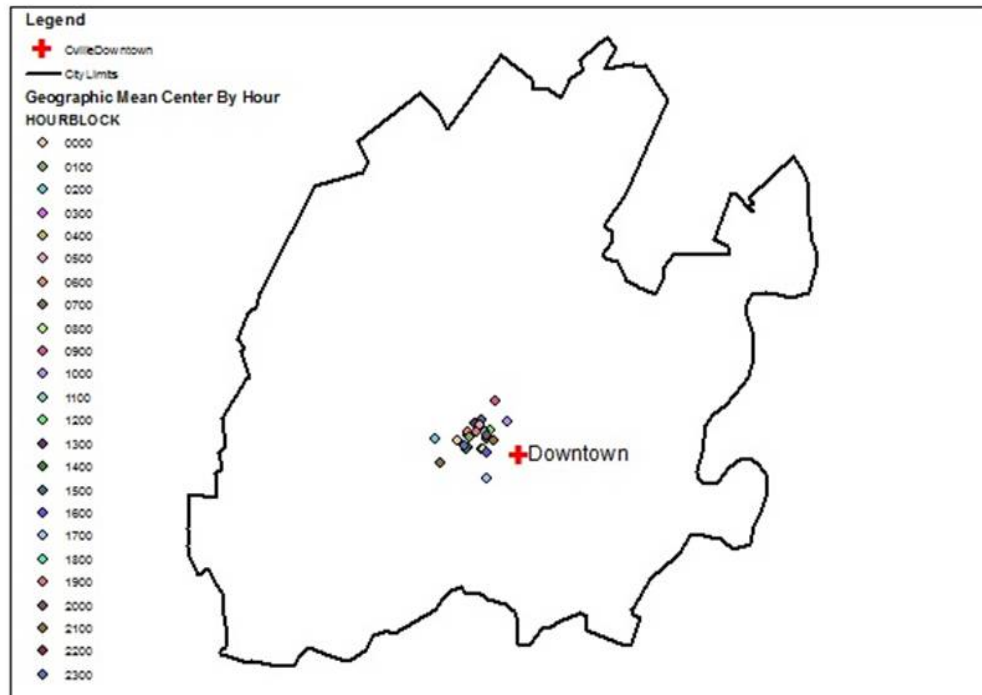


FIG. 7.7: Daily Distribution of Geographic Mean Centers



**FIG. 7.8: Hourly Distribution of Geographic Mean Centers**

of events. Using the Average Nearest Neighbor index in ArcGIS, we can determine the likelihood that the events are clustered versus dispersed. A ratio closer to 1 will indicate that the events are more likely dispersed; a ratio closer to 0 will indicate clustering that is statistically unlikely. Using the nearest neighbor test, we quantify the likelihood that certain days are more likely to have event clustering (Gorr and Kurland, 2011). Table 7.2 depicts the results for the daily temporal intervals.

The lower values for Friday and Saturday indicate a slight increase in the likelihood of clustering for these temporal blocks when compared with the other days of the week. Table 7.3 depicts some slightly more distinct patterns in the hourly intervals.

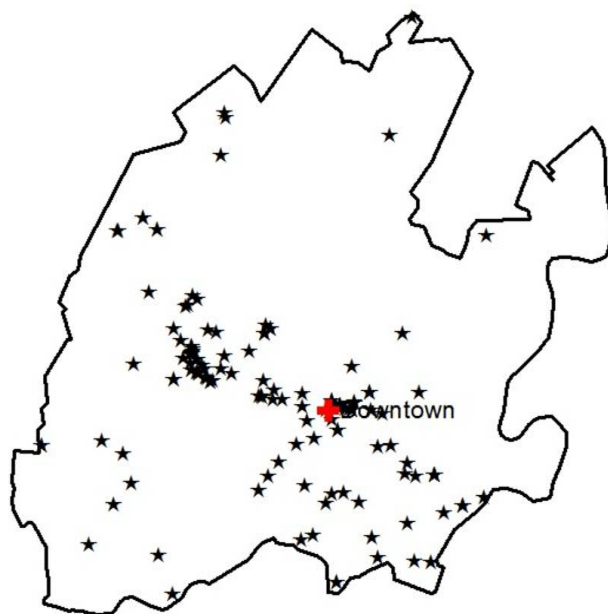
A visual comparison of the events supports the results from the nearest neighbor test by comparing the spatial distribution of events occurring at 0100 hours with events occurring at 0600 hours.

**TABLE 7.2: Average Nearest Neighbor Index Results – Daily**

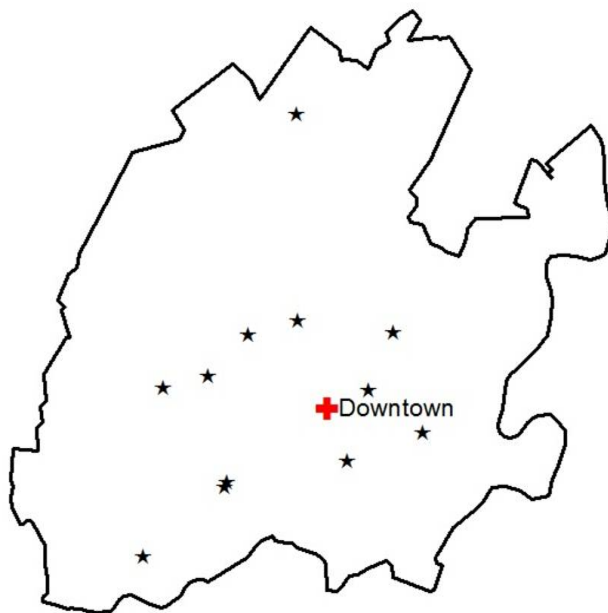
<b>Day of the Week</b>	<b>Observed Mean Distance / Expected Mean Distance</b>
Sunday	.50
Monday	.58
Tuesday	.50
Wednesday	.53
Thursday	.56
Friday	.47
Saturday	.46

**TABLE 7.3: Average Nearest Neighbor Index Results – Hourly**

<b>Hour</b>	<b>Obs. Mean Distance / Exp. Mean Distance</b>	<b>Hour</b>	<b>Obs. Mean Distance / Exp. Mean Distance</b>
0000	.55	1200	.65
0100	.50	1300	.58
0200	.66	1400	.68
0300	.94	1500	.71
0400	.93	1600	.69
0500	.97	1700	.62
0600	.98	1800	.55
0700	.88	1900	.73
0800	.81	2000	.64
0900	.64	2100	.67
1000	.72	2200	.69
1100	.64	2300	.66



(a) 0100 hours

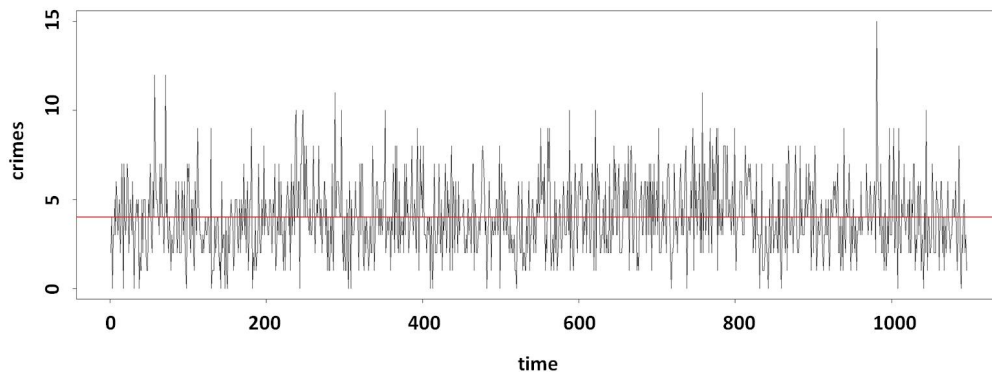


(b) 0600 hours

FIG. 7.9: Comparison of Hourly Interval Dispersion

Figure 7.9(a) depicts a distinct clustering of events around downtown and the vicinity of the “Corner” at UVA during the specified temporal interval. After combining the results of the autocorrelation plots with the nearest neighbor analysis, we have high confidence that temporal patterns might exist within the CSS process for assaults in the study region. However, experimentation has shown that the greatest understanding of the CSS process can be found in the model development process.

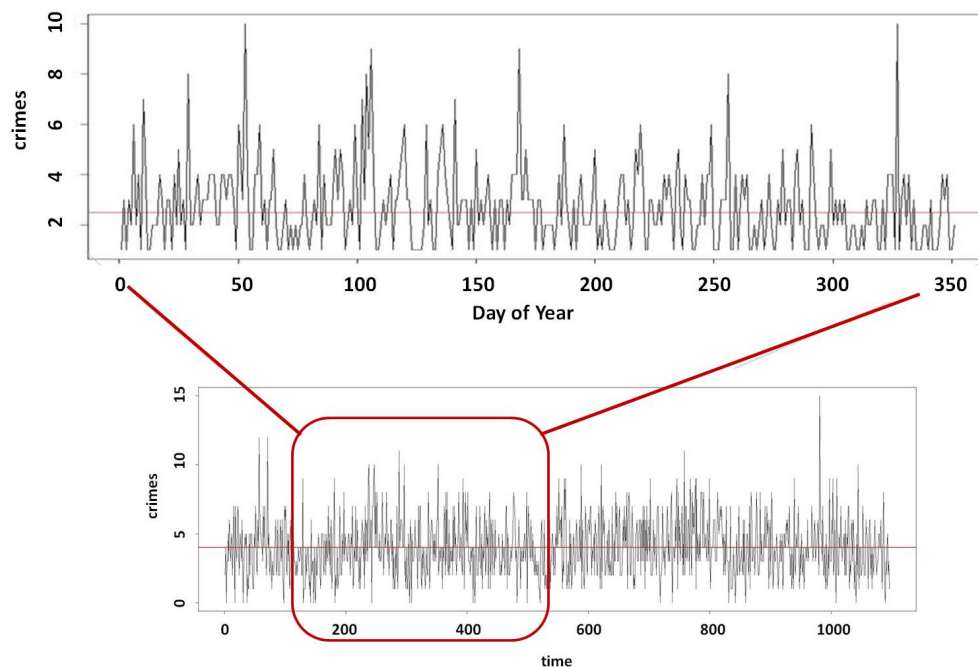
Before we move to the model development process, however, we examine one more component of spatial and temporal analysis. Chapter 5 reviewed the impact of spatial-temporal pulse events on the CSS process. We now return to the complete temporal horizon of event data to identify periods of correlation between potential pulse events and changes to the spatial-temporal patterns of the assault data. Starting with an overview of the time series for assaults, we determine which periods offer significant departure from the mean intensity level. Figure 7.10 reveals a significant departure late in the study horizon.



**FIG. 7.10: Time Series Plot of All Assaults**

Using the cross-correlation function (CCF), we compare a time series of our potential pulse events with our time series of assaults. Unfortunately, the initial results were inclusive. The preliminary model that included the pulse events, mentioned in Chapter 5, failed to provide a significant improvement in fit for our prediction from the base feature space model. Figure 7.11 depicts the lack of definitive impact of

each of the selected special events. In the zoomed window, the arrows depict the a-priori events, such as holidays and school closings, that we believed might cause an increase in criminal activity.



**FIG. 7.11: Time Series Plot with Zoomed Window for Possible Pulse Events**

This initial disappointment is not entirely unexpected. As discussed in Chapter 2, spatial-temporal analysis requires focused efforts on periods of temporal transition and local knowledge of the environment (Kerchner, 2000; Bernasco and Block, 2009). A more appropriate methodology for including the temporal features reduces the scope of the temporal horizon to those intervals with the greatest variance in crime rates. Reducing the temporal horizon to a smaller scale—such as a month before and after large spikes in crime rates—makes it easier to examine the impact of the potential pulse events. More importantly, we also include additional data from the Charlottesville police department to take advantage of its local knowledge of the temporal environment. Using this local knowledge of events, we develop another



**TABLE 7.4: Potential Spatial-Temporal Pulse Events – Charlottesville Assaults**

Temporal Event List	Description	Spatial Component
Football	UVA home football games	UVA stadium
Basketball	UVA home basketball games	John Paul Jones Arena
City – School	City school days (spring break, prom, last day, etc)	City schools
UVA – School	UVA school days (spring break, first day, holidays)	UVA central
Special Events	Foxfield Races, large-venue concerts, etc.	Various

series of potential spatial-temporal pulse events.

Starting with the overview of the time series for assaults seen in Figure 7.10, we determine which periods offer significant departure from the mean intensity level. Again, we use the CCF to measure the correlation between our incident time series and our potential pulse event time series, which in this case includes events from Table 7.4. Figure 7.12 depicts the visual result for the CCF exploration of assaults with the potential pulse effect time series.

On the basis of the distinct spike in the CCF for the Foxfield pulse events, we select the Foxfield Races for an extended analysis. Held every year in the spring and the fall, the Foxfield Races occur on the outskirts of Charlottesville, however, each race weekend brings an increase in visitors to the Charlottesville area and an increase in opportunity for criminal activity. Across the temporal horizon of this event dataset, we have three fall Foxfield Races and three spring Foxfield Races. In Section 7.4, we conduct a slightly more detailed CCF analysis, select training and test data, and modify the hierarchical feature-space model to account for the how the temporal proximity to the Foxfield Races can impact on the CSS process.

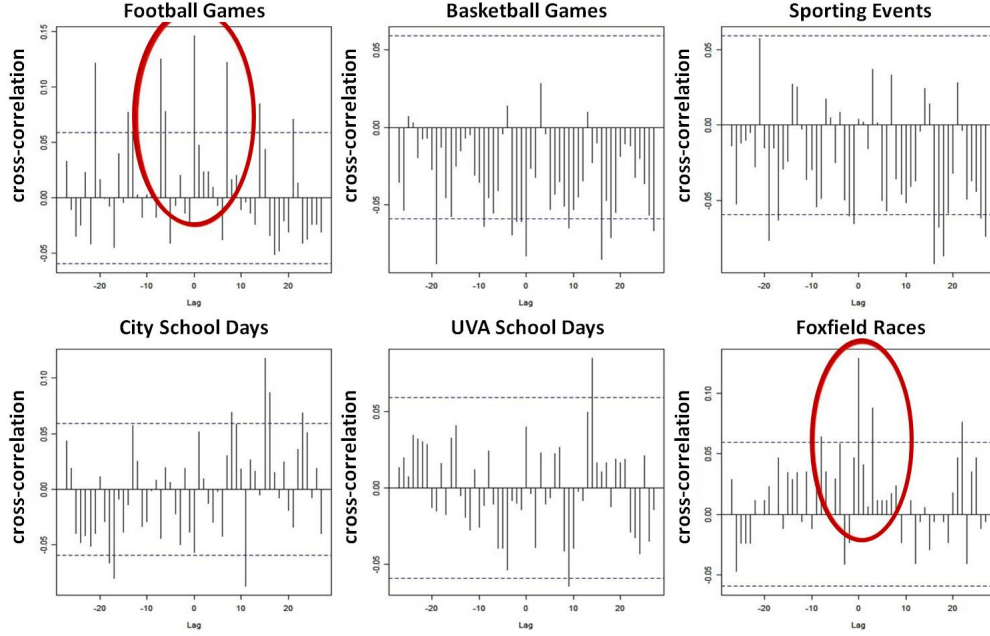


FIG. 7.12: Cross Correlation Results – Assaults

## 7.3 Seasonality and Patterns

### 7.3.1 Temporal Conditioning

In section 7.2, we used visual-analysis and the ACF and PACF plots to identify temporal intervals with focused patterns. Again, we are looking primarily for the intersection between focused temporal patterns with spatial clustering as seen in Figure 3.2. Since the ACF and PACF plots depict such a strong weekly seasonality and the average nearest neighbor analysis showed a slight increase in clustering probability for the weekend, we first refine Equation 4.2 to model the CSS process at a daily temporal interval:

$$\begin{aligned}
\Pr(Y_{s,j} = 1|\mathbf{X}) &= \text{logit}^{-1}(\alpha_j + \beta_{j[k]}\mathbf{X}_s), \\
&\text{for } s = 1, \dots, S \\
&\text{for } k = 1, \dots, K \\
\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} &\sim N \left( \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix} \right), \text{ for } j = 1, \dots, 7
\end{aligned} \tag{7.1}$$

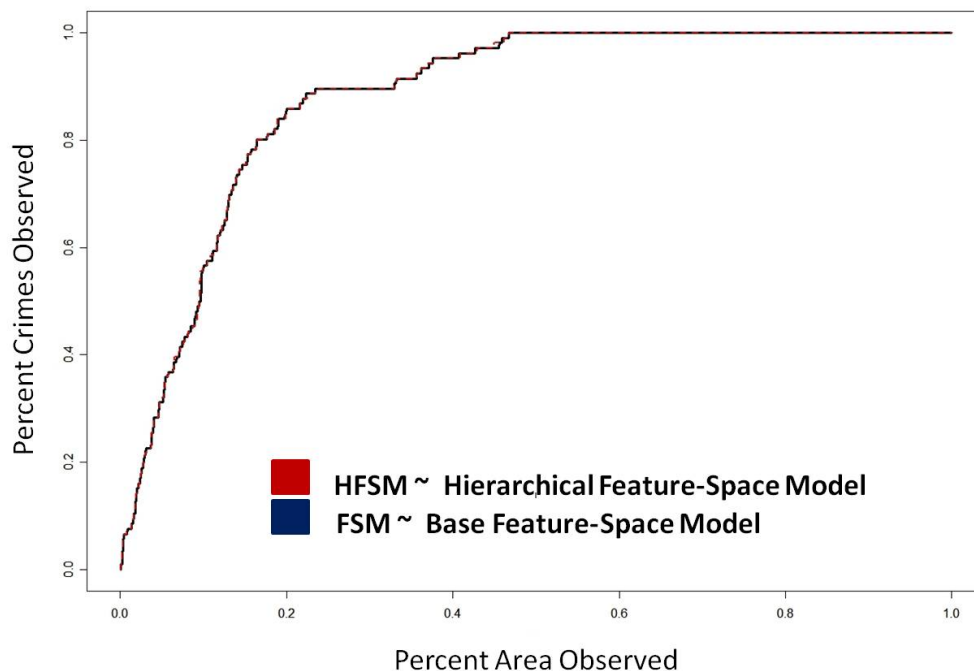
The varying slope and varying intercept model shown in Equation 7.1 continues to include the spatial feature-space components of Liu and Brown (2003) and uses the hierarchical structure from Gelman and Hill (2007) to model the impact of temporal patterns on the CSS process. Equation 7.1 uses day-of-week intervals, but we can easily modify the hierarchical component to represent different temporal intervals on the basis of the seasonality patterns identified in the two-stage analysis.

### 7.3.2 Significance Testing

Using the Charlottesville crime data introduced in Section 7.1, we trained the model and completed feature selection using the assault data aggregated at the daily level. We used 365 days of training data and the subsequent 30 days for testing. The base feature-space model is built against a dataset containing only the incidents specific to the examined temporal interval. The test set also only contains events occurring during the specific temporal interval. The hierarchical model is built against the entire training set and uses the temporal interval specific coefficients to predict against the same test set for the base model.

Unfortunately, the daily interval HFSM does not provide any significant im-

provement over the base feature-space model for predicting across the test set. Using the surveillance plot in Figure 7.13, in fact, we fail to identify any improvement between the HFSM and the base feature-space model without temporal information.



**FIG. 7.13: Surveillance Plot – Assaults – Daily Interval**

Although Figure 7.13 does not show any significant improvement of the HFSM over the base feature-space model when compared against the test set, the point of modeling in this case is not just for predictive power but also for understanding how the criminal's preferences shift regarding space and time. The estimated coefficients indicate a slight impact on some of the selected culture variables introduced in Section 7.1.

A review of Table 7.5 indicates that the predictive surface should change on the basis of the varying coefficients for the culture variables. Figure 7.14 shows only a slight variation between the predictive surfaces for Monday and Saturday. More importantly, the HFSM identifies a change in the CSS process regarding areas that

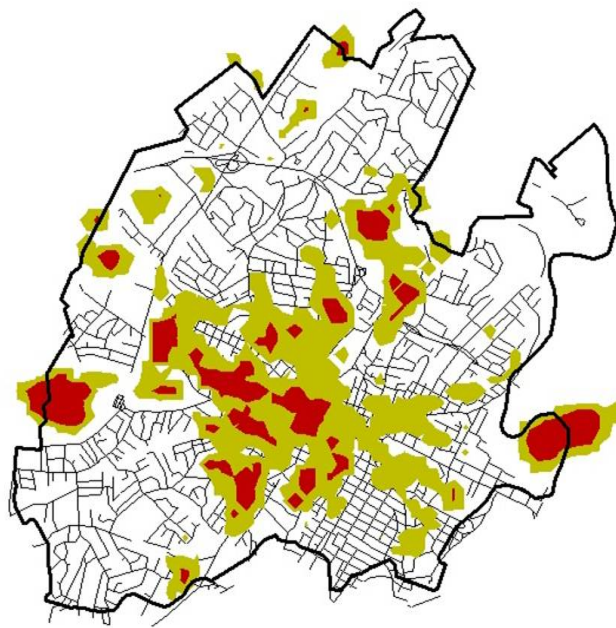
**TABLE 7.5: Varying Coefficients for HFSM – Daily Interval**

<b>Inter- val</b>	<b>Percent Black</b>	<b>Percent Asian</b>	<b>Percent Rental</b>	<b>Percent Vacant</b>
<b>SUN</b>	-0.32052	-17.4269	-2.23923	-0.52675
<b>MON</b>	-0.31455	-17.8428	-2.28203	-1.1237
<b>TUE</b>	-0.29848	-17.7634	-2.25905	-0.87659
<b>WED</b>	-0.35681	-17.4234	-2.26528	-0.7927
<b>THU</b>	-0.34302	-17.6954	-2.27667	-1.02558
<b>FRI</b>	-0.27004	-17.4155	-2.24681	-0.44902
<b>SAT</b>	-0.23102	-17.5743	-2.21917	0.038362

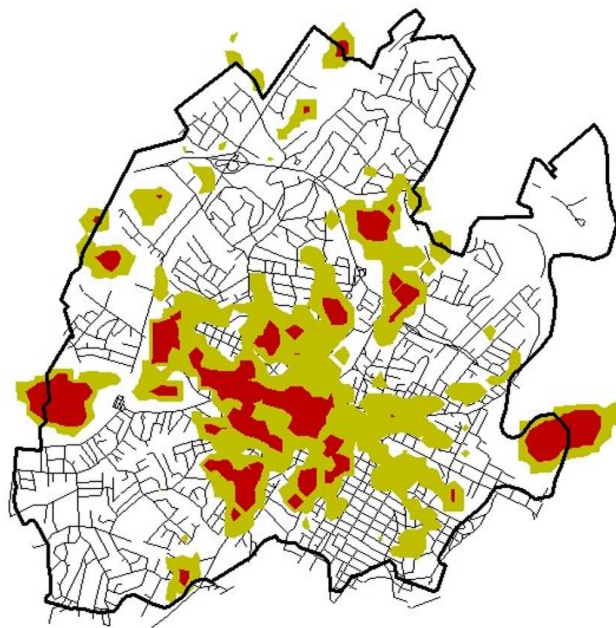
have high rental vacancies. During the week, areas with high vacancies serve as attractors to criminal activity but on the weekend, the CSS process shifts to favor areas with a lower percentage of vacant rentals.

Although the changes are not dramatic, the changes in intensity across the study region reflect opportunities for improving manning strategies or surveillance resource allocation. For a finer resource allocation strategy, however, we need to examine the temporal effects at a more detailed resolution.

To do so, we can return to Equation 7.1 and modify the hierarchical component to reflect conditioning upon an hourly interval. This new model is more computationally intensive since the hierarchical component includes seventeen additional levels when compared to the daily model. And while the hourly HFSM might not provide significant predictive performance across the test set, we do gain further insight into the impact of temporal patterns in the CSS process. Features with negative value coefficients can be seen as criminal attractors (areas that make crime more attractive for criminals). Features with positive coefficients are spatial areas that are less desirable to a criminal. Changes in a coefficient value across temporal intervals indicates that certain areas become more, or less, desirable to a criminal at certain times. Table 7.6 shows a distinct impact on four culture variables and three spatial variables when conditioned on the hour of the day.



(a) Monday



(b) Saturday

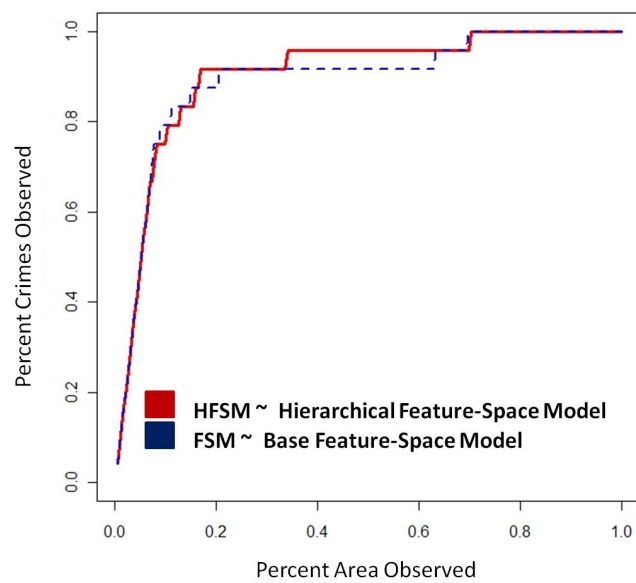
FIG. 7.14: Comparison of Predictive Surfaces from HFSM – Daily Intervals

TABLE 7.6: Varying Coefficients for HFSM – Hourly Interval

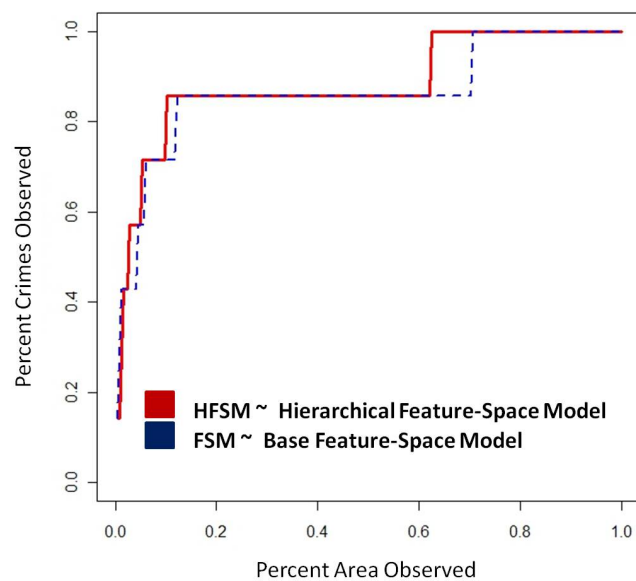
Hour Interval	Dist2Corner	Dist2Downtown	Dist2Schools	PerBlack	PerAsian	PerRental	PerVacant
0000	-0.0004166	-0.0001189	-0.0002802	-0.32555	-17.45900	-1.65150	-1.38723
0100	-0.0004264	-0.0001378	-0.0002873	-0.43318	-17.86900	-1.66430	-0.75545
0200	-0.0004491	-0.0001466	-0.0002925	-0.43419	-17.57200	-1.69120	-1.62384
0300	-0.0004453	-0.0001658	-0.0002906	-0.51150	-17.85200	-1.73630	-2.23316
0400	-0.0004589	-0.0001729	-0.0002960	-0.50322	-18.02200	-1.74930	-2.29617
0500	-0.0004697	-0.0001867	-0.0002986	-0.49087	-18.20200	-1.77170	-2.64802
0600	-0.0004704	-0.0001921	-0.0003042	-0.49072	-18.34500	-1.77930	-2.67733
0700	-0.0004613	-0.0001819	-0.0003003	-0.49896	-18.19400	-1.77400	-2.32532
0800	-0.0004455	-0.0001627	-0.0002944	-0.39597	-18.08400	-1.74420	-2.31418
0900	-0.0004311	-0.0001461	-0.0002898	-0.45027	-18.17900	-1.75400	-2.60981
1000	-0.0004371	-0.0001563	-0.0002885	-0.45414	-18.00400	-1.73720	-2.30912
1100	-0.0004370	-0.0001576	-0.0002918	-0.45662	-17.90300	-1.74010	-2.50458
1200	-0.0004093	-0.0001256	-0.0002884	-0.42780	-17.82000	-1.72440	-2.30454
1300	-0.0004395	-0.0001470	-0.0002943	-0.40488	-18.06600	-1.73680	-2.13659
1400	-0.0004369	-0.0001495	-0.0002896	-0.42414	-17.83400	-1.73310	-2.23767
1500	-0.0004140	-0.0001213	-0.0002863	-0.37253	-17.96700	-1.70980	-1.81440
1600	-0.0004226	-0.0001476	-0.0002821	-0.33225	-17.88900	-1.70250	-1.65103
1700	-0.0003974	-0.0001328	-0.0002834	-0.33053	-18.06200	-1.71680	-1.53720
1800	-0.0004077	-0.0001319	-0.0002814	-0.31609	-17.75300	-1.69310	-1.53554
1900	-0.0004192	-0.0001308	-0.0002848	-0.39350	-17.55400	-1.70390	-1.68235
2000	-0.0004030	-0.0001314	-0.0002763	-0.31439	-18.00800	-1.70660	-1.39760
2100	-0.0003960	-0.0001306	-0.0002813	-0.33836	-17.99500	-1.70650	-1.33747
2200	-0.0004130	-0.0001336	-0.0002792	-0.35285	-17.79300	-1.69610	-1.31970
2300	-0.0004203	-0.0001315	-0.0002819	-0.38367	-17.74400	-1.68360	-1.24579

As with the simulated data, we use the surveillance plots and the area under the curve as a measure of the model’s predictive performance for a specific time period. Overall, we do not see the same dramatic improvement using the sample data as was seen with the simulation data. The largest increases in predictive performance are seen during the temporal intervals (0100, 1300, and 1800 hours) when the nearest neighbor ratio indicated a greater likelihood of clustering. Table 7.7 highlights these temporal intervals.

Although the surveillance plots in Figure 7.15 show slight improvement of the HFSM over the base feature-space model, the point of modeling in this case is not just for predictive power, but also understanding how the criminal’s preferences shift in space and time. An examination of the estimated coefficients reveals the temporal patterns on some of the selected culture variables used for modeling. Of special note



(a) 0100 hours



(b) 0600 hours

**FIG. 7.15: Surveillance Plots for Comparison of HFSM and FSM During Periods of High and Low Activity**



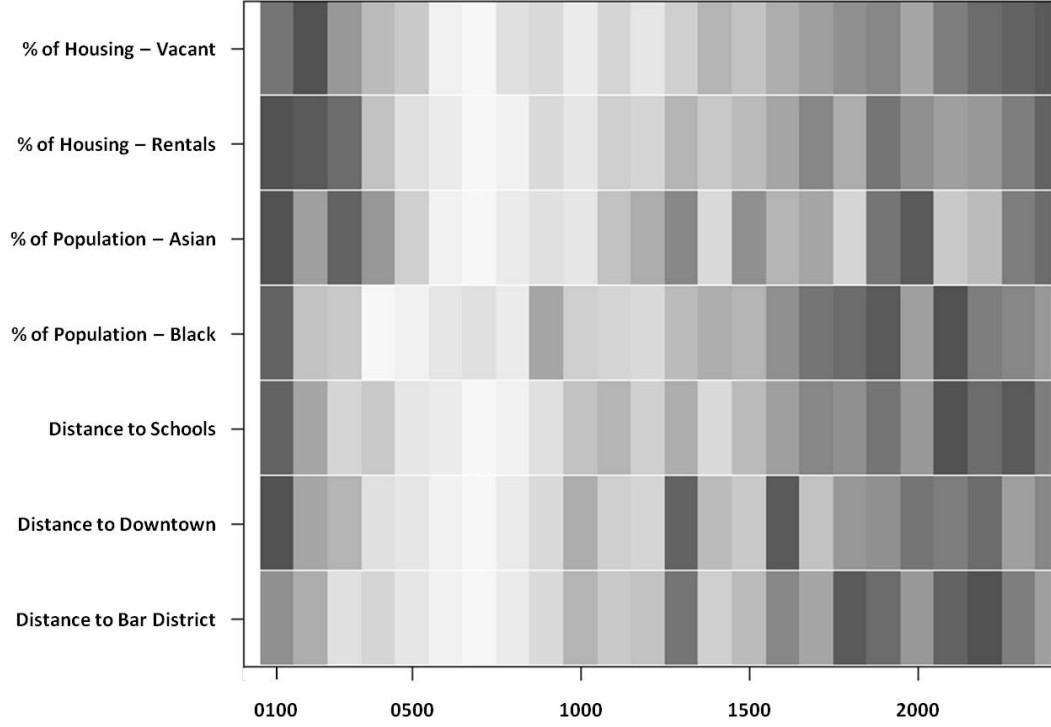
**TABLE 7.7: Area Under the Curve (AUC) Results for HFSM and FSM - Hourly Interval**

Hour	HFSM AUC	FSM AUC	Hour	HFSM AUC	FSM AUC
0000	.9498	.9487	1200	.9647	.9672
0100	<b>.9509</b>	.9379	1300	.9383	.9218
0200	.9375	.9362	1400	.9739	.9616
0300	.9663	.9715	1500	.9439	.9318
0400	.8415	.8374	1600	.9492	.9502
0500	.8931	.8757	1700	.9583	.9591
0600	.9722	.9761	1800	.9661	.9611
0700	.8714	.8721	1900	.9506	.9554
0800	.9853	.9909	2000	.9469	.9512
0900	.9673	.9794	2100	<b>.9275</b>	.9063
1000	.8980	.8854	2200	.9637	.9532
1100	.9335	.9391	2300	.9479	.9407

is the change identified regarding the role each feature plays as a potential attractor or repeller of criminal activity throughout the day. We depict these temporal patterns within the CSS process by using a multivariate time series plot. We normalize the modeled coefficients to compare impact of the features on the CSS across the temporal intervals (Peng, 2008). An example of a distinct temporal pattern is seen in Figure 7.16 showing the varying impact of four culture features and three spatial features across each hour of the day.

Darker areas indicate that the spatial feature is more of an attractor in the CSS while lighter areas indicate a lessening attraction. In most cases, this temporal shift matches the daily activity rhythms present in most cities, especially with the slight increase in the impact of distance to downtown bracketing the lunch hour.

In Chapter 4, we introduced the idea that the greatest value in predictive performance for the HFSM would be in examining multiple temporal intervals across a complex hierarchical model. We can extend our two-stage analysis from Section 7.2 to examine the variation across months, days-of-week, and hours-of-day as a



**FIG. 7.16: Varying Coefficients for HFSM – Hourly Interval**

suitable aggregated starting point. For ease of computation, we further aggregate the time-of-day data into a shift variable that is based on the hour-of-day occurrence. In Equation 7.2 below, we extend our feature-space model to a more complex, non-nested, hierarchical model that examines the varied spatial relationships within the monthly ( $m$ ), daily ( $d$ ), and shift ( $s$ ) temporal intervals:

$$\Pr(Y_{m[i],d[i],s[i]} = 1|X) = \text{logit}^{-1}(\alpha_{m[i],d[i],s[i]} + \beta_{m[i],d[i],s[i]}X_i), \text{ for } i = 1, \dots, n. \quad (7.2)$$

This data-level model can be decomposed to represent the intercepts and slopes for the temporal blocks minus the interaction terms

$$\begin{pmatrix} \alpha_{m,d,s} \\ \beta_{m,d,s} \end{pmatrix} = \begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix} + \begin{pmatrix} \gamma_{0m}^{month} \\ \gamma_{1m}^{month} \end{pmatrix} + \begin{pmatrix} \gamma_{0d}^{day} \\ \gamma_{1d}^{day} \end{pmatrix} + \begin{pmatrix} \gamma_{0s}^{shift} \\ \gamma_{1s}^{shift} \end{pmatrix} \quad (7.3)$$

along with the variations across the temporal levels:

$$\begin{aligned}
 \begin{pmatrix} \gamma_{0m}^{month} \\ \gamma_{1m}^{month} \end{pmatrix} &\sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \left( \Sigma^{month} \right), \right) \text{ for } m = 1, \dots, 12 \\
 \begin{pmatrix} \gamma_{0d}^{day} \\ \gamma_{1d}^{day} \end{pmatrix} &\sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \left( \Sigma^{day} \right), \right) \text{ for } d = 1, \dots, 7 \\
 \begin{pmatrix} \gamma_{0s}^{shift} \\ \gamma_{1s}^{shift} \end{pmatrix} &\sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \left( \Sigma^{shift} \right), \right) \text{ for } s = 1, \dots, 3.
 \end{aligned} \tag{7.4}$$

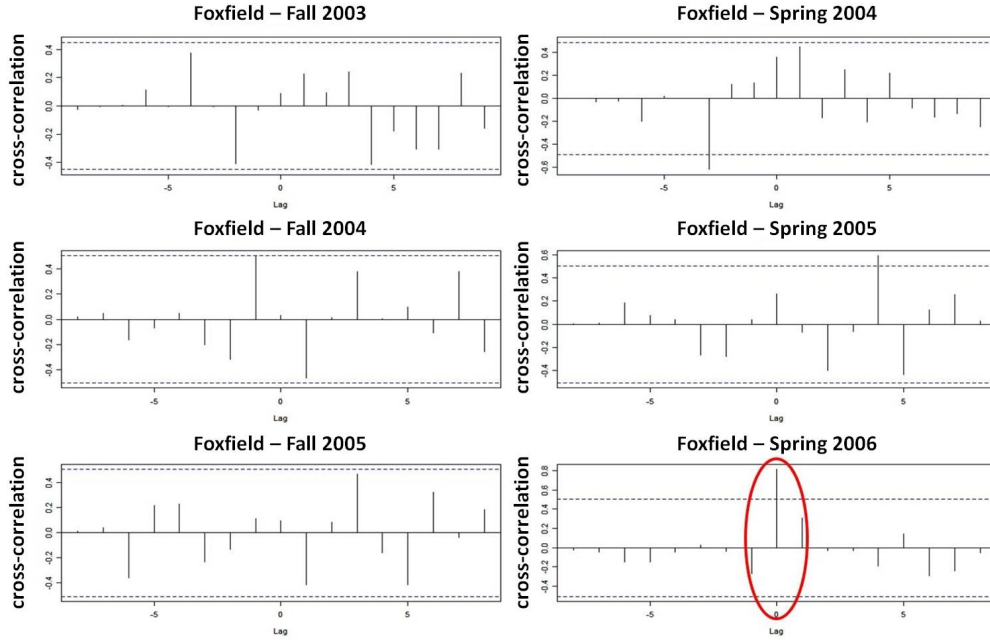
We compare the performance of our new HFSM to the base FSM without temporal features by using the surveillance plot. At the 20% of area observed surveillance point, the HFSM offers an approximately 5% increase in predictive performance for events across the modeled intervals.

## 7.4 Indicators and Interactions

### 7.4.1 Temporal Pulse Events

In Section 7.2, we used the CCF plots to select the Foxfield Races for our extended analysis. The Foxfield Races are horse racing events that occur every spring and fall on the outskirts of Charlottesville. More specifically, the Foxfield Races are large public gatherings that bring large crowds into the Charlottesville area. An examination of a reduced temporal window around each of the six Foxfield events in our dataset shows a distinct correlation between the spring Foxfield Races and the changes in the assault crime rate.

As seen in Figure 7.12, the fall Foxfield Races do not seem to be associated with an increase in the number of assaults. Given that the fall Foxfield event is widely viewed and marketed as a family event, this lack of correlation between the fall



**FIG. 7.17: Cross-Correlation Results – Foxfield Races and Assaults**

Foxfield events and shifts in the temporal rates of assaults makes sense. However, the large CCF response for the spring Foxfield event warrants extended analysis.

### 7.4.2 Significance Testing

We use the temporal proximity to the Foxfield Race as the hierarchical component of the HFSM. Since the Foxfield Race is held on a one-day event held on Saturday, we use days as our temporal resolution. We change the notation from Equation 5.3 slightly to reflect this new range of temporal distance for a varying-slope and varying-intercept model. In this case, the hierarchical component represents the temporal distance (i.e., the number of days) to the Foxfield event. We are now modeling how the CSS process changes on the basis of the temporal proximity to

the pulse event.

$$\begin{aligned} \Pr(Y_{s,j} = 1|X) &= \text{logit}^{-1}(\alpha_j + \beta_{j[k]}\mathbf{X}_s), \\ \begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} &\sim N\left(\begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix}\right), \\ &\text{for } s = 1, \dots, S \\ &\text{for } k = 1, \dots, K \\ &\text{for } j = 1, \dots, 8 \end{aligned} \tag{7.5}$$

An examination of the coefficients shows a slight impact on four of the feature-space variables used in the HFSM. While each of the coefficient values change slightly on the basis of the decreasing temporal proximity to the Foxfield event, only the feature-space variable for percent of vacant houses shows a complete change from being an attractor to a repeller of criminal activity. For this case, we use the feature variable, *Distance to the Corner*, to account for a spatial shift in preferences for the CSS process since the majority of the post Foxfield activity moves from the event location outside of Charlottesville city limits to the primary entertainment district vicinity of the UVA grounds. Although the changes in the feature-space

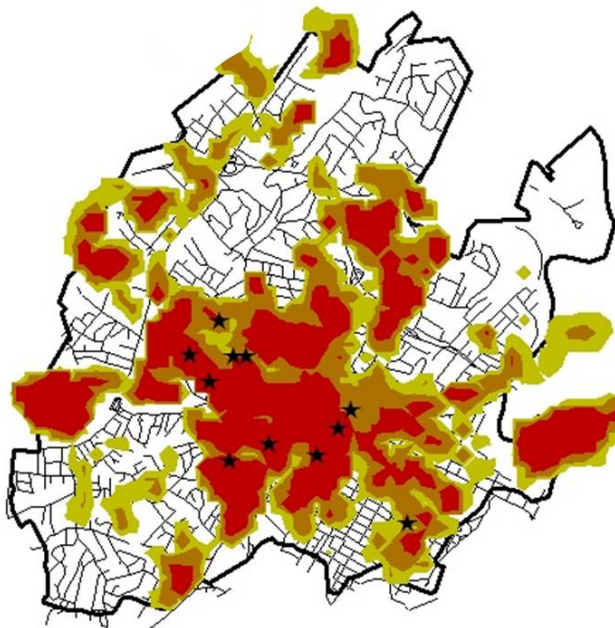
**TABLE 7.8: Varying Coefficients for Foxfield Pulse Event – HFSM**

<b>Days to Foxfield</b>	<b>Distance to Corner</b>	<b>Distance to Downtown</b>	<b>Percent Rental</b>	<b>Percent Vacant</b>
<b>7</b>	-0.001044	-0.001084	1.970	-0.01721
<b>6</b>	-0.001044	-0.001081	1.969	-0.08524
<b>5</b>	-0.001045	-0.001078	1.973	-0.06351
<b>4</b>	-0.001047	-0.001085	1.969	-0.07564
<b>3</b>	-0.001038	-0.001080	1.971	-0.04964
<b>2</b>	-0.001044	-0.001080	1.972	-0.01072
<b>1</b>	-0.001046	-0.001081	1.969	-0.09368
<b>0</b>	-0.001037	-0.001075	1.983	0.258530

coefficients are not dramatic, the introduction of the hierarchial framework to the

feature-space model offers significant improvement for law enforcement personnel focused on patrol assignments and other resource allocation problems. Figure 7.18 depicts the predictive threat surface built using the base FSM without any temporal information. Without the temporal information, the base FSM produces a threat surface for the day of the Foxfield event that is exactly the same as the threat surface for one week prior to Foxfield. Figure 7.19, however, depicts the predictive threat built using the HFSM for the day of the Foxfield event.

Adding the temporal proximity to the Foxfield event as the hierarchical component, we can now produce a series of daily predictive threat surfaces that adjust on the decreasing temporal distance to the approaching event. Refinement of the predictive surface allows law enforcement personnel to adjust patrols and manning levels since both can be focused on a smaller spatial area for that specific temporal window surrounding the special event.



**FIG. 7.18: Base FSM for Foxfield Races with Incidents for Race Day**

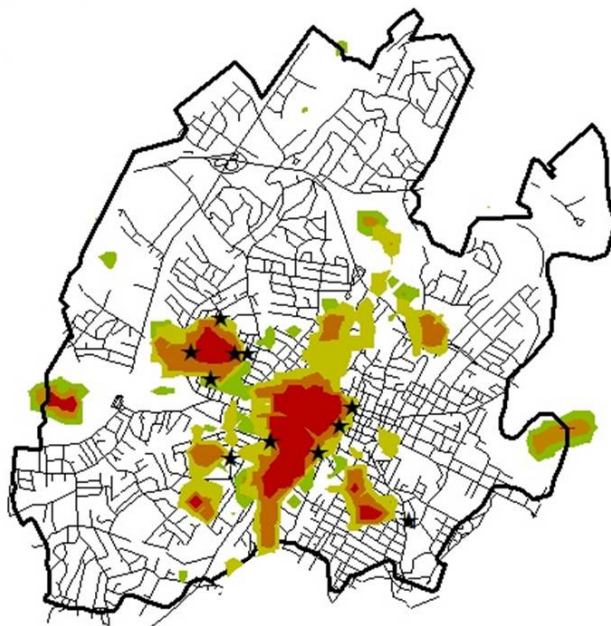


FIG. 7.19: HFSM for Foxfield Races with Incidents for Race Day

## 7.5 Neighbors and Recent Past

One motivation for analyzing criminal incidents is to improve resource allocation of police forces to either reduce the likelihood of crime at certain places and times or to mitigate the effects from a crime that has already occurred. Although the hierarchical modeling techniques discussed so far offer improvements in the prediction of crimes with temporal patterns, we have not yet considered time as a continuous element. Examining the trends of criminal incidents in both space and time concurrently may provide insight into the underlying risks for certain areas and help first responders pre-position assets in order to respond quickly during expected peak times. The Bayesian hierarchical feature-space model (HFSM) and the expanded feature-space model with spatial-temporal indicator functions (EFSM) provide two methods for addressing the continuous nature of time in criminal site-selection (CSS) problems.

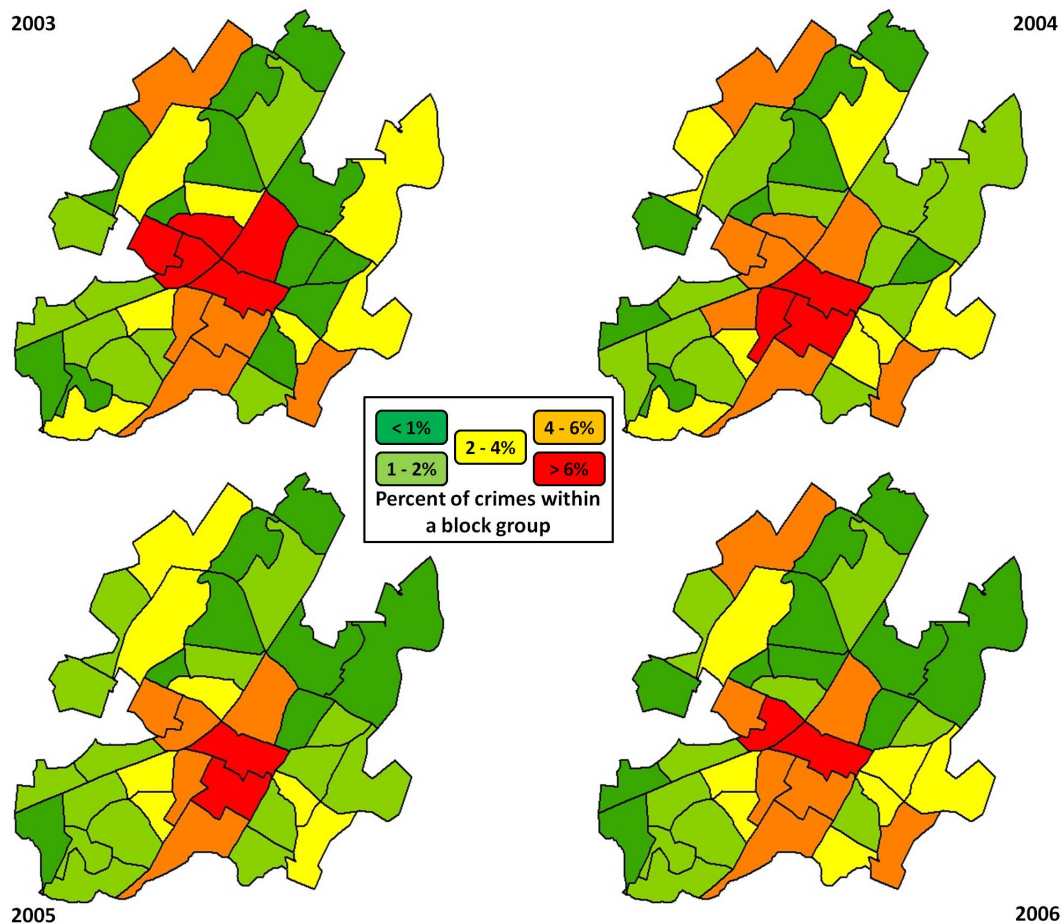
### 7.5.1 Data Aggregation and Preparation

Using the data from Section 7.1, we sample the complete dataset to develop a subset that contains a time horizon spanning three years with over 2,000 incidents. We restrict the crime types analyzed for this methodology to assaults, both simple and aggravated. We drape an irregular lattice over the study area and aggregate the criminal incidents at the daily level. Although the aggregation introduces some level of discreteness, we treat the temporal intervals as continuous points along the temporal horizon and not as a temporal bin for applying the conditioning method. The irregular lattice structure is based on the 37 US Census block-groups for the city. Using the block group lattice structure facilitates inclusion of demographic information at the block-group level that serve as features considered by criminals in their site-selection process. Figure 7.20 depicts the study region draped with the irregular lattice and showing spatial-temporal patterns of assaults over four distinct temporal intervals. The analysis that follows uses a second order neighbor structure over the irregular lattice depicted in Figure 7.20.

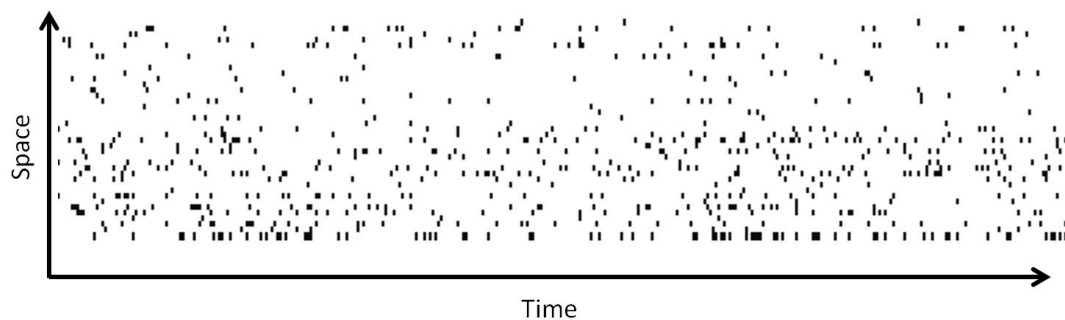
We set  $Y_{i,t} = 1$  if a criminal assault occurs within the specified block group  $i = 1, \dots, 37$  during one of the days  $t = 1, \dots, 1095$  of the study horizon. The block-group and daily aggregation results in a  $37 \times 365$  matrix for a total of 40,515 observations in space-time. Figure 7.21 depicts a one year snapshot of criminal events across the entire spatial region.

We develop an undirected graph of the study region to meet the conditions for using a MRF construct within the parameter model of the Bayesian hierarchical framework. The graph in Figure 7.22 does not represent the spatial layout of the study region but rather shows the 37 block-group nodes and the relationships built



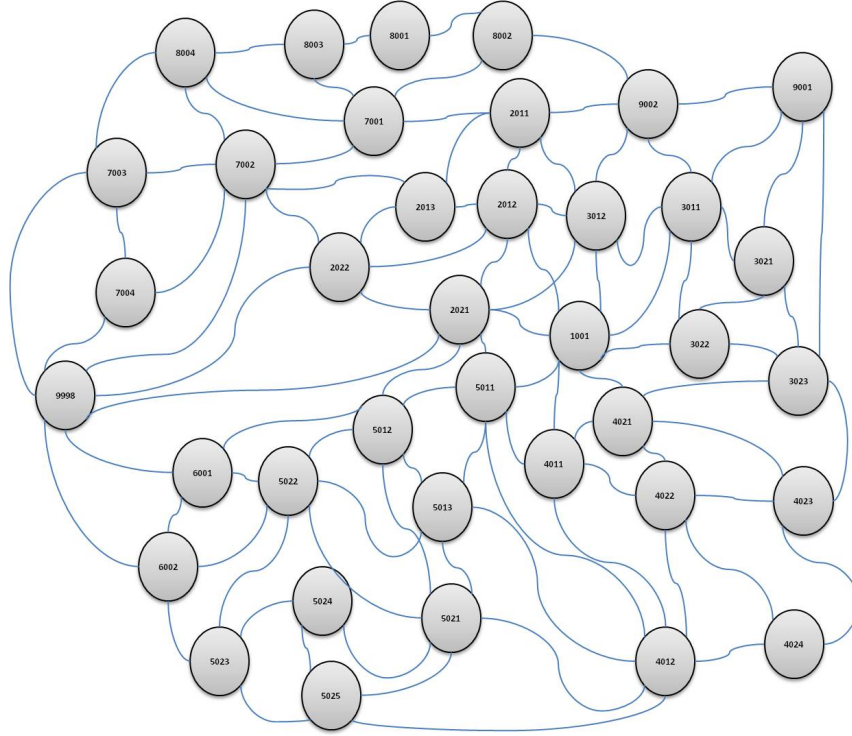


**FIG. 7.20: Evolution of Spatial Patterns Over a Continuous Temporal Horizon.** Since we identify changes in the map over time, we hypothesize that we have spatial and temporal effects within the criminal site selection process.



**FIG. 7.21: Space-time Lattice for Study Domain.** The arrangement of spatial regions along the  $y$  axis might falsely identify spatial clusters, however, the temporal horizon along the  $x$  axis does allow for visual identification of temporal clusters for assaults within the study region.

upon shared borders for a total of 186 neighbors. Returning to the space-time model construct depicted in Figure 2.10, we use our undirected graph to account for the impact of neighborhood activity on the CSS process.



**FIG. 7.22: Undirected Labeled Graph for Study Region.** Each node represents one of the thirty-seven block-groups within the city of Charlottesville, Virginia. The layout of the graph is not intended to show the spatial layout of the block groups but to rather to show the second-order neighbors. We label each node with the US Census code for the block-group and depict the neighbor connections via the lines.

## 7.5.2 Significance Testing

### Long Term Study Results

For the long term study, the block-group and daily aggregation results in a  $37 \times 365$  matrix for a total of 13,505 observations in space-time. We use a second-order

neighbor model to account for all the criminal activity in the surrounding census blocks. Table 7.9 outlines the specific models examined using both the demographic features and a feature-space prior obtained from the basic feature-space model of Chapter 4. We consider four alternatives to model the random effects using our variable  $\theta_{s,t}$  in the process model of Equation 6.2: 1) a time-varying trend; 2) a Markov random field accounting for the sum of the neighboring effects at a previous time; 3) a time-varying trend with a Markov random field; 4) a time-varying trend with a Markov random field and an interaction term. For every alternative we include a space-time independent noise term. For the first four alternatives, we attempt to account for the criminal site-selection preference by modeling  $\beta$  as seen in Figure 6.2. After model fitting, we evaluate performance using the MSPE discussed in Chapter 6.

**TABLE 7.9: Bayesian Hierarchical Feature-Space Model Development for Long Term Data Study**

Model	Predictors	Time	Deviance	MSPE
Spatial Choice and Trend	35	3049	4632	0.0430
Spatial Choice and MRF	35	2587	4622	0.0429
Spatial Choice and MRF and Trend	36	4694	4625	0.0429
Spatial Choice and MRF and Trend and Interaction	43	19481	4609	0.0428
Feature-Space Prior and Trend	31	2273	4785	0.0435
Feature-Space Prior and MRF	30	2314	4632	0.0430
Feature-Space Prior and MRF and Trend	38	7219	4614	0.0429
Feature-Space Prior and MRF and Trend and Interaction	40	9515	4619	0.0429

Although the predictive performance of the BHFSM is not significantly bet-

ter than the base feature-space model, we were expecting to see significant lift in the parameter estimation related to identifying criminal preferences for certain spatial features. In fact, even with all four models converging, the only feature-space variable with significantly better estimation was the preference for areas with high percent vacancy. However, (Lawson, 2009) shows that the combination of spatially-referenced explanatory variables within a Markov random field construct often yields poor estimates of the regression coefficients and produces computational challenges related to multi-collinearity. Both of our approaches to reduce the impact of correlation created additional challenges. First, removing the features that are spatially dependent limits our insight into the criminal site selection process for identifying feature-space preferences. Second, introducing new variables that have a stationary spatial attribute but are non-stationary temporally limits our ability to identify how the criminal's feature-space preferences evolve over time. Overall, the Bayesian approach offers promise to reducing uncertainty in the predictive surfaces. However, the computational time required for sampling from the posterior distribution for Bayesian inference for criminal site-selection problems remains a major drawback (Withers, 2002; Zhu et al., 2008).

We discuss an alternative approach in the conclusion that offers computational advantages while remaining sufficiently accurate for prediction. In the next section, we scale down the horizon of the study period as an additional step in examining the BHFSM.

### **Short Term Study Results**

Although applying the Bayesian framework to the long term study data did not result in significant gains in predictive performance, the initial disappointment was not entirely unexpected. Previous research shows that spatial-temporal analysis focused on criminal site-selection requires focused efforts on periods of temporal

transition and local knowledge of the environment (Kerchner, 2000; Bernasco and Block, 2009). A more appropriate methodology for including temporal information into the BHFSM reduces the scope of the temporal horizon to those intervals with the greatest variance in crime rates. Research has also shown that spatial regions experience great variance in crime rates for certain locations depending on the temporal proximity to special events (Cohen et al., 2007). Reducing the temporal horizon to a smaller scale — such as a thirty day window before and after large spikes in crime rates — makes it easier to examine the impact of these special events on the criminal site-selection process. More importantly, including additional data from local law enforcement personnel takes advantage of their local knowledge of the temporal environment (Cressie and Wikle, 2011).

**TABLE 7.10: Bayesian Hierarchical Feature-Space Model Development for Short Term Data Study**

Model	Predictors	Time	Deviance	MSPE	PLR
Feature-Space Model	7	5	423	0.0479	0.46
Expanded Feature-Space Model	9	5	420	0.0453	0.55
Spatial Choice and Trend	19	182	423	0.0479	0.52
Spatial Choice and MRF	19	303	423	0.0479	0.53
Spatial Choice and MRF and Trend	22	332	422	0.0478	0.53
Spatial Choice and MRF and Trend and Interaction	23	900	421	0.0477	0.53
Feature-Space Prior and Trend	8	180	420	0.0477	0.49
Feature-Space Prior and MRF	8	138	420	0.0477	0.47
Feature-Space Prior and MRF and Trend	10	371	419	0.0476	0.48
Feature-Space Prior and MRF and Trend and Interaction	12	1022	417	0.0475	0.50

Returning to the scenario from Section 7.4, we use a thirty-day study horizon surrounding the spring 2005 Foxfield event as our training set and then use the thirty-day horizon surrounding the spring 2006 Foxfield race as our test set. As with the long term study, we consider all four alternatives to model the random effects using our variable  $\theta_{s,t}$  in the process model. Table 7.10 outlines the specific models examined using both the demographic features and a feature-space prior obtained from a generalized linear regression similar to the work of Liu and Brown (2003) and as seen in our visual graph from Figure 6.2. Again, the only feature-space variable with significantly better estimation was the preference for areas with high percent vacancy. After model fitting, we evaluate performance using the PLR discussed in Section 3.7. The PLR plots seen in Figure 7.23 depict the spatial-temporal accuracy of each model’s performance in predicting future events. Although each BHFSM performs better than the base feature-space model, the computational time required for sampling from the posterior distribution for Bayesian inference is still several orders of magnitude greater than the time required for using the base feature-space models. The expanded feature-space model (EFSM) incorporates both temporal and neighborhood information and outperforms all other methods for accurately identifying areas with a high likelihood of crime.

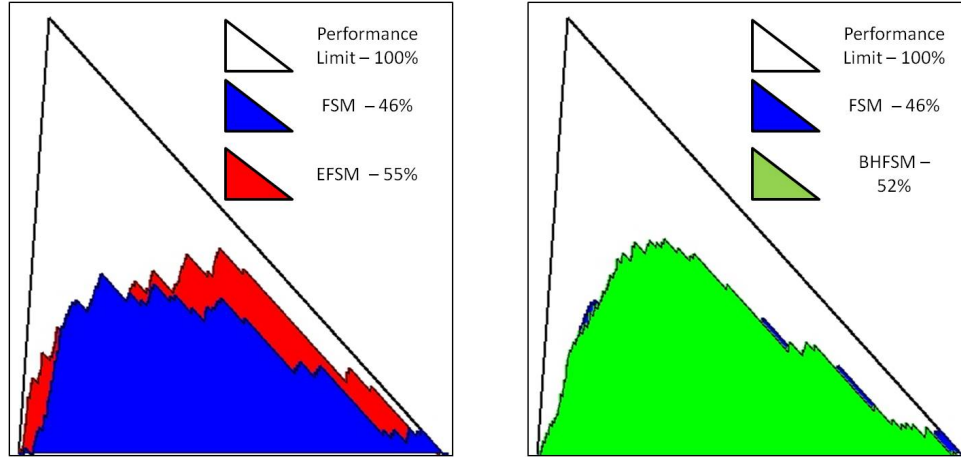


FIG. 7.23: Performance Limit Ratio

## 7.6 Discussion

Incorporating temporal information into the feature-space model improves both predictive inference and parameter estimation in criminal site-selection (CSS) problems. Enhanced understanding of the criminal site-selection process allows law enforcement personnel to adjust resource allocation strategies to better mitigate short term changes in the criminal-site selection process. For temporal intervals with focused patterns matched to spatial clustering, the hierarchical feature-space model (HFSM) provides improved predictive performance across discrete temporal intervals. The HFSM also enhances the understanding of temporal patterns in the CSS process. Although aggregation limits the predictive ability for non-aggregated data, the relative computational simplicity of hierarchical modeling provides the analyst a way to identify and understand space-time patterns in a dataset. Hierarchical modeling can be combined with classification information to examine distinct spatial-temporal patterns of different types of crimes on the basis of severity or community impact.

We observe a significant improvement when we include temporal proximity to identified spatial-temporal pulse events to account for changes from the baseline intensity rate of criminal assaults. Conducting a time series analysis of the event data

allows us to identify potential change periods and possible pulse events. Adding the temporal distance as a hierarchical component to the feature space model improves the accuracy of the predictive surface approximately 16% over the predictive surface without temporal features.

For strategic-level crime monitoring, incorporating the feature-space model into a Bayesian hierarchical framework accounts for both spatial and temporal components of the CSS process across continuous time. Although the overall predictive performance is not significantly improved, by reducing the variance on estimates for a criminal's feature-space preferences, we gain understanding into the temporal variations of the criminal site-selection process. Using spatial and temporal indicator functions to expand the base feature-space model (EFSM) provides a method that is computationally swift and sufficiently accurate. The performance of the EFSM dominates the others when we compare them using either the surveillance plot or the PLR across the space-time domain of interest.



# CHAPTER 8

## Conclusions

*The ideal reasoner . . . would, when he has once been shown a single fact in all its bearing, deduce from it not only all the chain of events which led up to it, but also all the results which would follow from it.*

– Sherlock Holmes in *A Scandal in Bohemia*

This chapter reviews the results and contributions of the research. Additional comments provide a focus for future work including the challenge of visualization and other potential application areas.

## 8.1 Contributions

This research develops three methods, and presents a major application, for considering temporal data as part of modeling and forecasting criminal site-selection (CSS) processes. Overcoming the challenges of spatial-temporal events offers benefits in both the academic and application arenas. Although previous research has shown that temporal patterns exist in criminal activity (Rossmo et al., 2005; Groff, 2007), most early work on CSS problems examined data aggregated across temporal intervals or within small two-week or monthly windows (Brown and Liu, 1999; Liu and Brown, 2003; Brown et al., 2004). To account for temporal patterns in the CSS process, we merge the traditional feature-space model with a hierarchical specification to improve current forecasting capabilities. Conditioning the base feature-space model with a hierarchical component built at a distinct temporal resolution gives law enforcement personnel a method to account for seasonality in the CSS process. The more general application of the hierarchical model includes non-nested temporal blocks at varied resolution in order to account for daily, weekly, or other seasonal patterns.

However, just as the criminal's preferences for certain locations might change depending on the time-of-day or the day-of-week, we also want to examine the impact of the special events on the CSS process. Previous research has shown that special events can impact the rate of crimes within a region or change the spatial distribution of crimes for a limited temporal window surrounding the special event (Cohen et al., 2003, 2007). Using the cross-correlation function (CCF) as a tool to

determine the relationship between two time series, we can identify potential spatial-temporal “pulse events” that affect the CSS process. Incorporating the temporal distance to the pulse event as a hierarchical component for the base feature-space model improves forecasting capabilities for law enforcement personnel during the temporal windows surrounding these special events.

Modifying the base feature-space model with a hierarchical framework provides the analyst with a method to account for temporal changes in the CSS process. Unfortunately, both the hierarchical and traditional feature-space methods fail to account for the impact of spatial and temporal recency on the CSS process. Using an irregular lattice structure of neighbors within a limited spatial-temporal region provides a methodological approach for identifying a criminal’s feature preferences across a continuous temporal horizon. Exploring the hierarchical feature-space model through a Bayesian framework provides a methodology for analyzing event initiations at higher spatial resolutions while reducing uncertainty in the feature preferences. Our more computationally efficient method uses temporal- and neighborhood-indicator functions with the base feature-space model to rapidly examine the CSS process across a continuous temporal horizon with predictive performance that exceeds the base feature-space model. The short-term study horizon analysis identified a significant decrease in the impact of certain variables to the model during non-active periods. Although we are not accounting for temporal changes in the criminals’ preferences, we are able to develop more accurate predictions that assist law enforcement personnel with adjusting staffing and asset allocation to meet the most likely patterns of criminal activity.

The examined methodologies were applied to a practical scenario using criminal data from Charlottesville, Virginia. The spatial-temporal patterns of varied crime types across the study region were modeled using the HFSM, the Bayesian HFSM, and the expanded feature-space model (EFSM). A non-nested structure for

different temporal blocks provides the most detailed understanding of the changes to the hierarchical model across both space and time. Using data aggregated at the census block-group level for a medium temporal resolution, the BHFSM allowed us to model an actor's spatial-temporal preferences within a limited temporal period. Incorporating elements of the feature-space methodology into the Bayesian construct allowed us to blend the benefits gained from understanding multiple covariates within the actor's spatial-temporal decision process with the basic elements of geographic recency and spatial dependence found in hotspot modeling.

The EFSM provides significant improvement in providing a decision maker with information of events within a space-time continuum over traditional modeling. The EFSM provided a computationally efficient and sufficiently accurate alternative to both the BHFSM and the HFSM. Using the PLR as a performance metric, the EFSM provides a 16% improvement in predictive performance over the base feature-space model without temporal information across the long term study horizon. Across a smaller study horizon, the BHFSM does provide improvement in predictive performance, but the increased computational requirements hinder the application of the BHFSM for "just-in-time" modeling. We expect that hierarchical modeling of inherently spatial-temporal events will improve resource allocation strategies for both crime management and crime response.

## 8.2 Future Work

### 8.2.1 Computational Alternatives

Several challenges remain for the further inclusion of temporal features in modeling the criminal's site-selection process. Two of the methodologies examined in this dissertation, temporal conditioning and the Bayesian hierarchical framework, are computationally intensive for the examined application area. The expanded feature-space model offers the performance benefits of improved predictive ability and faster modeling, but other approaches may deliver similar results with even greater gains in computational time. Structural vector autoregressive models (SVARs) show promise for forecasting employment rates given spatially based economic indicators (Rickman et al., 2009). Using an SVAR construct for modeling CSS problems might improve predictive ability if temporal changes in other features affect a criminal's temporal considerations for certain sites. However, the computational requirements for the SVARs, much like the Bayesian hierarchical model, are still rather demanding (Petrakis et al., 2009).

The social sciences offer another approach for reducing the computational demands of CSS modeling. Spatial-temporal designs for environmental research often include panel methods for monitoring and detecting temporal patterns and spatial relationships (Dobbie et al., 2008). We are not designing a method for collecting criminal-event data, but rather examining historical collections of crime data. After building the null-grid dataset, as reviewed in Section 7.1, we have a large dataset that requires inversion of the covariant matrix for the hierarchical feature-space model (HFSM). Using a variation of stratified sampling (Dobbie et al., 2008; Gilbert, 1987) on the resulting dataset might reduce the computational time while retaining comparable predictive performance.

## 8.2.2 Vehicle Crash Modeling

A current study from the American Automobile Association estimates the annual cost of traffic accidents to the American public as approximately \$160 billion. In addition, the study reports the national mortality count from automobile accidents as almost 43,000 (Thomas, 2008). Academic, government, and corporate efforts have focused on improving road conditions, vehicle safety, and driver awareness. Using a variation of the temporal considerations introduced in this dissertation for feature-space modeling provides a new foundation of thought for crash prediction.

Much like the feature-space model introduced in Chapter 2 focused on including proximity information in a clustering analysis, for a crash modeling application, we would include the distance to key features in a model to reflect the likelihood of accidents within certain spatial locations at specified temporal intervals. We would not be examining the preference structure of drivers; rather we would use the feature-space construct to build an empirical data-driven model that provides insight into the spatial-temporal environments of accidents. We would then derive the spatial clustering of accidents on the basis of the environments and conditions surrounding past events. The use of hierarchical correlation clustering would potentially aid in predictive analysis.

Much as crime depends on the interaction of a criminal, a victim, and an opportunity, an accident involves the interaction of several factors: drivers, traffic, roads, vehicles, and the environment. Of course, driver error stands out as the main factor that contributes to the majority of vehicle crashes (Miaou et al., 2003), but how do these other factors contribute to the likelihood of crashes? Recent studies model vehicular-accident rates as a function of traffic volumes, speed, and roadway geometry and they use random-effect models to predict the likelihood of accidents within specific spatial regions (Qi et al., 2007; Garber and Wu, 2001). A large body of

crash literature models incident rates of accidents as a realization of a Poisson process. Bayesian hierarchical methods similar to the crime-event modeling described in Chapter 6 have been used to model crash-count data with varying degrees of success (Tunaru, 2002; MacNab, 2003; Miaou et al., 2003; Qin et al., 2005). The methods introduced in Chapters 4 and 5 might offer benefits to accident-modeling research through the inclusion of temporal information in a feature-space model for determining event likelihood. Similarly, the Bayesian formulation introduced in Chapter 6 might offer an alternative approach to those used in current research if adjusted to model event intensity rather than event likelihood.

### 8.2.3 Other Areas of Future Research

From a theoretical perspective, the inclusion of temporal considerations into feature-space event modeling offers benefits not only in the law enforcement community but also within commercial, environmental and other governmental sectors. In the commercial sector, variations of the feature-space framework might be used for modeling temporal patterns in the retail or real estate markets. For example, using the expanded feature-space model would give analysts a method for exploring housing-price levels with respect to urban migration patterns over time. Public health officials could employ either the Bayesian hierarchical framework or the expanded feature-space model to assist with predicting times and places of non-contiguous disease outbreaks.

Equally daunting as the modeling of spatial-temporal data is the challenge of providing the results in a manner appropriate for and usable by an analyst. As Tukey (1977) notes statistical modeling results are not practical or beneficial until the analyst can visualize them effectively: “the greatest value of a picture is when it forces us to notice what we never expected to see.” Tools like ArcGIS do offer

a variation of the screen-dimension solution discussed in Chapter 2. The tracking analyst tool available in ArcGIS allows the analyst to examine the spatial-temporal evolution of events and supports the monitoring of real-time databases for warning of specific patterns (Allen, 2011; Gorr and Kurland, 2011). Using a set of evolution rules that match earlier work in spatial-temporal data mining, analysts may use the model builder tools to set thresholds for monitoring events that follow, coincide, or parallel previously developed historical patterns (Roddick and Lees, 2001).



## References

- Ackerman, R. (2006). Geospatial intelligence grows with conflicts and new allies. *SIGNAL*.
- Allen, D. W. (2011). *GIS Tutorial: Spatial Analysis Workbook*. ESRI Press, Redlands, California.
- Baily, T. and Gatrell, A. (1995). *Interactive Spatial Data Analysis*. Longman Scientific and Technical, Harlow, United Kingdom.
- Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall, Boca Raton.
- Bannatyne, J. C. and Edwards, H. P. (2003). A Bayesian explorations of the relationship between crime and unemployment in New Zealand for the time period: 1986–2002. In *International Workshop on Bayesian Data Analysis*, Santa Cruz.
- Benth, F. E., Saltyte-Benth, J., and Jalinskas, P. (2005). A spatial-temporal model for temperature with seasonal variance. Statistical Research Report No. 5.
- Bernasco, W. and Block, R. (2009). Where offenders choose to attack: A discrete choice model of robberies in Chicago. *Criminology*, pages 93–130.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, pages 192–236.
- Bivand, R. S., Pebesma, E. J., and Gomes-Rubio, V. (2008). *Applied Spatial Data Analysis with R*. Springer, New York.
- Block, R., Galary, A., and Brice, D. (2007). The journey to crime: Victims and offenders converge in violent index offences in Chicago. *Security Journal*, 20:123–137.
- Boba, R. (2005). *Crime Analysis and Crime Mapping*. Sage Publications, London.
- Booth, B. (2000). *Using ArcGIS 3D Analyst*. Environmental Systems Research Institute, Redlands.
- Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30:1145–1159. Retrieved July 2010.

- Brantingham, P. J. and Brantingham, P. L. (1984). *Patterns in Crime*. Macmillan Publishing Company, New York.
- Brown, D. E. (1998). Data mining to catch criminals: The regional crime analysis system (ReCAP). In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, San Diego. IEEE.
- Brown, D. E., Dalton, J., and Holye, H. (2004). Spatial forecast methods for terrorist events in urban environments. In *Proceedings of the Second NSF/NIJ Symposium on Intelligence and Security Informatics*, Heidelberg. Springer-Verlag.
- Brown, D. E. and Liu, H. (1999). A new approach to spatial-temporal criminal event prediction. In *Proceedings of the Environmental Systems Research Institute International User Conference*, San Diego. ESRI.
- Brown, D. E., Liu, H., and Xue, Y. (2001). Mining preferences from spatial-temporal data. In *Proceedings of the SIAM Conference*, Chicago. Society for Industrial and Applied Mathematics.
- Burnett, P. (1976). Behavioral geography and the philosophy of mind. In Golledge, R. and Rushton, G., editors, *Spatial Choice and Spatial Behavior*, pages pp. 23–50. Ohio State University Press, Columbus.
- Burrough, P. A. and McDonnell, R. A. (1998). *Principles of Geographic Information Systems*. Oxford University Press, Oxford.
- Chainey, S. and Ratcliffe, J. (2005). *GIS and Crime Mapping*. John Wiley and Sons, Ltd., West Sussex.
- Chatfield, C. (1975). *The Analysis of Time Series: Theory and Practice*. Chapman and Hall, London.
- Chawla, S., Shekhar, S., Wu, W., and Ozesmi, U. (2001). Modeling spatial dependencies for mining geospatial data. In Miller, H. J. and Han, J., editors, *Geographic Data Mining and Knowledge Discovery*, pages 131–159. Taylor and Francis, London.
- Chen, L., Fuentes, M., and Davis, J. M. (2006). Spatial-temporal statistical modeling and prediction of environmental processes. In Clark, J. S. and Gelfand, A. E., editors, *Hierarchical Modelling for the Environmental Sciences*, pages 121–144. Oxford, New York.
- Christakos, G. (1992). *Random Field Models in Earth Sciences*. Dover Publishing, Mineola, New York.
- Christensen, R. (1991). *Linear Models for Multivariate, Time Series, and Spatial Data*. Springer-Verlag, New York.

- Clark, R. V. (1980). Situational crime prevention: Theory and practice. *British Journal of Criminology*, pages 136–147.
- Clark, R. V. and Eck, J. E. (2005). Crime analysis for problem solvers in 60 small steps. Technical report, Center for Problem-Oriented Policing.
- Cliff, A. D. and Haggett, P. (1998). On complex geographical space: Computing framework for spatial diffusion processes. In Longley, P. A., Brooks, S. M., McDonnell, R., and Macmillan, B., editors, *Geocomputation: A Primer*, pages 231–256. John Wiley and Sons, Chichester.
- Cliff, A. D. and Ord, J. K. (1981). *Spatial Processes*. Pion, London.
- Cliff, A. D. and Ord, J. K. (1973). *Spatial Autocorrelation*. Pion, London.
- Cohen, J., Gorr, W., and Durso, C. (2003). Estimation of crime seasonality: a cross-sectional extension to time series classical decomposition. Working paper.
- Cohen, J., Gorr, W., and Olligschlaeger, A. (2007). Leading indicators and spatial interactions: a crime forecasting model for proactive police deployment. *Geographical Analysis*, 39:105–127.
- Crawley, M. J. (2007). *The R Book*. John Wiley and Sons, West Sussex.
- Cressie, N. A. (1993). *Statistics for Spatial Data*. John Wiley and Sons, New York.
- Cressie, N. A. and Wikle, C. K. (2011). *Statistics for Spatial-Temporal Data*. John Wiley and Sons, Hoboken, New Jersey.
- Dalton, J. (2005). Temporal Analysis and GIS. *Geospatial Solutions*.
- Dangermond, J. (2003). Forward. In Longley, P. A. and Batty, M., editors, *Advanced Spatial Analysis*, page vii. ESRI Press, Redlands.
- Deadman, D. and Pyle, D. (1997). Forecasting recorded property crime using a time-series econometric model. *British Journal of Criminology*, 37:437–445.
- Demers, M. N. (2005). *Fundamentals of Geographic Information Systems*. John Wiley and Son, Hoboken.
- Demsar, U. (2009). Geovisualization and geovisual analytics. In Fotheringham, A. S. and Rogerson, P. A., editors, *The SAGE Handbook of Spatial Analysis*, pages 41–62. Sage, Los Angeles.
- Diggle, P. J. and Ribeiro, P. J. (2007). *Model-based Geostatistics*. Springer Science and Business, New York.
- Diggle, P. J., Tawn, A. J., and Moyeed, R. A. (1998). Model-based geostatistics. *Applied Statistics*, pages 299–350.

- Dobbie, M. J., Henderson, B. L., and Stevens, D. L. (2008). Sparse sampling: Spatial design for monitoring stream networks. *Statistics Surveys*, 2:113–153.
- Eck, J. E., Chainey, S., Cameron, J. G., Leitner, M., and Wilson, R. E. (2005). Mapping crime: Understanding hot spots. Technical report, National Institute of Justice.
- Elsner, J. B., Jagger, T. H., Dickinson, M., and Rowe, D. (2008). Improving multi-season forecasts of North Atlantic hurricane activity. *Journal of Climate*, 21:1209–1219.
- Elsner, J. B., Murnane, R. J., and Jagger, T. H. (2006). Forecasting U.S. hurricanes 6 months in advance. *Geophysical Research Letters*, 33.
- Enders, W. and Sandler, T. (2000). Is transnational terrorism becoming more threatening? *Journal of Conflict Resolution*, 44(3):307–332.
- Enders, W. and Sandler, T. (2002). Patterns of transnational terrorism, 1970-1999: Alternative time-series estimates. *International Studies Quarterly*, 46:146–165.
- Enders, W. and Sandler, T. (2006). *The Political Economy of Terrorism*. Cambridge University Press, Cambridge.
- Ewing, G. O. (1976). Environmental and spatial preferences of interstate migrants in the United States. In Golledge, R. G. and Rushton, G., editors, *Spatial Choice and Spatial Behavior*, pages pp. 250–270. Ohio State University Press, Columbus.
- Faraway, J. J. (2006). *Extending the Linear Model with R*. Chapman and Hall / CRC, Boca Raton.
- FBI (2010). Famous cases. Federal Bureau of Investigations.
- Fotheringham, A. S., Brunsdon, C., and Charlton, M. (2000). *Quantitative Geography*. Sage, London.
- Galton, A. (1997). Space, time, and movement. In Stock, O., editor, *Spatial and Temporal Reasoning*, pages 319–352. Kluwer Academic Publishers, Dordrecht.
- Garber, N. J. and Wu, L. (2001). Stochastic models relating crash probabilities with geometric and corresponding traffic characteristics data. Research Report No. UVACTS-5-15-74.
- Gatrell, A. (1994). Density estimation and the visualization of point patterns. In Hearnshaw, H. M. and Unwin, D. J., editors, *Visualization in Geographical Information Systems*, pages 65–75. John Wiley and Sons, Chichester.
- Gelman, A. and Hill, J. (2007). *Data Analysis using Regression and Multilevel / Hierarchical Models*. Cambridge University Press, Cambridge.

- Getis, A. and Boots, B. (1978). *Models of Spatial Processes*. Cambridge University Press, Cambridge.
- Gilbert, R. O. (1987). *Statistical Methods for Environmental Pollution Monitoring*. John Wiley and Sons, New York.
- Goodchild, M. F. (1998). Different data sources and diverse data structures: Metadata and other solutions. In Longley, P. A., Brooks, S. M., McDonnell, R., and Macmillan, B., editors, *Geocomputation: A Primer*, pages 61–72. John Wiley and Sons, Chichester.
- Goodchild, M. F. (2010). Twenty years of progress: GIScience in 2010. *Journal of Spatial Information Sciences*, (1):3–20.
- Gorr, W., Olligschlaeger, A., and Thompson, Y. (2003). Short term forecasting of crime. *International Journal of Forecasting*, 19:579–594.
- Gorr, W. L. (2009a). Cloudy with a chance of theft. *Wired*.
- Gorr, W. L. (2009b). Forecast accuracy measures for exception reporting using receiver operating characteristic curves. *International Journal of Forecasting*, 25(1):48–61.
- Gorr, W. L. and Kurland, K. S. (2011). *GIS Tutorial for Crime Analysis*. ESRI Press, Redlands, California.
- Gorr, W. L., Olligschlaeger, A. M., Szczypula, J., and Thompson, Y. (1999). Forecasting crime. working paper.
- Groff, E. R. (2007). Situating simulation to model human spatio-temporal interactions: An example using crime events. *Transactions in GIS*, 11(4):507–530.
- Groff, E. R. and LaVigne, N. G. (2002). Forecasting the future of predictive crime mapping. In Tilley, N., editor, *Analysis for Crime Prevention*, volume 13 of *Crime Prevention Series*, pages 29–57. Lynne Rienner Publishers, Monsey, NY.
- Gunderson, L. and Brown, D. (2000). Using a multi-agent model to predict both physical and cyber criminal activity. In *2000 IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 2338–2343, Nashville.
- Guptill, S. (1995). Temporal information. In Gupta, S. and Morrison, J. L., editors, *Elements of Spatial Data Quality*. Elsevier Science Limited, Tarrytown.
- Hadlak, S., Tominski, C., Schulz, H.-J., and Schumann, H. (2010). Visualization of hierarchies in space and time. In *GeoVA(t) - Geospatial Visual Analytics Proceedings: 10–11 May 2010*.

- Haining, R. (2009). The special nature of spatial data. In Fotheringham, A. S. and Rogerson, P. A., editors, *The SAGE Handbook of Spatial Analysis*, pages 5–25. SAGE, Los Angeles.
- Higdon, D. (2007). A primer on space-time modeling from a Bayesian perspective. In Finkenstadt, B., Held, L., and Isham, V., editors, *Statistical Methods for Spatio-Temporal Systems*, pages 217–281. Chapman and Hall/CRC, Boca Raton.
- Huddleston, S. and Brown, D. E. (2009). A statistical threat assessment. *Systems, Man, and Cybernetics, Part A.*, 39:1307–1315.
- Huddleston, S., Learmonth, G., and Fox, J. (2008). Changing knives into spoons. *Systems and Information Engineering Design Symposium 2008*, pages 255–260.
- Huffer, F. W. and Wu, H. (1998). Markov Chain Monte Carlo for Autologistic Regression Models with Application to the Distribution of Plant Species. *Biometrics*, 54:509–524.
- Hughes, J., Haran, M., and Caragea, P. C. (2011). Autologistic models for binary data on a lattice. *Environmetrics*.
- Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). *Statistical Analysis and Modeling of Spatial Point Patterns*. John Wiley and Sons Ltd, West Sussex.
- Jacob, B., Lefgren, L., and Moretti, E. (2007). The dynamics of criminal behavior. *The Journal of Human Behavior*, 42:489–527.
- Jagger, T. H., Niu, X., and Elsner, J. B. (2002). A space-time model for seasonal hurricane prediction. *International Journal of Climatology*, 22:451–465.
- Jesse, L. (2000). Web-Enabled Temporal Analysis System. Retrieved January 10, 2007 from webtas.com.
- Johnston, K., Ver Hoef, J. M., Krivoruchko, K., and Lucas, N. (2001). *Using ArcGIS Geostatistical Analyst*. ESRI, Redmond.
- Kaluzny, S. P., Vega, S. C., Cardoso, T. P., and Shelly, A. A. (1997). *S+ SPATIAL-STATS User’s Manual for Windows and Unix*. Springer, New York.
- Keim, D. A. and Ward, M. (2003). Visualization. In Berthhold, M. and Hand, D. J., editors, *Intelligent Data Analysis*, pages 403–428. Springer-Verlang, Berlin.
- Kerchner, S. H. (2000). Spatial-temporal event prediction. Master’s thesis, University of Virginia, Charlottesville.
- Kery, M. (2010). *Introduction to WinBUGS for Ecologists*. Elsevier, Amsterdam.
- Kewley, R. H. and Evangelista, P. (2007). Evaluating machine learning methods for geospatial prediction problems. Being prepared for submission to IEEE.

- Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W. (2004). *Applied Linear Statistical Models*. McGraw-Hill, New York.
- Langran, G. (1992). *Time in Geographic Information System*. Taylor and Francis, London.
- Lawson, A. B. (2009). *Bayesian Disease Mapping*. CRC Press, Boca Raton.
- Le, N. D. and Zidek, J. V. (2006). *Statistical Analysis of Environmental Space-Time Processes*. Springer Science and Business Media, New York.
- Levine, N. (2009). Crimestat: A spatial statistics program for the analysis of crime incident locations. In *GeoComputation99*, Washington, DC. Ned Levine and Associates and the National Institute of Justice.
- Liang, J. (2001). *Simulating Crimes and Crime Patterns Using Cellular Automata and GIS*. PhD thesis, University of Cincinnati, Cincinnati.
- Liu, H. and Brown, D. E. (1998). Spatial-temporal event prediction: A new model. In *Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics*, volume 3, pages 2933–2937, San Diego. IEEE.
- Liu, H. and Brown, D. E. (2003). Criminal incident prediction using a point-pattern based density model. *International Journal of Forecasting*, 19:603–622.
- Liu, L., Wang, X., Eck, J., and Liang, J. (2005). Simulating crime events and crime patterns in a RA/CA model. In Wang, F., editor, *Geographic Information Systems and Crime Analysis*. Idea Group Publishing, Hershey, Pennsylvania.
- Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10:325–337.
- MacEachren, A., Buttenfield, B., Campbell, J., DiBiase, D., and Monmonier, M. (1992). Visualization. In Abler, R. F., Marcus, M. G., and Olson, J. M., editors, *Geography's Inner Worlds*, pages 101–137. Rutgers University, New Brunswick.
- MacNab, Y. C. (2003). A bayesian hierarchical model for accident and injury surveillance. *Accident Analysis and Prevention*, 35:91–102.
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., and Winkler, R. (1982). *The accuracy of extrapolation (time series) methods: Results of a forecasting competition*, volume 1. John Wiley and Sons, Ltd.
- Marin, J.-M. and Robert, C. P. (2007). *Bayesian Core: A Practical Approach to Computational Statistics*. Springer, New York.

- McCulloch, P. and Nelder, J. (1989). *Generalized Linear Models*. Chapman and Hall, London.
- McFadden, D. (1986). The choice theory approach to market research. *Marketing Science*, 5:275–297.
- Meeker, J. W., Parsons, K. J., and Vila, B. J. (2002). Developing a GIS-based regional gang incident tracking system. In Decker, S. H. and Reed, W. L., editors, *Responding to Gangs: Evaluation and Research*, pages 289–329. National Institute of Justice, Washington DC.
- Miaou, S.-P., Song, J. J., and Mallick, B. K. (2003). Roadway traffic crash mapping: A space-time modeling approach. *Journal of Transportation and Statistics*, 6:33–57.
- Mickolus, E. F. (1987). Comment - terrorists, governments, and numbers: Counting things versus things that count. *The Journal of Conflict Resolution*, 37.
- Midlarsky, M. I., Crenshaw, M., and Yoshida, F. (1980). Why violence spreads. *International Studies Quarterly*, 24.
- Miller, H. J. (2009). The special nature of spatial data. In Fotheringham, A. S. and Rogerson, P. A., editors, *The SAGE Handbook of Spatial Analysis*, pages 397–418. SAGE, Los Angeles.
- Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., and Tita, G. E. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108.
- Morgan, J. D. (2009). A visual time geographic approach to crime event profiling. *Crime Mapping: A Journal of Research and Practice*, 1:8–39.
- Mosher, C. J., Miethe, T. D., and Hart, T. C. (2011). *The Mismeasure of Crime*. Sage, Los Angeles.
- National Research Council, N. (2003). *IT Roadmap to a Geospatial Future*. The National Academies Press, Washington, DC.
- Nelder, J. and Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A*:370–384.
- Olligschlaeger, A. M. (1997). Artificial neural networks and crime mapping. In Weisburd, D. and McEwen, T., editors, *Crime Mapping and Crime Prevention*, volume 8 of *Crime Prevention Series*, pages 313–347. Criminal Justice Press, Monsey, NY.
- Olligschlaeger, A. M. and Gorr, W. L. (1997). Spatio-temporal forecasting of crime: application of classical and neural network methods. Working paper.



- Ormsby, T., Napoleon, E., Burke, R., Groessl, C., and Feaster, L. (2001). *Getting to Know ArcGIS*. ESRI Press, Redlands, California.
- Peng, R. D. (2008). A method for visualizing multivariate time series data. *Journal of Statistical Software*, 25.
- Petris, G., Petrone, S., and Campagnoli, P. (2009). *Dynamic Linear Models with R*. Springer, Dordrecht.
- Porter, M. D. (2006). *Detecting Space Time Anomalies in Point Process Models of Intelligent Site Selection*. PhD thesis, University of Virginia, Charlottesville.
- Provence, L. (2007). Rampage: Home invasion adds to attack fears. *The Hook*. Retrieved 24 October 2007.
- Qi, Y., Smith, B. L., and Guo, J. (2007). Freeway accident likelihood prediction using a panel data analysis approach. *Journal of Transportation Engineering*, 133:149–156.
- Qin, X., Ivan, J. N., Ravishanker, N., and Liu, J. (2005). Hierarchical Bayesian estimation of safety performance functions for two-lane highways using Markov chain Monte Carlo modeling. *Journal of Transportation Engineering*, pages 345–351.
- Ratcliffe, J. (2000). Aoristic analysis: The spatial interpretation of unspecific temporal events. *International Journal of Geographical Information Science*, 14:669–679.
- Ratcliffe, J. (2002). Aoristic signatures and the temporal analysis of high volume crime patterns. *Journal of Quantitative Criminology*, 18:23–43.
- Ratcliffe, J. H. (2004). The hotspot matrix: A framework for the spatio-temporal targeting of crime reduction. *Police Practice and Research*, 5:5–23.
- Ratcliffe, J. H. and McCullagh, M. J. (1998). Aoristic crime analysis. *International Journal of Geographical Information Science*, 12:751–764.
- Rengert, G. F. (1997). Auto theft in central philadelphia. In Homel, R., editor, *Policing for Prevention: Reducing Crime, Public Intoxication and Injury*, volume 7 of *Crime Prevention Series*, pages 199–219. Lynne Rienner Publishers, Monsey, NY.
- Rex, B. and Rasmussen, R. (2000). Starlight and map objects for data mining crime information. In *ESRI Users Conference*. ESRI.
- Rickman, D. S., Miller, S. R., and McKenzie, R. (2009). Spatial and sectoral linkages in regional models: A Bayesian vector autoregression forecast evaluation. *Papers in Regional Science*, 88(1):29–41.

- Riese, S. R. (2001). *Estimating the Probability of Landmine Contamination in a Non-Combat Environment*. PhD thesis, University of Virginia, Charlottesville.
- Ripley, B. D. (1981). *Spatial Analysis*. John Wiley and Sons, Inc, New York.
- Roddick, J. and Lees, B. G. (2001). Paradigms for spatial and spatio-temporal data mining. In Miller, H. and Han, J., editors, *Geographic Data Mining and Knowledge Discovery*. Taylor and Francis, London.
- Rogerson, P. A. (2006). *Statistical Methods for Geography*. SAGE, London.
- Rossmo, K., Laverly, I., and Moore, B. (2005). Geographic profiling for serial crime investigation. In Wang, F., editor, *Geographic Information Systems and Crime Analysis*, pages 137–152. Idea Group Publishing, Hershey, Pennsylvania.
- Rue, H. and Follestad, T. (2003). Gaussian Markov Random Field models with applications in spatial statistics. In *19th Nordic Conference on Mathematical Statistics*, Stockholm. Norwegian University of Science and Technology.
- Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields*. Chapman and Hall / CRC, Boca Raton.
- Schabenberger, O. and Gotway, C. A. (2005). *Statistical Methods for Spatial Data Analysis*. Chapman and Hall / CRC, Boca Raton.
- Shaw, S.-L. (2010). Time geography: Its past, present, and future. In *Proceedings of the American Association of Geographers Meeting*.
- Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T. (2005). ROCr: Visualizing Classifier Performance in R. *Bioinformatics*, 20:3940–3941.
- Smith, M. A. and Brown, D. E. (2004). Hierarchical choice modeling of terror attack site selection. *Decision Support Systems*.
- Snodgrass, R. (1992). Temporal databases. In Frank, A., Campari, I., and Formentini, U., editors, *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. Springer-Verlag, Berlin.
- Swets, J. A., Dawes, R. M., and Monahan, J. (2000). Better decisions through science. *Scientific American*, 283(4):82–87.
- Thomas, K. (2008). Study: Traffic crashes cost billions. Retrieved April 5, 2008.
- Townsley, M., Homel, R., and Chaseling, J. (2000). Repeat burglary victimisation: Spatial and temporal patterns. *Australian and New Zealand Journal of Criminology*, 33(1):37–63.

- Townsley, M. and Pease, K. (1997). Hot spots and cold comfort: the importance of having a working thermometer. In Tilly, N., editor, *Analysis for Crime Prevention*, volume 13 of *Crime Prevention Series*, pages 59–69. Lynne Rienner Publishers, Monsey, New York.
- Tseloni, A. (2006). Multilevel modeling of the number of property crimes: household and area effects. *Journal of the Royal Statistical Society*, 169:205–233.
- Tufte, E. (1997). *Visual Explanations*. Graphic Press, Cheshire, Connecticut.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley Publishing Company, Reading, Pennsylvania.
- Tunaru, R. (2002). Hierarchical Bayesian models for multiple count data. *Austrian Journal of Statistics*, 31:221–229.
- Unwin, D. (1981). *Introductory Spatial Analysis*. Methuen, London.
- Wickle, C. K. (2003). Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology*, 84(6):1382–1394.
- Withers, S. D. (2002). Quantitative methods: Bayesian inference, Bayesian thinking. *Progress in Human Geography*, 26:553–566.
- Wolpert, R. L. and Ickstadt, K. (1998). Poisson - gamma random field models for spatial statistics. *Biometrika*, 85:251–267.
- Xue, Y. and Brown, D. E. (2003). Decision based spatial analysis of crime. In *Proceedings of the NSF/NIJ Symposium on Intelligence and Security Informatics*, pages 153–167, Tuscon. Springer-Verlag Heidelberg.
- Xue, Y. and Brown, D. E. (2006). Spatial analysis with preference specification of latent decision makers for criminal event prediction. *Decision Support Systems*, 41:560–573.
- Yuan, M. (2009). Challenges and critical issues for temporal GIS research and technologies. In Karimi, H. A., editor, *Handbook of Research on Geoinformatics*, pages 144–153.
- Zhu, J., Huang, H. C., and Wu, J. (2006). Modeling spatial-temporal binary data using Markov random fields. University of Washington – Department of Statistics.
- Zhu, J., Zheng, Y., Carroll, A. L., and Aukema, B. H. (2008). Autologistic regression analysis of spatial-temporal binary data via Monte Carlo maximum likelihood. *Journal of Agricultural, Biological and Environmental Statistics*, 13:84–98.

## VITA

Jonathan M. Fox

2034 Locke Lane

Phone: (434) 566-2921

Charlottesville, VA 22911

Email: [jmf3a@virginia.edu](mailto:jmf3a@virginia.edu)

## Education

- B.S. Engineering, United States Military Academy, 1991.
- M.A. Military Science, United States Army Command & General Staff College, 2003.

## Employment

- United States Army 1991–present.

## Publications

### Archival Journal Articles

- Fox, J. and Brown, D.E., Hierarchical feature-space modeling of temporal patterns in criminal site-selection problems, 2012, *Crime Mapping: A Journal of*

*Research and Practice*, Accepted for publication – expected publication Summer of 2012.

## Peer-Reviewed Conference Proceedings

- Fox, J. and Brown, D.E., Using temporal indicator functions with generalized linear models for spatial-temporal event prediction, *Proceedings of 2012 Conference on Systems Engineering Research*, 2012.
- Fox, J., Learmonth, G.P., and Brown, D.E., Simulating spatial-temporal pulse events in criminal site selection problems, *Agent Directed Simulation Symposium*, 2012.
- Fox, J., Huddleston, S.H., Gerber, M., and Brown, D.E., Investigating a Bayesian hierarchical framework for feature-space modeling of criminal site-selection problems, *Proceedings of the 23rd Midwest Artificial Intelligence and Cognitive Science Conference*, 2012.

## Conference Proceedings – Not Peer Reviewed

- Fox, J.M. and Huddleston, S.H., Mathematical literacy as fundamental component of training cyber warriors, *Proceedings of the 2008 Information and Cyberspace Symposium*, 2008.
- Huddleston, S.H., Learmonth, G.P., and Fox, J.M., Changing knives into spoons, *Proceedings of the 2008 IEEE Systems and Information Engineering Design Symposium*, 2008.

## Conference Posters

- Fox, J.M., Hierarchical modeling of vehicular accidents with spatial-temporal considerations for improved resource allocation, Abstract and Poster Presentation for the 2008 University of Virginia Research Symposium.
- Fox, J.M. and Huddleston, S.H., Examining the oil-spot methodology of fighting an insurgency using hybrid agent-based modeling, Abstract and Poster Presentation for the 2008 Winter Simulation Conference.
- Fox, J.M., Spatial-temporal event prediction in an asymmetric environment, Abstract and Poster Presentation for the 2007 University of Virginia Research Symposium.

## Conference Presentations

- Fox, J.M. and Huddleston, S.H., Training cyber warriors – moving beyond tools and towards fishing, Military Operations Research Symposium (June 2011 Monterey, CA).
- Fox, J.M., Cyber fratricide, Invited presentation at the Air Force Research Laboratory Workshop on Cyber Fratricide (April 2010 San Antonio, TX).
- Brown, D.E., Huddleston, S.H., and Fox, J.M., Cultural-temporal analysis and forecasting of asymmetric attacks, KIMCOE Knowledge Fusion Workshop (November 2007 Cambridge, MD).