A Critical Exploration, Analysis, and Contextualization of Discrete and Continuous Models of Criminal Behaviour for the Prediction of Future Real-World Crime Occurrence

By
Jakob Conrad

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Bachelor of Science with Honours in Mathematics

April 2024, Halifax, Nova Scotia

Approved: ____________________
Dr. Paul Muir
Supervisor

Approved: ____________________
Dr. Wendy Finbow-Singh
Reader

Approved: ____________________
Dr. Cathy Conrad
Reader

Date: April 22, 2024

Acknowledgement

I owe the completion of this project to so many people, without whom I would not be sitting here today with a completed thesis. Though the past two years of work on this project have some of the most challenging of my life up to this point, it is with an extremely grateful heart that I am able to look back to all those who have supported both me and this work and who have continued to encourage me to fight through to the end.

To Dr. Paul Muir, who has not only helped shape this project from its very beginning but who has pushed me to fundamentally reconsider the ways in which I understand teaching, learning, and what I am personally capable of. To Dr. Neil Balan, who has helped review, shape, and advise a number of chapters within this completed thesis, and who has constantly addressed his students with an unparalleled level of kindness and support. To Dr. Rachel Zellars, who gave me the confidence years ago to begin this work and who has shown me so much compassion throughout my years at Saint Mary's.

To Dr. John Irving and Dr. Diego Rojas, whose love for both mathematics and teaching has been infectious. To Dr. Benita Bunjun and Dr. Val Marie Johnson, who helped challenge my preconceptions and who have pushed to greatly expanded the original scope of this project. To Dr. Mitja Mastnak and Dr. Luigi Gallo, who both took a chance on me as a young undergraduate student to engage me in the world of research.

To Amy Kehoe, who has too-many times had to endure my nonsense, and for her seemingly never-ending kindness toward all those around her. To Hannah Thomas, who has seen me at my worst and who regardless of circumstance has encouraged me to keep going. To Andrew Hutchins, who without knowing helped me regain my passion for this project at a time when I was contemplating abandoning it altogether.

To Scott Conrad, who throughout my 23 years of life has taught me eternal kindness and love. To Cathy Conrad, who has been my number-one ally and supporter in life and who has made me into the person I am today. To Ellie and Martin Suso, who have given my life a meaning that I could not have previously known. To Bubby, I love you so much more than you could know.

And lastly but certainly not least to Kevin, for spending so many of these long nights by side.

A Critical Exploration, Analysis, and Contextualization of Discrete and Continuous Models of Criminal Behaviour for the Prediction of Future Real-World Crime Occurrence

Abstract

This interdisciplinary project explores "predictive policing", a blanket term given to a number of crime prediction algorithms and tools used by police departments across the globe in an effort to predict and pre-empt crime occurrence. This project attempts to cut through the profound amount of both positive and negative rhetoric surrounding predictive policing software to understand what theory they are based on, how they are actually implemented in software, and how they interface with police officers working on-the-beat. Through a literature review of empirical environmental criminology research, a theory of how crime self-concentrates in space and time is discussed, as well as the potential explanations for this behaviour. Using this literature, two predictive models of criminal behaviour are introduced, explained, tested, and analysed, to understand how empirical crime observations can be translated into software. Using numerical results obtained from these models in conjunction with existing meta-critiques of predictive policing tools, the argument is made that while current predictive policing tools may hold theoretical value in the field of crime prediction, they have enough significant drawbacks as to cast doubt on their use to predict real world crime.

Contents

List of Tables

List of Figures

Chapter 1

Introduction: Predictable Crime & Predetermined Punishment

First published in 2021 by investigative reporter Matt Stroud, "Heatlisted" [45] tells the story of Chicago citizen Robert McDaniel's interactions with the Chicago Police Department (CPD) leading up to his eventual shootings. Summarizing his narrative, a 22-year-old McDaniel living in Austin, a predominantly Black neighbourhood of Chicago, was visited at his home by CPD officers sometime in mid-2013. This police visit was not in response to any misconduct on the part of McDaniel, but was prompted based on results from a new, secretive policing pilot program. Although it was known only at the time as the "CPD heat list", McDaniel was victimized by what has come to be known as the CPD's "Strategic Subjects List" (SSL). Under this program the CPD implemented a "Strategic Subject Algorithm" created by researchers from the Illinois Institute of Technology designed to attribute a numerical score to those with criminal records in order to quantify their risk of being involved in a future violent crime [9]. It was due to the SSL pilot program that Robert McDaniel was identified and singled-out by CPD officers: Although McDaniel had no previous record of violent crime, having been charged only with selling marijuana and street gambling, the SSL identified McDaniel as being more likely than over 99.9% of Chicago's population to be involved in a deadly shooting, either as the victim or as the perpetrator [45]. Commenting on the nature of the SSL author Matt Stroud invokes the imagery of "pre-crime" from Philip K. Dick's 1956 novella "Minority Report", in that the CPD was attempting to figure out "who would commit a crime before they actually did" [45].

Following his visit from the CPD, Robert McDaniel became someone of suspect within his community. Based on their algorithmic predictions, the CPD began to surveil McDaniel intensely, following him throughout his community, staking him out at his work, and routinely questioning those close to him about his daily activities [45]. As a result of the intense police presence surrounding him, McDaniel's neighbours, friends, and community turned against him, suspecting him of snitching to the CPD. McDaniel spent years living in the shadow of the CPD, and one day in early-2017 was the victim of a drive-by shooting just blocks away from his home. One bullet pierced McDaniel in his knee, leading him to be hospitalized with non-life-threatening injuries [45]. Following this, in the Summer of 2020 McDaniel was the victim of yet another shooting, luckily surviving for a second time. When questioned by Stroud on the motivation behind his two shootings, McDaniel comments decisively that while the SSL was designed to reduce violent crimes, "it brought violence directly to (me)" [45].

I chose to begin this project by introducing Robert McDaniel's story for two key reasons: Firstly, Chicago's SSL is just a small piece in a much larger network of what has come to be known as the regime of "predictive policing". Using advancements in data science and computing power achieved over the past number of decades, both researchers and private technology firms have been hard at work, developing software tools designed to be able to predict crime before it occurs in the first place. While a number of these tools (such as the SSL) are "individual-based", attempting to quantify either the likelihood with which an individual will be involved in a future crime or will re-offend, the second class of these tools are "place-based", attempting to quantify where crime is likely to occur in a given geographic region. Regardless of regime, Robert McDaniel's story serves as a gateway into the ideological warfare surrounding predictive policing today. On one side of this conflict are those strongly advocating against the use of predictive policing tools: Not only Matt Stroud, but a number of other investigative reporters have jumped at the chance to criticize the seemingly dystopian nature of predictive policing (see [15, 30, 16, 10, 3] to name but a handful). In addition, larger institutions have broadly advocated against the use of predictive policing technologies, including both the RAND Corporation in [17] and Data for Black Lives in [4]. Perhaps most damningly, an open letter published in the October 2020 issue of Notices of the American Mathematical Society urges mathematicians at large to avoid providing police departments with mathematical resources or research, specifically indicting predictive policing as a problematic technology that deserves a broad audit from the mathematical community at large [25]. On the other side of the predictive policing debate, many police departments have openly spoken in favour of the technology, commenting that predictive policing has contributed to noticeable reductions in overall crime occurrence [38]. It is no small task to sift through the mountain of rhetoric surrounding predictive policing, to understand whether such tools are actually of merit or whether they serve only to entrench existing police violence.

Concerning predictive policing, there are a number of central issues that I intend to unpack and analyse to get at the nature of predictive policing as an area of study. Before outlining these central issues, I would like to note that there are a great number of "predictive policing tools" that are in active use today: I opt here to focus on place-based predictive policing tools over individual-based predictive policing tools, for the sake of scope. Even within the category of place-based predictive policing tools however, one runs up against the "blue wall" of policing, in that it is often difficult to obtain concrete information about day-to-day police operations. I thus restrict the scope of the algorithms I intend to analyse within this project to two key models of residential burglary, first introduced by Martin B. Short et al. in [44] and most notably used in some form within the predictive policing software *PredPol* [37]. While my ultimate conclusions pertaining to these two models are in no way intended to be statements about predictive crime modelling in general, later insights from an analysis of these two models do have results generalizable to other place-based predictive policing models.

Turning specifically then to place-based predictive policing and the two models of residential burglary introduced by Short et al. in [44], I am interested to question the theoretical foundations of these models. Upon what criminological work are these models based, and are these models able to be motivated by data obtained by empirical criminologists? That is, does crime in the real world display definitive and predictable behaviour that can be leveraged to create a mathematical model of criminal behaviour? In the context of Short et al.'s two models of residential burglary, I am curious to understand the mathematics that governs these models at their most fundamental level: What is the basis for the expectation that these models are able to form predictions about where crime will occur, and what behaviours can be expected from these models? Further, do these models accurately represent what real-world criminal behaviour looks like? Can these models feasibly be used to generate crime predictions that are accurate with respect to the real world? Finally, I look to see how these models are actually implemented within real-world police departments. How do police officers interface with predictive policing tools, and how do predictive policing tools translate crime predictions into concrete suggestions for police officers? How does historic policing data play a role in tuning predictive policing models to real-world environments, and do these models hold up against empirical field trials? These questions will be discussed in sequence throughout the following chapters.

Charting a more detailed roadmap for this thesis, I begin in Chapter 2 by discussing environmental criminology, the sub-area of criminology concerned with understanding the geography of crime. I examine the work of a number of environmental criminologists to quantify both spatial and temporal trends in criminal behaviour that appear to be present regardless of context. Building on this work, I discuss criminologist David Weisburd's "physical law of crime concentrations" [49] as a foundational theory to motivate predictive policing as an area of study. I then briefly discuss potential explanations for the aforementioned trends in criminal behaviour, from both criminology and behavioural psychology. In Chapter 3, I use both these potential explanations and work from Short et al. to motivate and describe Short et al.'s discrete model of residential burglary. I discuss the details of this agent-based model of crime, as well as possible parameter choices for this model. I then use my implementation of this model in Python to perform a number of simulations of criminal behaviour, the results of which are included at the end of Chapter 3. In Chapter 4, I begin by introducing a general discussion of partial differential equations (PDEs), and their potential solution behaviour. I then follow Short et al.'s derivation of a continuous model of residential burglary, based on the discrete model of residential burglary. Before performing simulations using this second model of criminal behaviour, I describe both the BACOL and BACOLI solvers for the computation of error-controlled numerical solutions to systems of parabolic PDEs. Using a special Python wrapper for BACOLI (BACOLI_PY), I perform a number of simulations using the one-dimensional variant of the continuous model of criminal behaviour, the results of which are included at the end of Chapter 4. In Chapter 5, I reintroduce the results from both the discrete and continuous models of residential burglary, to compare and contrast against each other. I discuss my results in the context of the environmental criminology literature introduced first in Chapter 2, to assess whether either model agrees with the work of empirical criminologists. In Chapter 6, I pull back from a detailed discussion of the discrete and continuous models of criminal behaviour to discuss implementation challenges with predictive policing tools at a more fundamental level. I discuss how predictive policing as an area of study is actually far more poorly understood than it may seem to an outsider. I speak to how regardless of the underlying mathematics, there are fundamental challenges to creating predictive policing tools, such as the issues of flawed historic policing data and problematic officer integration. I finally discuss the few empirical field trials of predictive policing tools that exist within the peer-reviewed literature, to assess whether there is any evidence pointing to predictive policing having a noticeable impact on crime reduction.

As a final note, I mentioned previously that I chose to include Robert McDaniel's story for *two* key reasons. The second of the two reasons why I raise McDaniel's story, is to underscore this entire project in the real systemic violence that many face at the hands of the police. Within the coming chapters I discuss the criminological theory contributing to predictive crime modelling, the applied mathematics behind building and solving discrete and continuous models of criminal behaviour, and some critical insights from sociologists on the efficacy of crime prevention, among other things. It can be easy when digging deep into these muddy details of predictive policing to forget that predictive policing tools are in use in the real world to this day, and are in at least some way contributing to police violence against racialized and otherwise marginalized communities. It is my hope that Robert McDaniel's story should serve as a continuing reminder throughout this project: That predictive policing is not simply some mathematical curiosity to be debated within the ivory halls of academia, but a regime of policing that has the capacity to enact extreme violence within innumerably many communities. Let Robert's story underline the gravity of predictive policing as a topic for discussion and analysis.

Chapter 2

Environmental Criminology and the Hotspot Behaviour of Crime

"Since time immemorial, humans seem to have had a natural pre-disposition for committing mayhem and coming up with novel sanctions to punish those who perpetrated the acts" [39, pg. 18]. For millennia humans have engaged in behaviour seen as antithetical to the societal norms of the time, and have received punishment according to their actions. Although the modern conceptions of law, order, crime, and punishment may be unrecognisable in the context of earlier societies, crime and criminality themselves are as old as the human race [39]. Developing in parallel to the existence of deviant behaviour, forms of law and policing have attempted to treat the endemic issue of crime for organized society. While the fight against crime has been ongoing for thousands of years, it has only been in the past handful of decades that strategies of crime prevention have shifted from being reactive to being proactive in nature. Whereas the crime problem was once primarily addressed by harshly punishing criminal behaviour [31] (and although in many ways this practice is all the more prevalent today [42, 13]), in more recent years we have begun to consider fighting crime via attempts to prevent it in the first place. By means of increased state surveillance and a more profound culture of police dominance, some argue that crime can now be addressed by police forces before it even occurs [29, 36]. Within the overall cognitive shift in policing towards crime prevention, the idea of "predicting" where crime will occur has begun to gain traction. Suppose that if we are able to use empirical data in tandem with criminological theory to understand exactly where, when, and why crime occurs, would we be able to stop crime at the source? There is a growing body of work analysing how crimes are distributed both in space and in time, in the effort to further develop and refine proactive policing strategies. By analysing the geography of crime it has been found that remarkably, urban crime displays profound spatial and temporal patterns. Such foundational work like that of David Weisburd in [49] (to be discussed in further detail in throughout this chapter) has laid the foundations for software marketed to be able to "predict" where crime will occur in the future (predictive policing software), a very strong claim that deserves a close critique. To perform this critique, it is first essential to understand the empirical motivation for and theory underlying this predictive policing software.

Within this chapter I consider the field of *environmental criminology*, the area of criminology concerned with understanding and mapping the geography of crime. I define a number of key concepts relating to environmental criminology and consider how environmental criminologists employ the idea of "microplaces" to understand the spatial dynamics of crime at the resolution of individual street segments. I then look at a number of empirical studies examining the dispersion of crime in urban environments at this small spatial resolution, highlighting that crime displays a profoundly consistent "hotspot" behaviour within cities. I consider a number of explanations for this hotspot behaviour, more specifically in the context of residential burglary. By using the insights of a number of environmental criminologists, I examine how we are in theory able to craft a mathematical model designed to "predict" crime.

2.1 The criminology of places

Before digging into environmental criminology and crime hotspots it is important to discuss the words often used to describe the geography of crime, and their precise meanings. Foremost, what precisely is meant by "crime"? While there is room to debate the nature of crime itself, what constitutes a "criminal act", and who is liable to punish such acts [23], for the purposes of discussing environmental criminology I opt here to define crime in a relatively simple sense. Following a legal precedent, I consider crime to be "anything prohibited under the criminal law - the criminal law being that branch of law dealing with state punishment" [23, pg. 609]. Of note is that there is a seemingly small but important distinction between the words "crime" and "criminality". Whereas the word "crime" most often refers to actual concrete criminal acts, the word "criminality" most often refers to the ability and desire for a person to commit crimes given the right circumstances [51]. In a way, the difference between questioning crime and questioning criminality can be understood as approaching the study of criminology from either the fields of geography or sociology. Studies of crime are focused more on the act itself, where and when crime is most likely to occur, what types of crimes prevail in certain environments, what environmental conditions lead to changes in crime distribution, etc. In contrast, studies of criminality are focused on the people who commit crimes, their motivations, life circumstances, economic standings, family history, etc. [51]. To focus on environmental criminology and hotspots of crime therefore focuses on questions of crime itself over questions of criminality.

Discussing the geography of crime involves "crime places": places where crime is likely to or has occurred. Hence, another important term to consider is that of "place". Criminologist David Weisburd understands the word place to mean an expanse that is "defined by its internal characteristics that differentiate it from other places" [51, p.17]. This conceptualization of place does not put restrictions on the scope nor the contents of a place. The idea of place is left as general as possible. In this sense, any specific and distinguishable expanse can be considered a place: the produce section of an individual grocery store can be understood as a place, just as the entirety of the state of Ohio can be understood as a place. As there are no restrictions placed on the context of a place, one is able to label and consider any specific area as a “crime place”, using the tools of environmental criminology accordingly.

In addition to precise definitions of “crime” and “place”, there are a few terms inherent to environmental criminology and more specifically. predictive policing, in specific that merit precise definitions. Often when discussing the geography of crime and predictive crime modelling, the “criminal attractiveness” (sometimes used interchangeably with “attractiveness”) of a given place is considered. In essence, criminal attractiveness is a qualitative measure of how likely a certain place is to be victimized, relative to the other places in a specific environment (often a city) [44]. Although in practice it is impossible to measure and ascribe a value indicating how attractive a certain place is to a prospective criminal, using the term criminal attractiveness gives language to discuss how differing places may appear to potential criminals. When studying crime and place, it is also important to consider the impacts of scale. Given the more general definition of “place”, one can study crime and place using entire neighbourhoods and cities as the spatial unit of analysis. I follow the convention set by David Weisburd in [51] to call places on this scale “macroplaces”, where one studying crime at this resolution would be considering the “macrogeography” of crime. Conversely, one is able to study crime and place at the spatial resolution of city blocks and individual street addresses. Places at this smaller resolution are often called “street segments”, or “microplaces”, where one studying crime at this smaller resolution would be considering the “microgeography” of crime [51].

Although there is a much to be learned by considering the dynamics of crime at the spatial resolution of macroplaces (see [52] for a broad overview of such research), I restrict the conversation here to that of microplaces. In the effort to predict individual crimes, any predictive model of criminal activity must clearly work on the scale of microplaces. As will be discussed in the coming sections, crime displays profound patterns at the spatial resolution of microplaces, with these patterns motivating the development of the predictive crime models in use today. I follow the convention set by Lawrence Sherman et al. in their influential 1989 article "Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place" to label the study of crime and place at a microgeographic scale as "the criminology of places". This is in contrast to the study of criminality, which is aptly labelled as "the criminology of people", or "the criminology of collectives" [43]. An important note on the criminology of places is that the type of crime need not be specified to study crime at place. The spatial dynamics of all different forms of crime can be considered at the resolution of microplaces, although, as will be discussed in further detail later, the microgeographic dynamics of residential burglary are of particular interest to those attempting to develop predictive models of crime. For a number of reasons (to be discussed in the following sections), the spatial dynamics of residential burglary are far simpler to quantify when compared with other forms of crime (for example, motor vehicle theft or hit-and-runs), hence this focus.

Within the wider sphere of criminology, there has been relatively little work considering the spatial dynamics of crime in general, let alone at the resolution of microplaces. In a 2015 study, David Weisburd finds that throughout the history of modern criminology, researchers have been far more interested in questioning criminality rather than questioning the nature of crime itself: Examining *Criminology*, one of the leading peer-reviewed journals dedicated to the study of criminology, Weisburd finds that approximately 66% of the 719 empirical articles published between 1990 and 2014 in *Criminology* focus on people and criminality, over focusing on place and crime [49]. Even further, whereas approximately 30% of the 719 empirical articles published in *Criminology* between 1990 and 2014 focus on studying crime and place at macrogeographic units of analysis, only approximately 4% of these 719 articles focus on studying crime and place at microgeographic units of analysis. Although Weisburd notes that the number of empirical articles considering microplaces in criminology published each year has been increasing, there is still a surprising absence of research considering microplaces within the context of criminology. Considering this disparity of articles published on microplaces, Weisburd asks the following key question: "Why should there be a turning point [in modern criminology] that would focus on microgeographic units of analysis?" [49, p.135].

2.2 The strong correlation between crime and microplaces

While there have been relatively few publications considering microplaces within the criminology of places when compared with the entirety of criminological publications [49], the growing body of literature shows that crime exhibits strong patterns at small spatial resolutions. Further, the works of a number of different scholars in different cities show that microgeographic trends in crime are remarkably consistent between cities [51]. This consistency has led to the development of a physical "law of crime concentration across cities" [49]. The empirical work shown in this section forms the impetus for the development of predictive models of crime: By quantifying and analysing the ways in which crime is arranged into hotspots at a microgeographic scale, allows for a discussion of the ways in which crime can potentially be predictively modelled.

Charting an abridged history of empirical studies on the criminology of places, as early as 1988 scholars began considering crime patterns at the resolution of microplaces. Criminologist Glenn L. Pierce led a project published in 1988 to assess the strategies of policing and crime control available at the time [35]. Pierce et al. made use of the Boston Police Department's (BPD) *Computer Aided Dispatch System* as a dataset to analyse crime occurrence in Boston at the spatial resolution of microplaces. Although Pierce et al. did not consider types of crime (they only measured the number of phone calls to the BPD by address, not the contents of the calls), by surveying how calls to the police were spatially distributed within Boston between the years 1977 to 1982, Pierce et al. found that while 80.79% of all street addresses within Boston produced at least one call to the BPD per-year over the period of study, only 18.32% of all street addresses within Boston produced five or more calls to the police per-year over the period of study [35, p.VI-9]. Even further, they found that just 0.89% of all street addresses within Boston produced fifty or more calls to the police per-year over the period of study [35, p.VI-9]. While it is to be expected that increasingly smaller percentages of overall street addresses should account for larger volumes of calls (it would be ridiculous to assume that a large number of street addresses would produce fifty or more calls to the BPD per-year), that such a small percentage of street addresses produced such a vast number of calls is statistically significant [35]. Put in a different way, over the period of study Pierce et al. found that 18.32% of all street addresses within Boston were responsible for 74.78% of the demand for BPD services, and that just 0.89% of all street addresses within Boston were responsible for 23.56% of the demand for BPD services. The entirety of the results from [35] are enumerated in Figure 2.1. Although this study was quite rudimentary (not considering types of crime nor performing any extensive statistical analysis), from this work alone there is enough cause to consider the ways in which crime is unequally distributed on a microgeographic scale. It is clear from this study that a vast majority of crime in Boston (over the period of study) occurred at a relative minority of street addresses, a finding that many authors would replicate in a number of different contexts in the following years.

Following the study on policing in Boston, a 1989 study by Lawrence Sherman et al. examined crime occurrence at the resolution of street segments in Minneapolis over the span of one year (From December 15, 1985 to December 15, 1986) [43]. In this study Sherman et al. were able to replicate the results of Pierce et al., this time in the context of Minneapolis: Over the year of study, although approximately 60% of street addresses generated at least one call to the Minneapolis Police Department (MPD), a remarkable 3.3% of street addresses generated a majority (50.4%) of the overall calls to the MPD [43]. Of note, is that Sherman et al. were also unable to consider types of crime, as their data was collected in much the same way to that of Pierce et al. before them. The entirety of the results from [43] are enumerated in Figure 2.2. Although their findings were similar to those from Pierce et al., Sherman et al. found that when compared against a normal distribution for the expected number of addresses generating a certain number of phone calls to the MPD, the real-world data showed that a disproportionate majority of places generated very few phone calls to the MPD whereas a disproportionate minority of places generated many phone calls to the MPD. In addition to their statistical analysis, Sherman et al. used their work to motivate the framework of the criminology of places, as discussed previously. Though there were probes into the microgeography of crime, it was with this 1989 study that the study of crime at microplaces became more widely accepted as potentially being able to offer valuable insights into the organization of crime. The work performed by Sherman et al. in Minneapolis solidified the trend that between cities, crime displays significant "hotspot behavior": That relatively small geographic regions within a city seem to fall victim to a majority of overall crimes.

Over the following decades, many more studies considering the spatial distribution of crime within cities would be published. In *Analysing Crime Patters: Frontiers of Practice*, a criminology text dedicated to the exploration of the hotspot behaviour of crime [11], authors Eck et al. examined crime occurrence in the Bronx, a borough of New York City. Agreeing with the work that came before them, Eck et al. discovered that over their period of study (October 1, 1995 to October 15, 1996) crime within the Bronx was arranged into strong hotspots, with a relatively small proportion of places being responsible for a relative majority of overall crime occurrence [11]. Note that within this study Eck et al. make the distinction between robberies, assaults, burglaries, grand larceny, and auto theft: In all cases, significant crime hotspot behaviour can be seen. The results from [11] are enumerated in Figure 2.3.

Table I

The Distribution of the Total Demand for Police Services Among Street Addresses Experiencing a Given Annual Rate of Demand in Boston for the Period 1977 to 1982

Annual Rate of Demand for Services per Street Address	Percent of All Street Addresses	Percent of Total Demand for Police Services
1 or more	80.79	100.00
2 or more	43.64	91.00
3 or more	30.10	84.44
4 or more	22.87	79.18
5 or more	18.32	74.78
10 or more	8.66	59.62
20 or more	3.64	50.13
30 or more	2.01	34.01
40 or more	1.27	27.86
50 or more	.86	23.56
75 or more	.40	16.79
100 or more	.22	13.13
150 or more	.09	9.36

Figure 2.1: Results from Pierce et al. demonstrating the disproportionate demand for the Boston Police Department coming from a minority of street addresses from the years 1977 to 1982 [35, Table VI-I].

No. of Calls	Observed No. of Places	Expected No. of Places	Cumulative % of Places	Cumulative % of Calls
0	45,561	6,854	100%	—
1	35,858	19,328	60.4	100.0
2	11,318	27,253	29.2	88.9
3	5,683	25,618	19.4	81.9
4	3,508	18,060	14.4	76.7
5	2,299	10,186	11.4	72.4
6	1,678	4,787	9.4	68.8
7	1,250	1,929	7.9	65.7
8	963	680	6.8	63.0
9	814	213	6.0	60.6
10	652	60	5.3	58.4
11	506	15	4.7	56.3
12	415	4	4.3	54.6
13	357	1	3.9	53.1
14	297	0	3.6	51.7
15 $\geq$	3,841	0	3.3	50.4

mean = 2.82 $X^2 = 301{,}376$ df = 14 $p < .0001$

Figure 2.2: Results from Sherman et al. demonstrating the disproportionate demand for the Minneapolis Police Department coming from a minority of street addresses over the span of a year [35, Table I].

# of	*All Crimes*		*Robberies*		*Assaults*		*Burglaries*		*Grand Larceny*		*Auto Theft*	
Events	*Places*	*Crimes*	*Places*	*Crimes*	*Places*	*Crimes*	*Places*	*Crimes*	*Places*	*Crimes*	*Places*	*Crimes*
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0.2851	0.5602	0.1703	0.3513	0.1677	0.3441	0.2200	0.4253	0.1010	0.2424	0.1288	0.2878
3	0.1291	0.3683	0.0547	0.1704	0.0477	0.1550	0.1303	0.2115	0.0321	0.1263	0.0391	0.1412
4	0.0696	0.2584	0.0209	0.0913	0.0176	0.0837	0.0525	0.1126	0.0128	0.0774	0.0142	0.0802
5	0.0394	0.1840	0.0110	0.0601	0.0069	0.0500	0.0202	0.0578	0.0070	0.0579	0.0080	0.0597
6	0.0243	0.1377	0.0062	0.0415	0.0031	0.0349	0.0103	0.0368	0.0048	0.0487	0.0052	0.0484
7	0.0163	0.1079	0.0043	0.0324	0.0019	0.0295	0.0059	0.0256	0.0031	0.0401	0.0033	0.0389
8	0.0107	0.0839	0.0028	0.0244	0.0011	0.0253	0.0036	0.0188	0.0029	0.0387	0.0022	0.0327
9	0.0073	0.0673	0.0019	0.0190	0.0008	0.0229	0.0017	0.0123	0.0022	0.0338	0.0013	0.0272
10	0.0053	0.0562	0.0011	0.0130	0.0004	0.0203	0.0004	0.0072	0.0017	0.0301	0.0012	0.0263
11	0.0042	0.0496	0.0010	0.0121	0.0004	0.0203	0.0002	0.0064	0.0017	0.0301	0.0012	0.0263
12	0.0034	0.0437	0.0005	0.0079	0.0004	0.0203	0.0002	0.0064	0.0017	0.0301	0.0011	0.0252
13	0.0028	0.0394	0.0002	0.0056	0.0004	0.0203	0.0002	0.0064	0.0012	0.0253	0.0006	0.0205
14	0.0023	0.0355	0.0001	0.0044	0.0004	0.0203	0.0002	0.0064	0.0012	0.0253	0.0006	0.0205
15	0.0018	0.0310	0.0001	0.0044	0.0004	0.0203	0.0002	0.0064	0.0012	0.0253	0.0005	0.0191
16	0.0016	0.0291	0.0001	0.0044	0.0004	0.0203	0.0002	0.0064	0.0010	0.0222	0.0004	0.0176
17	0.0014	0.0274	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0007	0.0189	0.0004	0.0176
18	0.0011	0.0245	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0007	0.0189	0.0004	0.0176
19	0.0009	0.0218	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0005	0.0153	0.0004	0.0176
20	0.0006	0.0190	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0005	0.0153	0.0004	0.0176
21	0.0006	0.0181	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0005	0.0153	0.0004	0.0176
22	0.0005	0.0172	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0002	0.0110	0.0004	0.0176
23	0.0004	0.0163	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0002	0.0110	0.0004	0.0176
>23	0.0003	0.0153	0.0001	0.0044	0.0002	0.0178	0.0002	0.0064	0.0002	0.0110	0.0002	0.0154
Total	28,848	46,891	8,213	10,504	5,236	6,644	9,095	12,343	4,137	4,909	8,293	10,145

Figure 2.3: Results from Eck et al. demonstrating the disproportionate demand for the New York Police Department coming from a minority of street addresses in the Bronx over the span of a year [11, Table 5.1].

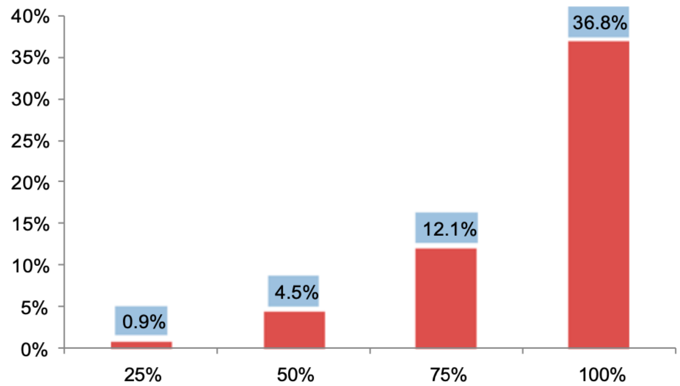

Figure 2.4: Results from Weisburd and Amram demonstrating the proportion of street addresses in Tel Aviv-Jaffa that account for %25, %50, %75 and %100 of crime, respectively [50, Figure 1].

Following these studies performed in the United States, further probes on crime hotspots considered cities outside of the United States: A 2014 study by David Weisburd and Shai Amram examined crime distribution in Tel Aviv-Jaffa, Israel. Perhaps unsurprisingly, Weisburd and Amram found that over the year 2010, approximately 4.5% of street addresses produced over 50% of the overall crime within Tel Aviv-Jaffa [50]. These results parallel the findings of previous authors, and contribute to the idea that regardless of context, crime is very distinctly arranged into tight spatial hotspots. The results from [50] are enumerated in Figure 2.4.

Although nowhere near exhaustive, the studies collected here all show a remarkably similar trend: That a majority of crime appears to occur at a minority of street addresses, regardless of city. Building on this foundation, in a 2015 article, criminologist David Weisburd examined eight cities in an effort to develop a more rigorous physical law governing the arrangement of overall crime within cities [49]. Of the eight cities considered by Weisburd, five of which (Sacramento, Seattle, New York City, Tel Aviv-Yafo, and Cincinanati) have populations ranging from approximately 300,000 people to 8,000,000 million people, and are considered by Weisburd to be “large cities”. The remaining three cities (Brooklyn Park, Redlands, and Ventura) have populations ranging from approximately 70,000 people to 100,000 people, and are considered by Weisburd to be "small cities". Weisburd notes that in the case of the five larger cities, crime is distributed unevenly as previously discussed, with a majority of crimes occurring at a tight "bandwidth" of street segments. On examining these five cities, Weisburd finds that 50% of all crime incidents within each larger city occurs at anywhere between 4.2% (Sacramento) and 6.0% (Cincinnati) of all street segments. Further, Weisburd finds that 25% of all crime incidents within each larger city occurs at anywhere between 0.8% (Sacramento) and 1.6% (Cincinnati) of all street segments. There is a remarkable observed consistency between these five large cities in terms of overall crime distribution: A given percentage of crime incidents (say, 50% or 25%) occur at a consistently small percentahe of overall street addresses, between multiple large cities. The results pertaining to these five larger cities are enumerated in Figure 2.5. In the case of the three smaller cities Weisburd finds that there is an almost identical pattern, with the only difference being that crimes are even more highly concentrated at a minority of street addresses. Weisburd finds that 50% of all crime incidents within each smaller city occur at approximately 3% of all street segments, and that 25% of all crime incidents within each smaller city occurs at approximately 0.5% of all street segments. Though there is a noticeable difference in the severity of crime concentrations between large and small cities as discussed by Weisburd, the conclusion for both cases is the same. The results pertaining to these three smaller cities are enumerated in Figure 2.6.

From both the wealth of research on the spatial distribution of crime performed before him and from his extensive 2015 study, David Weisburd proposes a "physical law of crime concentration" [49]. Moving further than simply noticing that crime displays simple hotspot behaviour, Weisburd proposes that "for a defined measure of crime at a specific microgeographic unit, the concentration of crime will fall within a narrow bandwidth of percentages for a defined cumulative proportion of crime" [49]. Unpacking this statement, Weisburd is claiming that not only are a majority of crimes concentrated at a minority of microgeographic units, but that regardless of context, the percentage of microgeographic units at

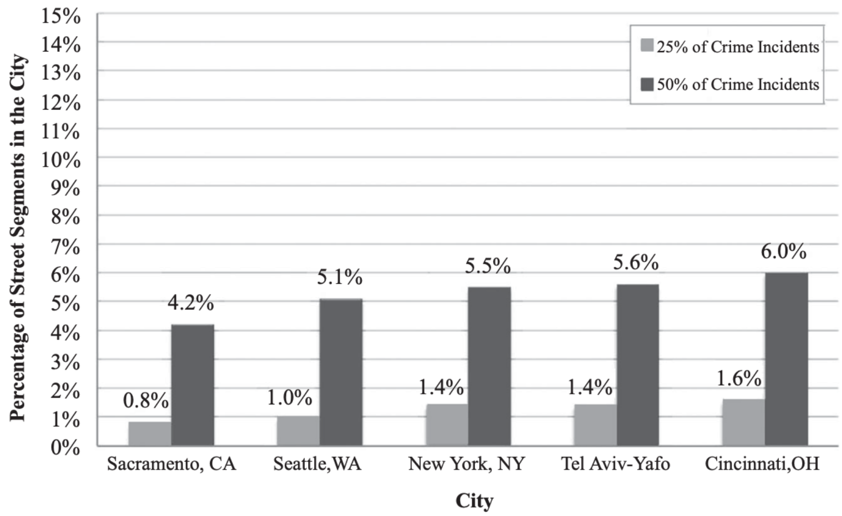

Figure 2.5: Results from Weisburd demonstrating the disproportionately small percentage of street addresses that generate %50 and %25 of all crime incidents, in five large cities [49, Figure 3].

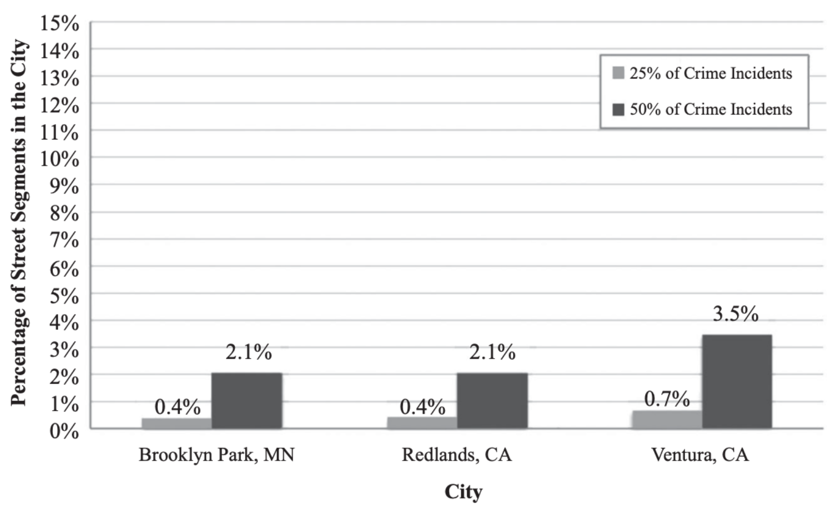

Figure 2.6: Results from Weisburd demonstrating the disproportionately small percentage of street addresses that generate %50 and %25 of all crime incidents, in three small cities [49, Figure 4].

which a certain overall amount of crime occurs, is consistent. Putting this law in context of the preceding results, regardless of the city examined, approximately 50% of crime occurs at anywhere between 2.1% to 6.0% of all street segments and approximately 25% of crime occurs at anywhere between 0.4% to 1.6% of all street segments. Through his work, Weisburd has found that the "bandwidth of percentages for a defined cumulative proportion of crime" representing the concentration of crime is quite small, a range of about 4% when considering 50% of all crimes and a range of less than 1.5% when considering 25% of all crimes [49].

Finally, it is important when building a law of crime concentration, to consider how crime is distributed not only spatially, but through time. David Weisburd was one of the first to consider this question, asking whether the tight hotspot behavior seen in crime distribution is consistent over time, or whether crime hotspots can grow and shrink in intensity over time. From the cities previously discussed, Weisburd was able to obtain sufficient microgeographic data over a period of many years for Tel Aviv-Yafo, Seattle, Brooklyn Park, and New York City. Examining Weisburd's data, it is first abundantly clear that the overall number of crime incidents in each city varies radically between years. Further, there is no discernible global trend in the overall number of crime incidents between years in each city. That is, when comparing overall crime levels per-year between each of the cities studied by Weisburd, there are no notable patterns in overall crime levels shared between any of the cities. *Of interest however, is that the percentage of street segments at which both 50% and 25% of crime occurs remains consistent over time, between all cities.* Although these findings are only preliminary, Weisburd comments that the stability of this percentage lends credence to his proposed law of crime concentration, going as far to argue that that there is sufficient evidence to consider this law as a scientific principle [49]. The results from this study on the consistency of crime hotspots are enumerated in Figure 2.7.

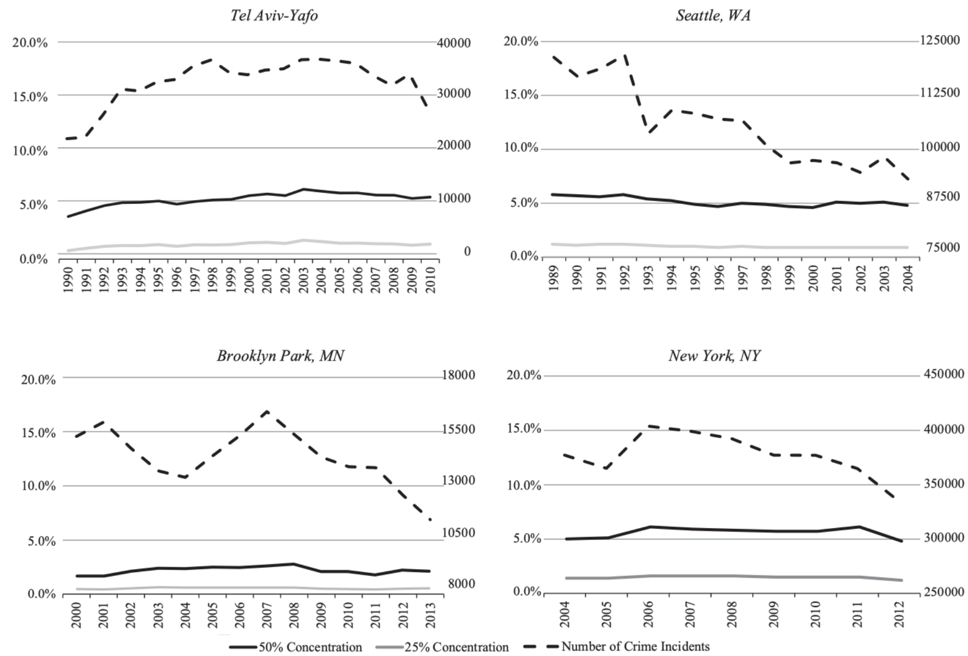

Figure 2.7: Results from Weisburd examining the variability in overall crime levels and in crime concentrations over time [49, Figure 5].

2.3 Explanations for the hotspot behaviour of crime and the possibility to predict crime hotspots

It is made clear by examining the empirical evidence that crime, regardless of type or context, is arranged into tight spatial hotspots. This is a good motivation for developing a predictive model of crime: The fact that crime displays such consistent patterns in space and time lend themselves to being predicted and replicated using mathematical models. To create these models however, requires a baseline understanding of the theoretical criminological frameworks designed to explain the observed hotspot behaviour of crime. Any mathematical model designed to "predict" criminal behaviour must not only make use of empirical data, but must also align with frameworks that explain how crime is spatially distributed. Although this is by no means an exhaustive review of research explaining the spatial behaviour of crime, understanding basic principles will allow for a much richer discussion of the development of the mathematical models for the prediction of crime.

Beginning by speaking more generally, there are two leading models to explain why arbitrary types of crime are generally arranged into tight spatial hotspots: The temporal cognition model (sometimes called the state-dependence model) and the population heterogeneity model (sometimes called the flag model) [51].

The basic principle motivating the temporal cognition model is that following a successful crime at a given place, a criminal is more likely to return to the same place and re-offend, a process known as *repeat-victimization* [20]. This idea generally follows from *rational choice theory*, a fundamental framework to understand criminal behaviour. Although not without its detractors, rational choice theory states, in simple terms, that a criminal will weigh the benefits and consequences of a potential crime, eventually making a rational and informed decision about whether they will or will not commit a crime [7]. If, for a person about to engage in a potential criminal act, the perceived benefits of a success outweigh the perceived consequences of a failure, rational choice theory states that the person will behave rationally and will follow through with the criminal act. Relating this framework back to the idea of the temporal cognition model of crime and repeat-victimization, suppose a person commits a successful residential burglary. In committing this crime, the person will learn a great deal of information about the layout, weaknesses, etc., of the residence where they committed their crime. Due to both the information gain and the boost in confidence following their successful criminal act, the perceived chance of encountering consequences in subsequent burglaries at the same residential address will decrease in the mind of the criminal. What results is that at this address, in the mind of the criminal, the perceived benefits remain the same where the perceived chance of encountering consequences decreases. Rational choice theory thus states that the criminal is significantly more likely to re-victimize this same address, motivating the process of repeat-victimization.

Using the framework of rational choice theory, the idea of repeat-victimization can be easily extended to that of *near-repeat-victimization*. Supposing again that a person commits a successful residential burglary, they will not only learn about the exact residence in which they burgled, but about the surrounding residences, and about the qualities of the neighbourhood as a whole. Thus, the perceived chance of encountering consequences in subsequent burglaries at neighbouring addresses will decrease in the mind of the criminal, leading to increased rates of victimization at nearby addresses [19]. Of importance, is that the temporal cognition model to motivate crime hotspots assumes nothing about the underlying addresses being victimized: The only assumptions being made are that potential criminals engage in rational choice theory. The temporal cognition model posits in fact that there is a "beginning state" of criminal attractiveness, where all addresses are equally as attractive to potential criminals. It is only after a number of seemingly "random" criminal acts have occurred, that certain places become hotspots for crime. The complete lack of reliance on geographic information in the temporal cognition model lends itself well to the fact that crime hotspots seem to form regardless of context, as seen in the preceding section. This framework to understand crime hotspots states that regardless of context, hotspots of crime are destined to form due to the underlying psychology of criminals.

For decades rational choice theory has been well accepted within criminology, serving as a strong framework with which to understand the hotspot behaviour of criminal activity. As mentioned previously however, this is not the only means with which we are able to justify hotspots of crime: The contrasting yet complimentary population heterogeneity model serves to justify crime hotspots not through the psychology of individual criminals, but through the geographic differences in the crime places themselves. The basic principle motivating the population heterogeneity model is that (in stark contrast to the temporal cognition model) there is no such "beginning state" where all places in a given environment are equally attractive to potential criminals. To the contrary, this model states that all places in a given environment have differing levels of criminal attractiveness due to potentially unquantifiable internal and external geographic factors. The population heterogeneity model states that factors such as the amount of neighbouring green space, average household income, style of architecture, distance to certain amenities, etc., all play a role in determining the criminal attractiveness of an individual address. According to this model, potential criminals encountering places of low criminal attractiveness are less likely to commit crimes, whereas potential criminals encountering places of higher criminal attractiveness are thus more likely to commit crimes.

Discussing these two general models to explain crime hotspots, David Weisburd comments that these two frameworks are in no way mutually exclusive, and in reality quite probably work in tandem to inform the developments of crime hotspots. Boiling down the insights from these two models, we can identify two major factors that will influence a given place's criminal attractiveness and thus the possibility of a crime hotspot forming: The number of successful crimes that have occurred recently at said place, and the baseline criminal attractiveness of said place. These two factors can and have been leveraged, to form the foundation for a mathematical model to predict crimes, to be discussed in the following sections.

Chapter 3

A Discrete Mathematical Model of Criminal Behaviour

In their 2008 paper "A Statistical Model of Criminal Behavior" [44], authors Martin B. Short et al. introduced a continuous mathematical model of residential burglary, building on recent advancements in the study of the criminology of places. This model uses a system of reaction-diffusion partial differential equations (PDEs) to model how both criminal attractiveness and criminal density evolve over time in idealized urban environments. The PDE model introduced by Short et al. has had deep implications not only within the study of environmental criminology, but within the day-to-day operations of many police departments: A number of technology start-ups took this model and implemented it into commercial software that police departments today are currently using to influence their decisions on what places within their jurisdictions to target [27, 38]. Though there have been critiques of regimes of proactive policing and predictive policing software more generally (see both Chapter 6 and [22, 40, 1] to name a very small few critical articles), there has been a relative lack of critical work examining the base model introduced by Short et al. in [44]. A critical examination of the model of criminal behaviour as introduced by Short et al. is the principal goal of this project, and the following three chapters will follow the development of and perform numerical trials with this model.

Before diving into a discussion of the continuous model of residential burglary introduced by Short et al. in [44], it is important to provide a detailed discussion of the origins of this model. Building on the criminological work from the previous chapter, we begin by describing how both the temporal cognition and population heterogeneity models of criminal behaviour discussed previously can be leveraged to create a discrete model of criminal behaviour. We then summarize the development of the discrete model of residential burglary provided by Short et al., in [44]. Following this, we provide the results from our own numerical simulations in Python using this discrete model. A discussion of the intricacies of this discrete model will allow for a much richer discussion of the continuous model of residential burglary in the following chapter.

3.1 Motivating a Discrete Model of Criminal Behaviour

Recall from the previous chapter that there is extensive empirical evidence indicating that regardless of type, crime appears to be arranged into tight spatial hotspots (see [35, 43, 11, 50, 49]). The discrete model of residential burglary proposed by Short et al. seeks to replicate this behaviour, building on criminological frameworks that explain this hotspot behaviour. Recall from the previous chapter that criminologists have proposed a number of theories attempting to explain why regardless of urban environment, crime is generally very concentrated in space. The two leading models justifying crime hotspots are the temporal cognition model of criminal behaviour and the population heterogeneity model of criminal behaviour [49]. Put in very simple terms, the temporal cognition model of criminal behaviour states that following a successful crime, a criminal is more likely to return to the same location or a nearby location to re-offend. These processes, known as repeat victimization and near-repeat victimization respectively are well accepted within criminology literature and are considered to be one of the fundamental reasons behind crime hotspot formation. The second of the two models to justify crime hotspots, the population heterogeneity model of criminal behaviour, states that fundamentally each potential crime place within a larger environment has a differing level of baseline criminal attractiveness and that potential criminals are far more likely to victimize a site that is relatively more attractive when compared to the other places in an environment. When combined, the temporal cognition and population heterogeneity models of criminal behaviour form the foundation used by Short et al. to build their numerical model of criminal behaviour.

It should be noted, that while the above two models work well as theoretical tools to explain the spatial dynamics of crime, they should only be understood as more general explanations for these dynamics. In reality, when examining specific types of crime there are a number of more complicated factors to take into consideration: As an example, consider the case of motor vehicle thefts. As the potential targets for motor vehicle thefts (the vehicles themselves) are able to move throughout space, it stands to reason that the dynamics underlying this type of crime are more complicated than say, the dynamics underlying residential burglary. Short et al. have taken this into consideration and have opted to restrict their initial model to the case of residential burglary. They comment that "residential burglary [is] in many ways the simplest crime type, since mobile offenders are coupled to stationary target sites, and further complexity arising from the relative movement between the agents at play may be ignored." [44].

3.2 Building a Discrete Model of Criminal Behaviour

The discrete model of residential burglary introduced by Short et al. (as originally described in [44]) is composed of two components: The houses where crimes may occur and a collection of criminals that are able to move freely between neighbouring houses. The model generally allows for the houses to be arranged arbitrarily, though for the sake of simplicity, Short et al. opt for the houses to be arranged on a rectangular grid. Each house has a value representing criminal attractiveness, which is the sum of the underlying attractiveness of the site (see the population heterogeneity model of criminal behaviour discussed previously) with a "dynamic component of criminal attractiveness" which is set according to the number of recent crimes occurring nearby (see the temporal cognition model of criminal behaviour discussed previously). Within this model, criminals are generated at a steady rate across the grid of houses and will be removed from the grid once they have committed a burglary (modelling the fact that criminals will often "lie low" for a period of time after offending). Each criminal has the potential to commit a burglary according to the criminal attractiveness of each site. If a criminal does not commit a burglary, they will move to an adjacent site biased towards neighbouring sites with higher criminal attractivenesses.

Moving now to discuss the discrete model of criminal behaviour in more detail, consider a discrete two-dimensional rectangular lattice with spacing l to represent potential residences. The position of each lattice point (or, site) is given by $s = (i, j)$ and the quantity δt is the discrete time interval over which the model is to be updated to a new state. Note that the choice of l and δt will specify the speed at which criminals move throughout the lattice, and thus the choice of the two quantities must align with real-world behaviour. The overall "criminal attractiveness" of a given lattice point s at a given time t is given by $A_s(t) \in [0, \infty)$. Note that $A_s(t)$ is able to vary over time, to allow for the criminal attractiveness of a given site to vary according to the repeat and near-repeat victimization effects discussed previously. $A_s(t)$ is thus defined to be the sum of two components:

$$A_s(t) = A_s^0 + B_s(t) \tag{3.1}$$

where A_s^0 represents the underlying criminal attractiveness of a given site s (a quantity that is not allowed to vary over time) and $B_s(t)$ represents the dynamic component of criminal attractiveness (a quantity to represent neighbourhood effects on the overall criminal attractiveness).

Throughout the lattice, criminals are generated at each lattice point at a rate Γ. Over each time-step of length δt criminals are able to perform one of two actions: They are able to either burglarize their current site or move to an adjacent site based on the attractiveness of adjacent sites. For a criminal beginning at site s, over the time step $[t, t+\delta t]$ the probability that the criminal will burglarize site s is given by:

$$p_s(t) = 1 - e^{-A_s(t)\cdot\delta t}. \tag{3.2}$$

This probability function is chosen by Short et al. using a standard Poisson process, given that the expected number of burglaries at site s over the time step $[t, t+\delta t]$ is $A_s(t)\cdot\delta t$. Note, that the higher the criminal attractiveness is at a site s, the higher the probability that a criminal will burglarize site s. Further, note that the longer the discrete time-step δt is the greater the chance site s will be burglarized. Upon a successful burglary, the criminal will be immediately removed from the grid. If however, for a criminal at site s, over the time step $[t, t+\delta t]$, no burglary occurs, the criminal will instead move to a random neighbouring lattice point with a bias towards neighbouring sites of higher criminal attractivenesses. Let $\mathcal{N}_s$ denote the set of all sites neighbouring s. Given that the criminal does not burglarize s during the current time step, the probability that the criminal will move to site $n \in \mathcal{N}_s$ is given by:

$$q_{s\to n}(t) = \frac{A_n(t)}{\sum_{i\in\mathcal{N}_s} A_i(t)}. \tag{3.3}$$

With this, the behaviour of criminals throughout the lattice has been entirely specified.

What remains is to discuss how the dynamic component of criminal attractiveness, $B_s(t)$, evolves over time. Recall from previous discussions that, after being victimized, a given location and its neighbours will experience a heightened risk of being re-victimized (this is the repeat and near-repeat victimization behaviour discussed in [49]). Put differently, we would expect that for a site s, after being victimized, its overall criminal attractiveness should increase. To first model the repeat victimization behaviour, Short et al. begin by introducing the quantity ω to represent the rate at which the dynamic component of criminal attractiveness decays. Further, the quantity θ is introduced to represent the amount by which an individual burglary will increase the criminal attractiveness of a site and $E_s(t)$ is introduced to represent the number of burglaries occurring at site s over the time step $[t, t + \delta t]$. $B_s(t)$ is updated between time steps according to:

$$B_s(t + \delta t) = B_s(t)(1 - \omega \cdot \delta t) + \theta E_s(t). \tag{3.4}$$

Describing equation (3.4) in words, between time-steps the dynamic component of criminal attractiveness will experience both some decay proportional to ω and δt and some growth due to the recent number of burglaries occurring at site s ($E_s(t)$). Equation (3.4) accounts for repeat victimization effects at an individual site s, though does not take into account the effect that burglaries at sites neighbouring s have on the criminal attractiveness of site s (near-repeat victimization effects). To account for this, Short et al. introduce the quantity η to represent the significance of neighbourhood effects on the dynamic component of criminal attractiveness of site s. Equation (3.4) can then be modified to account for near-repeat victimization effects as follows:

$$B_s(t + \delta t) = \left[(1 - \eta)B_s(t) + \frac{\eta}{z}\sum_{i \in \mathcal{N}_s} B_i(t)\right](1 - \omega \cdot \delta t) + \theta E_s(t), \tag{3.5}$$

where z is simply the number of sites that neighbour s (put differently, $z = |\mathcal{N}_s|$). Describing equation (3.5) in words, it is functionally identical to equation (3.4) but instead of the dynamic component of criminal attractiveness at site s decaying based on its value during the previous time-step, it is instead modified based on a weighted average of each of the dynamic components of criminal attractiveness from all sites neighbouring s. If the average

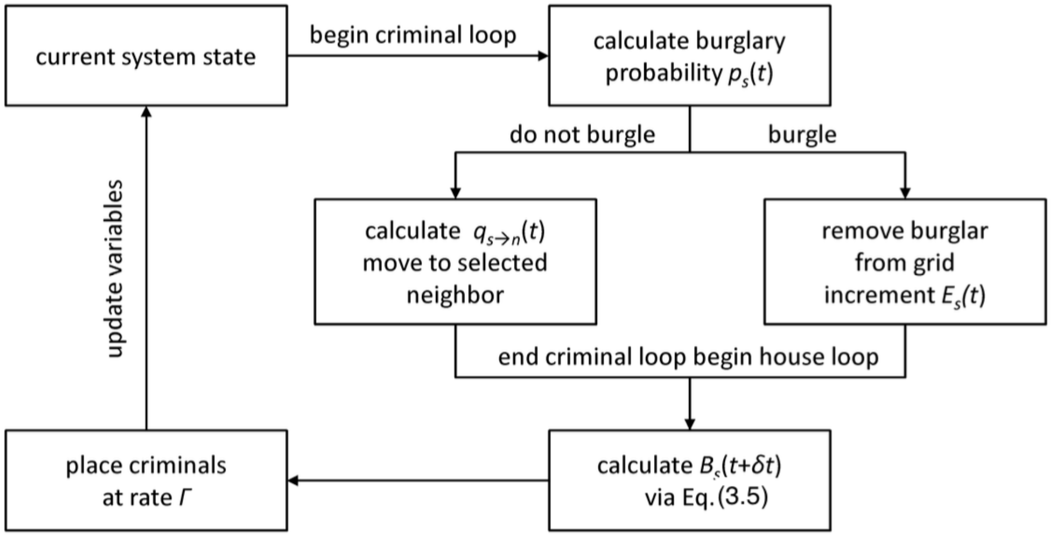

Figure 3.1: Flowchart from Short et al. representing a single time step from the discrete model of residential burglary, modified slightly to reference the correct equations in this document [44, Figure 2]

.

values of the dynamic components of criminal attractiveness over each site neighbouring s are actually greater than the dynamic component of criminal attractiveness at site s, then $B_s(t + \delta t)$ will increased based on the higher attractivenesses at neighbouring sites. At the end of each time step the dynamic component of criminal attractiveness is updated according to the above rule to simulate the process of repeat *and* near-repeat victimization.

A summary flowchart of the discrete model adapted from Short et al. is provided in Figure 3.1. In addition, a table of all free parameters of the discrete model along with their potential ranges is included in Table 3.1.

Though the discrete model of residential burglary was ultimately designed by Short et al. to derive their continuous model of residential burglary, there are still benefits to analysing the discrete model in isolation. A number of simulations of residential burglary are included in [44], though I opt here to perform a more extensive number of simulations using a wider variety of parameters. My Python code implementing this discrete model is included in

Param.	Desc.	Range
l	Distance between adjacent lattice points	$(0, \infty)$
δt	Duration of one time step	$(0, \infty)$
A_s^0	Baseline crim. attractiveness	$M_{m\times n}(\mathbb{R}^+)$
Γ	Rate of generation of criminals at each lattice point	$[0, \infty)$
ω	Decay rate of dynamic component of crim. attractiveness	$[0, \infty)$
θ	Crim. attractiveness increase amt. due to one burglary	$[0, \infty)$
η	Significance of neighbourhood effects on crim. attractiveness	$[0, 1]$

Table 3.1: All free parameters in the discrete model of residential burglary along with descriptions and available ranges.

Appendix A, and results from my simulations will be discussed in the following section.

3.3 Results from Simulations Using the Discrete Model of Criminal Behaviour

Using my implementation of Short et al.'s discrete model of residential burglary [44] included in Appendix A, I performed four large-scale simulations with different parameter sets. For all four trials I used a 100×100 rectangular lattice, with $l = 1$ and $\delta t = 0.01$. Each simulation also set $A_s^0 = 1/30$ for all lattice points s and set $\omega = 1/15$. Following the work done by Short et al. in [44], the only parameters allowed to vary between trials were η (the significance of neighbourhood effects), Γ (the rate of criminal generation), and θ (the amount to increase a site's criminal attractiveness due to one burglary). The choices of parameters are taken from the trials performed in [44] and are shown in Table 3.2.

Each of the four simulations using the discrete model of residential burglary was run for 1000 time steps, with no criminals initially present on the grid. Snapshots from each of the four trials are included in Figures 3.2, 3.3, 3.4, and 3.5 respectively. In each of the four figures, criminal attractiveness is plotted as a heat-map: Black regions represent areas where the dynamic component of criminal attractiveness is equal to zero (that is to say, the

Trial	η	Γ	θ
A	0.2	0.019	0.56
B	0.2	0.002	5.6
C	0.03	0.019	0.56
D	0.03	0.002	5.6

Table 3.2: Parameter choices obtained from [44] for each of the four simulations using the discrete model of residential burglary as shown in Figures 3.2 through 3.5.

criminal attractiveness at such a site is only composed of the static component, $A_0 = 1/30$). In contrast, lighter regions represent areas of higher criminal attractiveness. We will discuss these results in detail in Chapter 5.

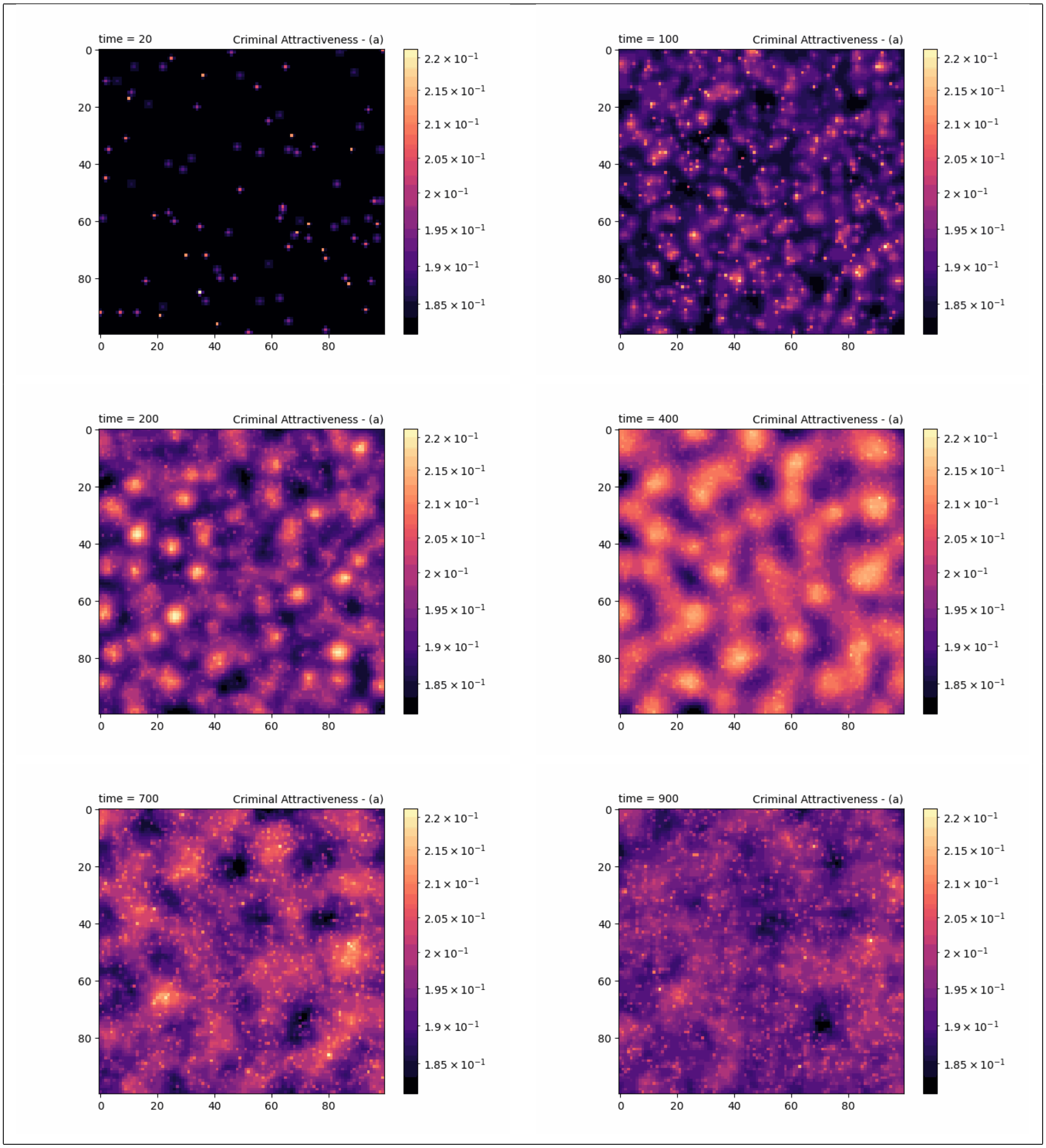

Figure 3.2: Trial A simulation results from the discrete model of residential burglary [44], using the parameters given for Trial A in Table 3.2.

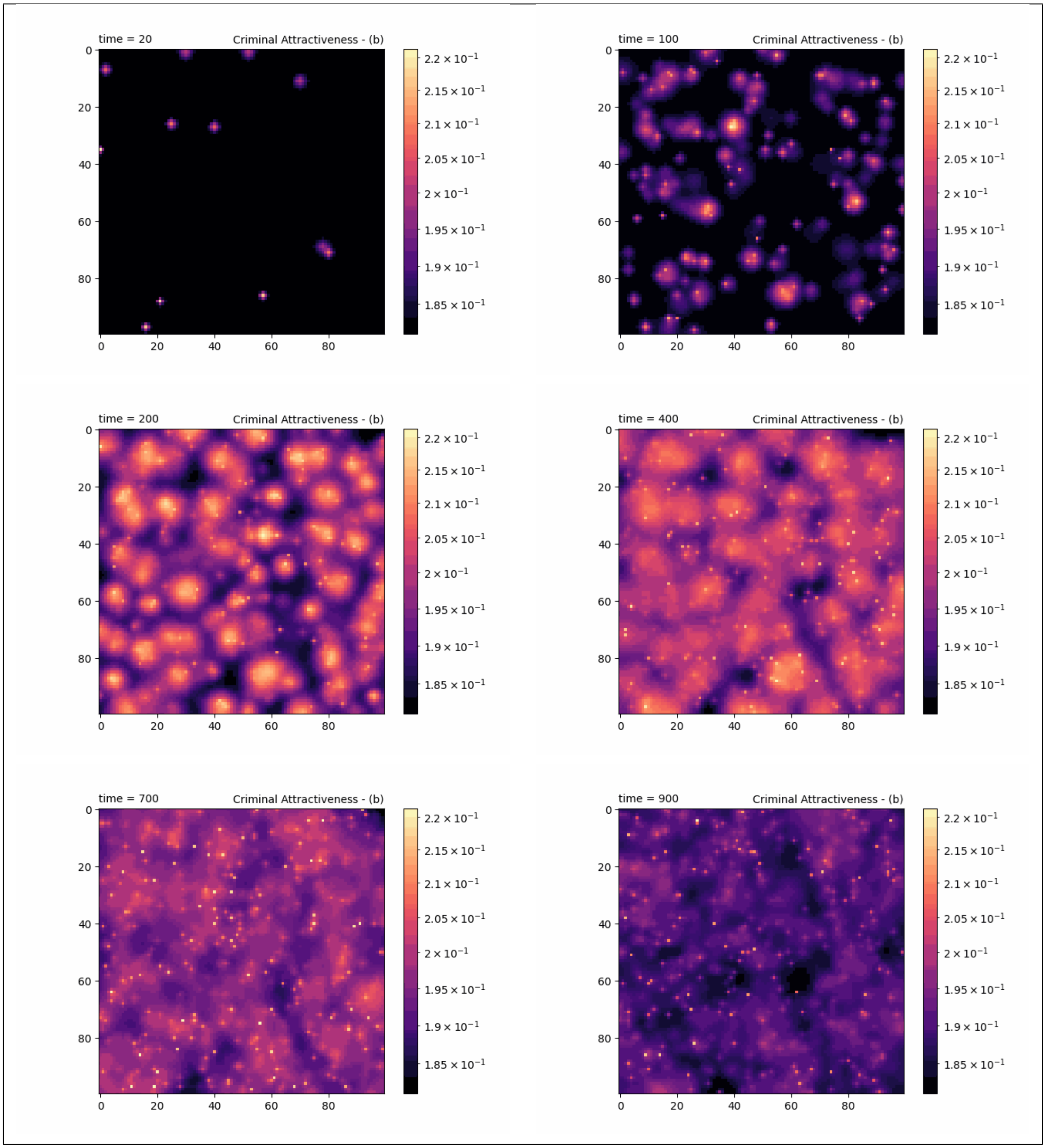

Figure 3.3: Trial B simulation results from the discrete model of residential burglary [44], using the parameters given for Trial B in Table 3.2.

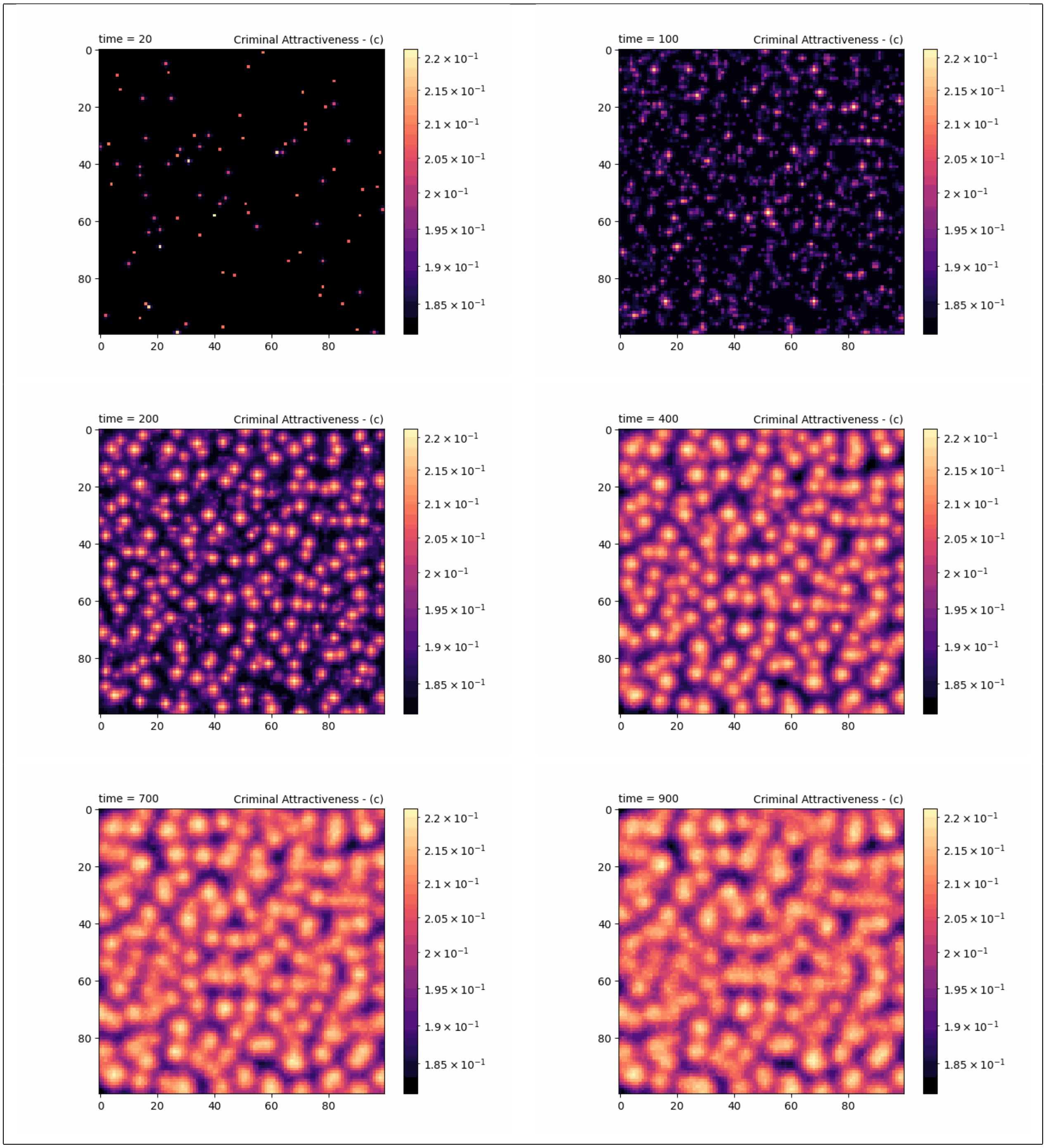

Figure 3.4: Trial C simulation results from the discrete model of residential burglary [44], using the parameters given for Trial C in Table 3.2.

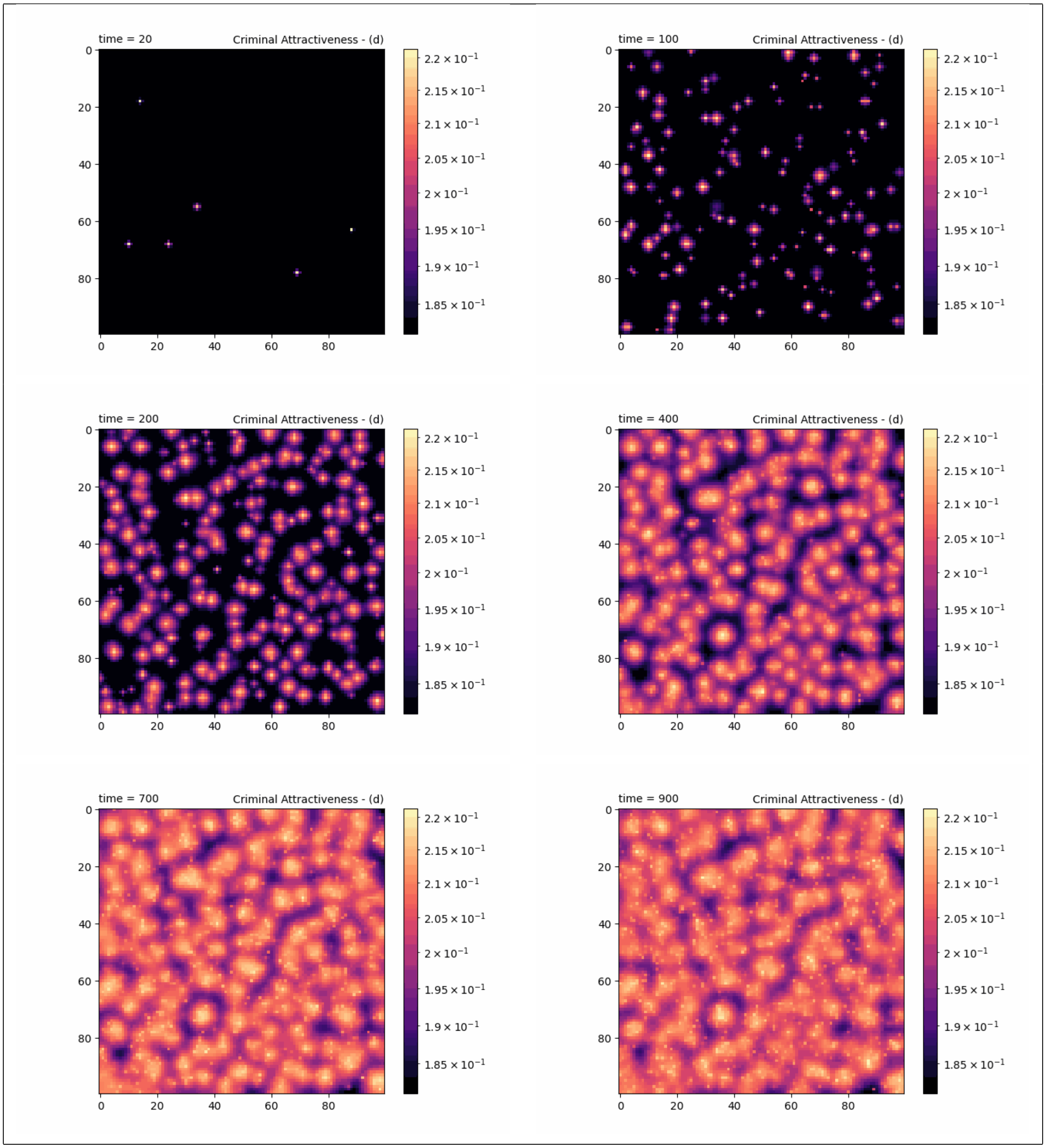

Figure 3.5: Trial D simulation results from the discrete model of residential burglary [44], using the parameters given for Trial D in Table 3.2.

Chapter 4

A Continuous Mathematical Model of Criminal Behaviour

4.1 Partial Differential Equations and Their Solutions

At a high level, a differential equation is simply an equation involving a function and its derivatives with respect to one or more independent variables. If a differential equation relates a function in one variable with its derivatives, then it is known as an *ordinary differential equation* (ODE). Further, the order of an ODE is simply the order of the highest derivative included in the ODE. Let $y : \mathbb{R} \to \mathbb{R}$ be dependant on x and C^n continuous, and let $y^{(k)}$ denote the k-th derivative of y. The general form of an n-th order ODE (sometimes called the implicit form of an n-th order ODE) is:

$$f(x, y, y', y'', \ldots, y^{(n)}) = 0, \tag{4.1}$$

where f is an arbitrary function f. As a more concrete example, consider the following first-order ODE:

$$y'(x) - \cos(x) \cdot y(x) = 0, \quad x \in \mathbb{R}. \tag{4.2}$$

It can be easily verified that the solution to the above ODE is $y(x) = e^{sin(x)} + C$ for an arbitrary $C \in \mathbb{R}$. Note that there is an infinite family of solutions to this ODE, parametrized by the constant C. When coupled with an *initial condition* of the form $y(x_0) = C$ for some $x_0 \in \mathbb{R}$, a first-order ODE, under appropriate assumptions, will have precisely one exact solution [14]. The combination of an ODE together with an initial condition is often called an *initial value problem* (IVP or IVODE). Consider again the same first-order ODE, this time with an initial condition:

$$y'(x) - \cos(x) \cdot y(x) = 0, \quad y(0) = 2, \quad x \in \mathbb{R}. \tag{4.3}$$

As before, the solution to this ODE is $y(x) = e^{sin(x)} + C$, but using the initial condition one can substitute $x = 0$ into this exact solution to obtain $C = 1$. Hence, the unique exact solution to the above IVP is $y(x) = e^{sin(x)} + 1$.

Continuing, note that for a general n-th order ODE there will be an infinite family of solutions, parametrized by n constants $\{C_1, \ldots C_n\}$. In this more general case there are several ways to assert a unique exact solution. Two notable examples being either by specifying a set of n initial conditions at the same $x_0 \in \mathbb{R}$ of the form $\{y(x_0) = C_1, y'(x_0) = C_2, \ldots, y^{(n-1)}(x_0) = C_n\}$ (resulting again in an IVP), or by specifying n conditions at different points over the problem domain of the form $\{y(x_0) = C_1, y(x_1) = C_2, \ldots, y(x_{n-1}) = C_n\}$, resulting in what is known as a *multi-point value problem* (BVP, or BVODE). While we do not include a discussion of the techniques to compute exact solutions to both IVPs and BVPs here, of importance is the idea that n additional conditions must be imposed on an n-th order ODE to assert an exact solution

If a differential equation relates a function in multiple variables with its partial derivatives, then said equation is known as a *partial differential equation.* As with ODEs, the order of a PDE is simply the order of the highest partial derivative included in the PDE. Often, PDEs that model real-world phenomena consist of a function with only two independent variables (normally one variable representing time and one representing space, or two variables representing space) with its partial derivatives. Let $u(x,t) : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ be C^2-continuous in both x and t. Further, let $u_x(x,t)$ and $u_{xx}(x,t)$ denote the first and second derivatives of u with respect to x, and let $u_t(x,t)$ and $u_{tt}(x,t)$ denote the first and second derivatives of u with respect to t. The general form (sometimes called the implicit form) for a second-order PDE in one spatial variable and one time variable is as follows:

$$f\left(x, t, u(x,t), u_x(x,t), u_t(x,t), u_{xx}(x,t), u_{tt}(x,t)\right) = 0, \tag{4.4}$$

where f is an arbitrary function. While a general theory concerning n-th order PDEs is outside the scope of this project, it should be noted that the behaviour of PDEs can be roughly categorized into three separate categories [14]:

1. *Hyperbolic PDEs* are those that are time-dependant and do not have solutions evolving towards a steady-state.

2. *Parabolic PDEs* are those that are time-dependant and do have solutions evolving towards a steady-state.

3. *Elliptic PDEs* are those that are time-independent, instead depending on only spatial variables.

For an arbitrary higher order PDEs, it should be noted that the PDE may exhibit different types of behaviour at a number of different regions throughout the problem domain. Regardless, the above terminology allows for a richer discussion of the solution behaviour of a number of PDEs.

As is the case with ODEs, a PDE by itself will generally have an infinite family of solutions parametrized by a number of constants. Depending on the general behaviour of a PDE, different combinations of initial and boundary conditions are required to assert the

Equation	Type	Examples of Req. Conditions
(4.5)	Hyperbolic	I.C. $u(x,0)$ and left B.C. $b_l(a,t)$
(4.6)	Parabolic	I.C. $u(x,0)$, left B.C. $b_l(a,t)$, and right B.C $b_r(b,t)$
(4.7)	Elliptic	Four B.C.s: $u(a,y)$, $u(b,y)$, $u(x,c)$, $u(x,d)$

Table 4.1: Different types of PDEs, and the additional initial conditions (ICs) and boundary conditions (BCs) required to assert unique solutions.

existence of a unique exact solution. Note that in contrast to the ODE case, for a PDE both the initial and boundary conditions are continuous functions instead of scalar quantities. Let $a, b, c, d \in \mathbb{R}$ such that $a < b$ and $c < d$. Consider the following three PDEs:

$$u_t(x,t) = u_x(x,t), \quad a \leq x \leq b, \quad t \geq 0 \qquad \text{(The 1D wave equation)} \tag{4.5}$$

$$u_t(x,t) = u_{xx}(x,t), \quad a \leq x \leq b, \quad t \geq 0 \qquad \text{(The 1D heat equation)} \tag{4.6}$$

$$0 = u_{xx}(x,y) + u_{yy}(x,y), \quad a \leq x \leq b, \quad c \leq y \leq d \qquad \text{(The 2D Laplace equation)} \tag{4.7}$$

The above three PDEs (equations (4.5), (4.6), (4.7)) are examples of hyperbolic, parabolic, and elliptic PDEs, respectively. To help describe how initial and boundary conditions must be used in each of these three cases, Table 4.1 describes examples of potential additional constraints that must be placed on each of these three PDEs to assert the existence of unique exact solutions. In addition to this, Figure 4.1 provides a reference for the problem domains for equations (4.5), (4.6), and (4.7).

As a final note, there are a number of different potential boundary conditions that can be applied to a PDE. Consider an arbitrary parabolic PDE in one space and one time dimension. As mentioned previously, to assert the existence of a unique exact solution this PDE requires two boundary conditions (as well as an initial condition). Table 4.2 uses this type of PDE as an example to outline the two most common types of boundary conditions that can be used in conjunction with a PDE to assert a unique exact solution.

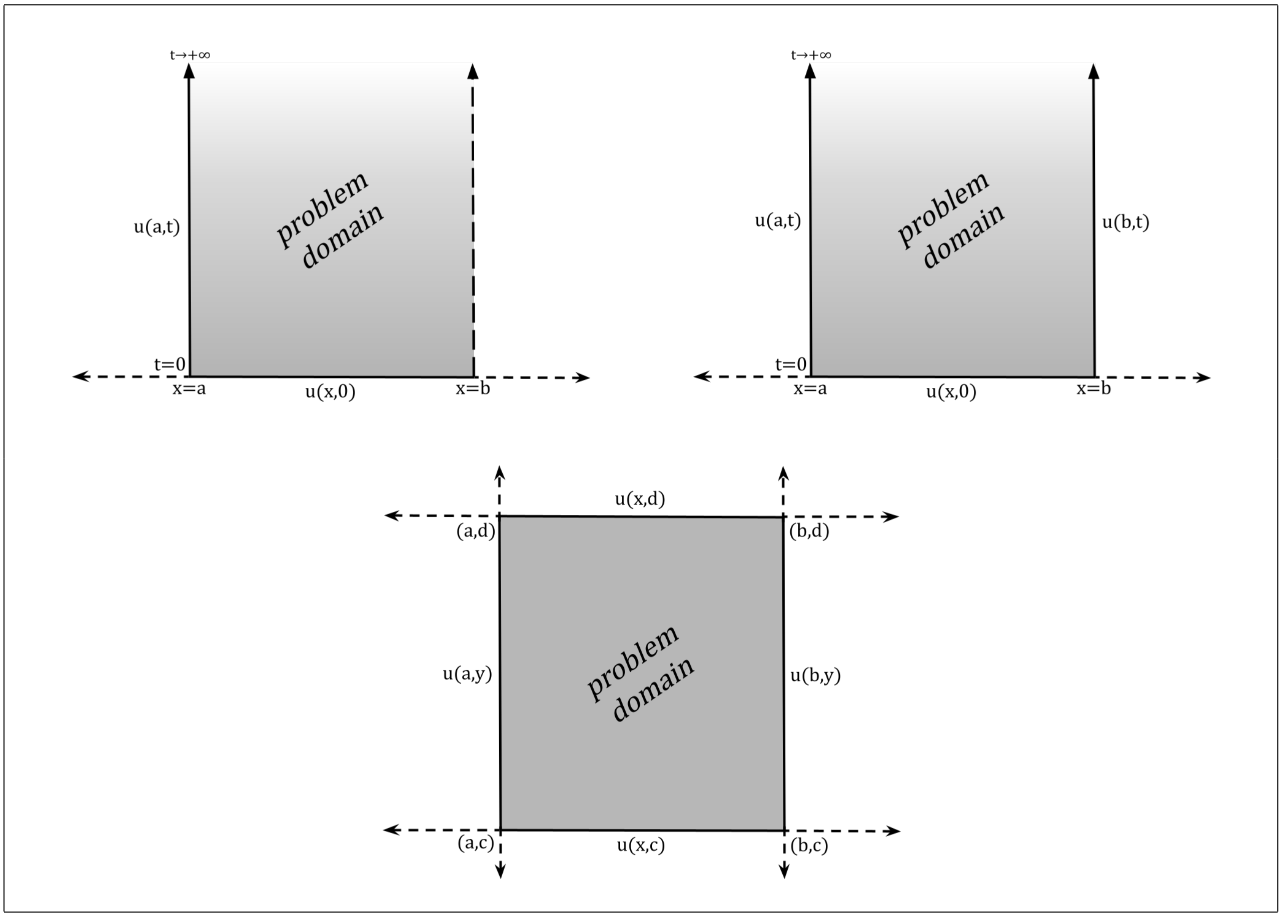

Figure 4.1: Visual depiction of the problem domains of hyperbolic, parabolic, and elliptic PDEs in one spatial dimension and one time dimension. The top left diagram shows a typical hyperbolic PDE problem domain. The top right diagram shows a typical parabolic PDE problem domain. The bottom diagram shows a typical elliptic PDE problem domain. Bolded lines indicate either initial or boundary conditions that must be imposed on the PDE to assert a unique exact solution.

Type of B.C.	General Form	Note
Dirichlet	$u(a,t) = b_l(a,t)$, $u(b,t) = b_r(b,t)$	Boundary of u specified explicitly
Neumann	$u_t(a,t) = b_l(a,t)$, $u_t(b,t) = b_r(b,t)$	Boundary of u_t specified explicitly

Table 4.2: The two most common types of boundary conditions that can be imposed on PDEs, with their general forms.

There are a number of other potential types of boundary conditions that can be imposed on a PDE, though these two are by far the most common when considering PDEs to model real-world phenomena. Very often as well, "no-flux Neumann boundary conditions" are considered, where the value of the PDE along the given boundary is uniformly zero. In terms of the notation included in Table 4.2, no-flux Neumann boundary conditions would take the form $u_x(a,t) = u_x(b,t) = 0$. Ultimately, the continuous model of residential burglary to be discussed in the following section will be based on two parabolic PDEs with no-flux Neumann boundary conditions.

4.2 Deriving a Continuous Model of Criminal Behaviour

As discussed in the previous section, differential equations lend themselves well to the mathematical modelling of continuous real-world phenomena. Turning back to the discrete model of residential burglary as introduced by Short et al. in [44], there are a number of key theoretical problems with this discrete model that can be addressed using a continuous model based on differential equations. Foremost, while crime occurrence is discrete in nature, criminal density and criminal attractiveness are better represented by functions that are continuous throughout space. The discrete model of residential burglary only considers criminal density and attractiveness at its discrete lattice points, a problem that is addressed by modelling residential burglary using PDEs. As well, we would prefer that a model of residential burglary to be deterministic in nature, not relying on weighted random walks as the discrete model does. Building a continuous model of residential burglary allows for an alternative and generally more robust model.

A number of modifications must be made to Short et al.'s discrete model of residential burglary to obtain a continuous model of residential burglary. The final continuous model is a system of two partial differential equations, one concerning how the dynamic component of criminal attractiveness evolves over time and space and the other concerning how criminal density evolves over time and space. I begin here with the simpler derivation as originally presented in [44], to obtain the PDE concerning the dynamic component of criminal attractiveness. To begin, recall that within the discrete model for a given site s, the dynamic component of criminal attractiveness at s is updated between time-steps according to the following rule:

$$B_s(t+\delta t) = \left[(1-\eta)B_s(t) + \frac{\eta}{z}\sum_{i\in\mathcal{N}_s} B_i(t)\right](1-\omega\cdot\delta t) + \theta E_s(t), \tag{4.8}$$

where we recall that $B_s(t)$ is the dynamic component of criminal attractiveness at site s at time t, $E_s(t)$ is the number of criminal events occurring at site s over the time-step beginning at time t, $\mathcal{N}_s$ is the set of all sites neighbouring s, δt is the time step, and η, z, ω, θ are all parameters (originally introduced in Table 3.1). Essentially, this equation governs how to update the dynamic component of criminal attractiveness at a given site s while simulating criminal behaviour over the time-step beginning at time t. The first step in deriving a PDE representing the dynamic component of criminal attractiveness is to modify the above equation to represent the *expected* value of $B_s(t)$ after one time-step. To do so, simply substitute $E_s(t) = n_s(t)p_s(t)$ as follows:

$$B_s(t+\delta t) = \left[(1-\eta)B_s(t) + \frac{\eta}{z}\sum_{i\in\mathcal{N}_s} B_i(t)\right](1-\omega\delta t) + \theta n_s(t)p_s(t), \tag{4.9}$$

where $n_s(t)$ is the number of criminals at site s at time t and $p_s(t)$ is the probability that given a criminal is at site s, they will burglarize s at time t. This new formula is further modified by converting $n_s(t)$ from an absolute measure to a density function, $\rho_s(t) = n_s(t)/l^2$ (recall, l is the spacing between adjacent lattice points). This leads to a modification of the above formula as follows:

$$B_s(t+\delta t) = \left[(1-\eta)B_s(t) + \frac{\eta}{z}\sum_{i\in\mathcal{N}_s} B_i(t)\right](1-\omega\delta t) + \theta l^2 \rho_s(t) p_s(t). \tag{4.10}$$

Next, consider the discrete spatial Laplacian of $B_s(t)$, defined as:

$$\Delta B_s(t) = \frac{\sum_{i\in\mathcal{N}_s}\left(B_i(t) - B_s(t)\right)}{l^2}. \tag{4.11}$$

The discrete spatial Laplacian is useful in this context as taking the limit of $\Delta B_s(t)$ as the mesh spacing l tends to zero results in the continuous Laplacian of $B_s(t)$. Put differently, for $s = (x, y)$, let $B(x, y, t) = \lim_{l,\delta t\to 0}\{B_s(t)\}$, and proceed as follows:

$$\lim_{l,\delta t\to 0} \Delta B_s(t) = \nabla^2 B = \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2}. \tag{4.12}$$

Using the discrete spatial Laplacian, equation (4.10) can be modified as follows (recalling that $\mathcal{N}_s$ denotes the set of sites that directly neighbour s and that z indicates the number of sites neighbouring s):

$$\begin{aligned}
B_s(t+\delta t) &= \left[(1-\eta)B_s(t) + \frac{\eta}{z}\sum_{i\in\mathcal{N}_s} B_i(t)\right](1-\omega\delta t) + \theta l^2 \rho_s(t) p_s(t),\\
&= \left[B_s(t) - \frac{\eta z B_s(t)}{z} + \frac{\eta}{z}\sum_{i\in\mathcal{N}_s} B_i(t)\right](1-\omega\delta t) + \theta l^2 \rho_s(t) p_s(t),\\
&= \left[B_s(t) + \frac{\eta}{z}\sum_{i\in\mathcal{N}_s}\left(B_i(t) - B_s(t)\right)\right](1-\omega\delta t) + \theta l^2 \rho_s(t) p_s(t),\\
&= \left[B_s(t) + \frac{\eta l^2}{z}\sum_{i\in\mathcal{N}_s}\left(B_i(t) - B_s(t)\right)/l^2\right](1-\omega\delta t) + \theta l^2 \rho_s(t) p_s(t),\\
&= \left[B_s(t) + \frac{\eta l^2}{z}\Delta B_s(t)\right](1-\omega\delta t) + \theta l^2 \rho_s(t) p_s(t).
\end{aligned} \tag{4.13}$$

Continuing, let $l^2/\delta t$ be fixed with a value of D, and let $\theta\delta t$ be fixed with a value of ϵ.

Subtract $B_s(t)$ from each side then divide each side by δt to obtain the following:

$$
\begin{aligned}
\frac{B_s(t+\delta t)-B_s(t)}{\delta t} &= \frac{\left[B_s(t)+\frac{\eta l^2}{z}\Delta B_s(t)\right](1-\omega\delta t)+\theta l^2\rho_s(t)p_s(t)-B_s(t)}{\delta t}, \\
&= \frac{\eta l^2}{z\delta t}\Delta B_s(t)-\frac{\omega\eta l^2}{z}\Delta B_s(t)-\omega B_s(t)+\frac{\theta l^2}{\delta t}\rho_s(t)p_s(t), \\
&= \frac{\eta D}{z}\Delta B_s(t)-\frac{\omega\eta l^2}{z}\Delta B_s(t)-\omega B_s(t)+\theta D\rho_s(t)p_s(t). \qquad (4.14)
\end{aligned}
$$

Finally, to obtain the continuum version of equation (4.8), take the limit of both sides of equation (4.14) as both l and δt tend to zero (recalling the definition of $p_s(t)$ from the previous chapter), as follows:

$$
\begin{aligned}
\lim_{l,\delta t\to 0}\left\{\frac{B_s(t+\delta t)-B_s(t)}{\delta t}\right\} &= \lim_{l,\delta t\to 0}\left\{\frac{\eta D}{z}\Delta B_s(t)-\cancel{\frac{\omega\eta l^2}{z}\Delta B_s(t)}-\omega B_s(t)+\theta D\rho_s(t)p_s(t)\right\}, \\
\frac{\partial B}{\partial t} &= \frac{\eta D}{z}\nabla^2 B-\omega B+\lim_{l,\delta t\to 0}\left\{\theta D\rho_s(t)p_s(t)\right\}, \\
&= \frac{\eta D}{z}\nabla^2 B-\omega B+\lim_{l,\delta t\to 0}\left\{\theta D\rho_s(t)\left[1-e^{-A_s(t)\delta t}\right]\right\}, \\
&= \frac{\eta D}{z}\nabla^2 B-\omega B+\lim_{l,\delta t\to 0}\left\{\theta D\rho_s(t)\left[1-\sum_{n=0}^{\infty}\frac{(-A_s(t)\delta t)^n}{n!}\right]\right\}, \\
&= \frac{\eta D}{z}\nabla^2 B-\omega B+ \\
&\quad \lim_{l,\delta t\to 0}\left\{\theta D\rho_s(t)\left[1-1+A_s(t)\delta t+\cancel{\sum_{n=2}^{\infty}\frac{(-A_s(t)\delta t)^n}{n!}}\right]\right\}, \\
&= \frac{\eta D}{z}\nabla^2 B-\omega B+\lim_{l,\delta t\to 0}\left\{\theta D\rho_s(t)A_s(t)\delta t\right\}, \\
&= \frac{\eta D}{z}\nabla^2 B-\omega B+\lim_{l,\delta t\to 0}\left\{\epsilon D\rho_s(t)A_s(t)\right\}, \\
\frac{\partial B}{\partial t} &= \frac{\eta D}{z}\nabla^2 B-\omega B+\epsilon D\rho A, \qquad (4.15)
\end{aligned}
$$

which is the final equation presented in [44] representing the continuous version of equation (4.8). A small modification to this equation is obtained by recalling that the overall criminal attractiveness A is the sum of the dynamic component of criminal attractiveness B with a component that is static with respect to time, A^0. Equation (4.15) can thus be modified to be written in terms of the overall criminal attractiveness, A as follows:

$$\frac{\partial A}{\partial t} = \frac{\eta D}{z}\nabla^2 A - \omega\left(A - A^0\right) + \epsilon D \rho A. \tag{4.16}$$

This is the first of the two PDEs composing the continuous model of residential burglary. To obtain the second PDE, more work is required than simply rearranging and limiting an existing equation. Note that while the goal is for the second PDE to represent how criminal density as a continuous function evolves over time, the discrete model does not contain an "update rule" similar to that of equation (4.8), governing how criminal density changes between time-steps in the discrete model. To begin the derivation of a PDE to represent how criminal density evolves with respect to time, the following equation is introduced to give the "expected" number of criminals at site s after the time-step beginning at time t. Note that, as $\mathcal{N}_s$ denotes the set of all sites neighbouring s, for an arbitrary $i \in \mathcal{N}_s$, $\mathcal{N}_i$ denotes the set of all sites neighbouring i:

$$n_s(t+\delta t) = A_s(t) \sum_{i \in \mathcal{N}_s} \left(\frac{n_i(t)\left[1 - p_i(t)\right]}{\sum\limits_{j \in \mathcal{N}_i} A_j(t)} \right) + \Gamma \delta t. \tag{4.17}$$

To unpack the meaning behind this "update rule", consider from the discrete model of residential burglary that there are only two ways by which a criminal can be at a given site s after a time-step: Either they moved from an adjacent site i during the time-step, or they were generated at site s during the time-step. To account for the expected number of criminals moving from adjacent sites during the time-step, consider that for an arbitrary site i neighbouring s, $n_i(t)\left[1 - p_i(t)\right]$ gives the expected number of criminals that will chose not to burglarize site i and that will instead opt to move to an adjacent site. Multiplying this value by $A_s(t) / \sum\limits_{j \in \mathcal{N}_i} A_j(t)$, the probability that a criminal will move from site i to site s given they opted to move sites, will give the expected number of criminals that moved from site i to site s over the time-step. Finally, considering the summation of this value over all sites neighbouring s, gives the expected number of criminals arriving at site s from adjacent sites over the current time-step. To account for criminals being generated at site s, the term $\Gamma\delta t$ is introduced, simply being the rate of criminal generation multiplied by the duration of the given time-step.

As was the case in deriving the continuum version of equation (4.8), the above "update rule" for $n_s(t+\delta t)$ will allow for the derivation of a continuum equation representing how criminal density evolves over time. This derivation involves computing the following limit:

$$\lim_{l,\delta t\to 0}\frac{\rho_s(t+\delta t)-\rho_s(t)}{\delta t}=\lim_{l,\delta t\to 0}\left\{\frac{1}{l^2\delta t}\left(A_s(t)\left(\sum_{s'\sim s}\frac{l^2\rho_{s'}(t)[1-p_{s'}(t)]}{\sum\limits_{s''\sim s'}A_{s''}(t)}\right)+\Gamma\delta t-l^2\rho_s(t)\right)\right\}. \tag{4.18}$$

Once evaluated, the above limit results in the following continuum equation (where Γ/l^2 is fixed and set equal to γ):

$$\frac{\partial\rho}{\partial t}=\frac{D}{z}\nabla\left[\nabla\rho-\frac{2\rho}{A}\nabla A\right]-\rho A+\gamma. \tag{4.19}$$

Together, equations (4.16) and (4.19) form a system of PDEs representing the continuous model of residential burglary, as initially introduced by Short et al. [44]. Both PDEs are parabolic in nature and thus require left and right boundary conditions along each of the spatial axes considered (these equations can either be considered in one or two spatial dimensions, along with one time dimension). No-flux Neumann boundary conditions suffice for both equations (i.e. for the one-dimensional case, $A_x(0,t)=A_x(L,t)=\rho_x(0,t)=\rho_x(L,t)=0$ with L being the length of the spatial domain). It should be noted that both the one-dimensional and two-dimensional cases are of theoretical interest: Whereas the two-dimensional version of the PDE system can be interpreted as representing crime throughout an entire city, the one-dimensional version of the PDE system can be interpreted as representing crime along an individual street.

It turns out that one is not able to simply take this system of PDEs and obtain exact solutions for A and ρ however. To address this issue, for the one-dimensional case we are able to use software to compute approximate error-controlled solutions to the continuous model of residential burglary. Before examining numerical solutions to the one-dimensional continuous model, the BACOL and BACOLI solvers are discussed as a means by which error-controlled numerical solutions to the continuous model of residential burglary may be computed.

4.3 The BACOLI Software Package to Compute Numerical Solutions to Partial Differential Equations

Though one is able to build mathematical models to represent real-world phenomena using systems of PDEs, it is an entirely separate challenge to obtain solutions to these PDE systems. Very often in fact, the PDE systems that are used in real-world mathematical models are impossible to solve exactly. Because of this difficulty there is a great deal of work dedicated to studying and developing methods to approximate solutions to systems of PDEs. As is the case with many of the other sub-fields of numerical analysis, an important problem is to not only compute approximate solutions to PDE systems but to compute accurate estimates of the error associated with these approximate solutions. If a software package is able to estimate the error associated with a computed solution, it can compute several approximate solutions, readjusting the parameters of the method until an appropriate solution is obtained for which the associated error estimate lies within a user-defined tolerance, a process known as *error control*.

Within this section I discuss BACOL and BACOLI, two robust solvers written in the Fortran programming language that can compute error controlled approximate solutions to one-dimensional time-dependant systems of PDEs. While these two solvers use the same underlying algorithm to compute their approximate solutions, they use related but different methods to generate their error estimates. I explain how these two software packages use DASSL [32], a differential-algebraic system solver, along with a collocation method to compute their approximate solutions. I then discuss how both solvers are able to generate error estimates for their approximate solutions and how they implement error control. I then include in Appendix B an approximate solution to the one-dimensional Burgers' equation using BACOLI_PY, a Python wrapper for BACOLI.

4.3.1 The collocation method to approximate solutions to partial differential equations

Underlying both the BACOL and BACOLI solvers is a *collocation method*, which is used to compute approximate solutions to systems of PDEs [14]. Consider the general explicit form for a system of second-order PDEs in one spatial dimension, where a bold symbol represents a vector quantity:

$$\boldsymbol{u}_t(t,x) = \boldsymbol{f}(t,x,\boldsymbol{u}(t,x),\boldsymbol{u}_x(t,x),\boldsymbol{u}_{xx}(t,x)), \quad x \in [0,1], \quad t \geq t_0, \tag{4.20}$$

with initial conditions:

$$\boldsymbol{u}(t_0,x) = \boldsymbol{u}_0(x), \quad x \in [0,1], \tag{4.21}$$

and separated boundary conditions:

$$\boldsymbol{b}_{left}(t,\boldsymbol{u}(t,0),\boldsymbol{u}_x(t,0)) = 0, \quad \boldsymbol{b}_{right}(t,\boldsymbol{u}(t,1),\boldsymbol{u}_x(t,1)) = 0, \quad t \geq t_0. \tag{4.22}$$

For simplicity of presentation, the spatial domain is assumed to be $[0, 1]$, though BACOL, and BACOLI are capable of dealing with arbitrary spatial domains of the form $[a, b]$, where $a, b \in \mathbb{R}$ [14, 48, 33]. The goal of the collocation method is to represent an approximate solution, $\boldsymbol{v}(t, x)$, using a linear combination of known basis functions. Generally speaking, given a set of C^1-continuous basis functions $\{\phi_i(x)\}_{i=1}^{NC}$ defined on $[0, 1]$, the collocation method looks to find an approximate solution $\boldsymbol{v}(t, x)$ of the form:

$$\boldsymbol{v}(t, x) = \sum_{i=1}^{NC} \boldsymbol{\alpha}_i(t)\phi_i(x), \tag{4.23}$$

where the $\{\boldsymbol{\alpha}_i(t)\}_{i=1}^{NC}$ are time-dependant coefficients to be determined and where NC will be defined shortly. To find these coefficients, we substitute $\boldsymbol{v}(t, x)$ into the PDE in equation (4.20) and require that $\boldsymbol{v}(t, x)$ satisfies the PDE at a predefined set of *collocation points* $\{\xi_i\}_{i=1}^{NC} \subset [0, 1]$ with $\xi_1 = 0$ and $\xi_{NC} = 1$. Performing this substitution for $i = 2, \ldots, (NC-1)$, results in the following system of ODEs:

$$\sum_{j=1}^{NC} \boldsymbol{\alpha}_j'(t)\phi_j(\xi_i) = \boldsymbol{f}\left(t, \xi_i, \sum_{j=1}^{NC} \boldsymbol{\alpha}_j(t)\phi_j(\xi_i), \sum_{j=1}^{NC} \boldsymbol{\alpha}_j(t)\phi_j'(\xi_i), \sum_{j=1}^{NC} \boldsymbol{\alpha}_j(t)\phi_j''(\xi_i)\right). \tag{4.24}$$

It is simple to obtain a set of initial conditions for the above system of $NC - 2$ ODEs. Imposing the condition that the approximate solution $\boldsymbol{v}(t, x)$ satisfies the initial condition for the PDE system as given in (4.21) at the collocation points $\{\xi_i\}_{i=2}^{NC-1}$ results in the following linear system:

$$\boldsymbol{v}(t_0, \xi_j) = \sum_{i=1}^{NC} \boldsymbol{\alpha}_i(t_0)\phi_i(\xi_j) = \boldsymbol{u}_0(\xi_j), \quad j = 2, \ldots, (NC - 2), \tag{4.25}$$

where the $\{\boldsymbol{\alpha}_i(t_0)\}_{i=1}^{NC}$ coefficients are easily determined using a linear system solver. Doing so gives a complete set of initial conditions for the system of ODEs given in 4.24. As the PDE itself is not defined for $x = 0$ nor for $x = 1$, instead of requiring that $\boldsymbol{v}(t, x)$ satisfies the PDE at $x = \xi_1 = 0$ and $x = \xi_{NC} = 1$, we require that $\boldsymbol{v}(t, x)$ satisfies the boundary conditions at $x = \xi_1 = 0$ and $x = \xi_{NC} = 1$ as follows:

$$\boldsymbol{b}_{left}(t, \boldsymbol{v}(t,\xi_1), \boldsymbol{v}_x(t,\xi_1)) = \boldsymbol{b}_{left}\left(t, \sum_{i=1}^{NC} \boldsymbol{\alpha}_i(t)\phi_i(\xi_i), \sum_{i=1}^{NC} \boldsymbol{\alpha}_i(t)\phi_i'(\xi_i)\right) = 0, \quad (4.26)$$

$$\boldsymbol{b}_{right}(t, \boldsymbol{v}(t,\xi_{NC}), \boldsymbol{v}_x(t,\xi_{NC})) = \boldsymbol{b}_{right}\left(t, \sum_{i=1}^{NC} \boldsymbol{\alpha}_i(t)\phi_i(\xi_{NC}), \sum_{i=1}^{NC} \boldsymbol{\alpha}_i(t)\phi_i'(\xi_{NC})\right) = 0. \quad (4.27)$$

Solving the IVODE system given by equations (4.24) and (4.25) together with the two algebraic equations in equations (4.26) and (4.27) gives the $\{\boldsymbol{\alpha}_i(t)\}_{i=1}^{NC}$. For this type of system, known as a *Differential-Algebraic Equation* (DAE) system, we can obtain solution approximations using numerical software. One such DAE solver is DASSL [32]; this error-control DAE solver is used in both BACOL and BACOLI [48, 33]. Using DASSL to compute each of the $\{\boldsymbol{\alpha}_i(t)\}_{i=1}^{NC}$ functions, one is thus able to obtain an approximate solution to the PDE system in equation (4.20).

Both BACOL and BACOLI use the same collocation method to compute approximate solutions to systems of PDEs. First, BACOL and BACOLI begin by introducing an increasing discrete spatial mesh:

$$0 = x_0 < x_1 < \ldots < x_N = 1. \quad (4.28)$$

BACOL and BACOLI consider approximate solutions represented in terms of a basis consisting of piecewise polynomials of degree p over each mesh sub-interval $[x_i, x_{i+1}]$, $i = 1, \ldots, N$, with C^1-continuity imposed at each of the internal mesh points x_i. The dimension of the piecewise polynomial space over all of the spatial intervals is thus given by:

$$NC = N(p-1) + 2. \quad (4.29)$$

Instead of using globally supported basis functions $\{\phi_i(x)\}_{i=1}^{N}$ in their collocation scheme, BACOL and BACOLI use a B-spline basis $\{B_i(x)\}_{i=1}^{NC}$, implemented using de Boor B-spline package [8]. The reason a B-spline basis is used, is to ensure that each of the basis functions is zero on a majority of the domain, greatly reducing the time complexity of the entire computation [14]. Let $\boldsymbol{v}(t,x)$ be an approximate solution to a system of PDEs computed using either BACOL or BACOLI. Using the B-spline basis, the s^{th} component of the approximate solution (given by $v_s(t,x)$) is of the form:

$$v_s(t,x) = \sum_{i=1}^{NC} \gamma_{i,s}(t) B_i(x), \tag{4.30}$$

where the $\{\gamma_{i,s}(t)\}_{i=1}^{NC}$ are time-dependant coefficients to be determined. Next, instead of using NC globally defined collocation points, both BACOL and BACOLI chose the $p-1$ Gaussian points defined relatively on each mesh subinterval $[x_i, x_{i+1}]$, $i = 1, \ldots, N$, to be used as the NC collocation points. As with the more general collocation method, by requiring that each component of the approximate solution $v_s(t,x)$ satisfy the PDE at the $p-1$ Gaussian points over each of the mesh subintervals, in addition to requiring that the boundary conditions are met, we obtain a DAE system which can be solved using DASSL. To solve systems of DAEs, DASSL implements a special type of *linear multi-step method*, known as the *backwards differentiation method* [8]. Of importance is the fact that DASSL functions by taking discrete steps through the time domain, returning the values of its error-controlled approximated solution at the end of each time step.

Both BACOL and BACOLI implement the same collocation method. The main difference between the two solvers lies in the way that they compute their error estimates and implement the process of error control, to be discussed in the following sections.

4.3.2 Standard error control and local extrapolation error control in BACOL

Though BACOLI is in a sense a refined version of BACOL, to understand BACOLI it is essential to first understand how BACOL implements error control. Recall from the previous section, that BACOL computes an approximate solution to a system of PDEs in the form of a piecewise polynomial of degree p over the spatial mesh. This degree p computed solution is what is ultimately returned to the user. In addition to this however, BACOL simultaneously computes a second solution in exactly the same manner as the first, in the form of a piecewise polynomial of degree $p+1$. This degree $p+1$ computed solution is hidden from the user entirely, and is only used to compute estimates of the error associated with the degree p computed solution. Let $\boldsymbol{v}(t,x)$ denote the approximate solution of degree p, where $v_s(t,x)$ denotes the s^{th} component of $\boldsymbol{v}(t,x)$. Note that the error associated with $\boldsymbol{v}(t,x)$ is $O(h^{p+1})$, where h is the maximum spatial mesh subinterval size [33]. Further, let $\bar{\boldsymbol{v}}(t,x)$ denote the approximate solution of degree $p+1$, where $\bar{v}_s(t,x)$ denotes the s^{th} component of $\bar{\boldsymbol{v}}(t,x)$. Note that the error associated with $\bar{\boldsymbol{v}}(t,x)$ is $O(h^{p+2})$, where h is again the maximum spatial mesh subinterval size [33]. Following each time step taken by DASSL, BACOL computes an estimate of the error associated with each of the $v_s(t,x)$, each error estimate denoted by $E_s(t)$. To do so, BACOL compares $v_s(t,x)$ with $\bar{v}_s(t,x)$ for all s using the continuous L2 norm as follows:

$$E_s(t) = \sqrt{\int_0^1 \left(\frac{v_s(t,x) - \bar{v}_s(t,x)}{ATOL_s + RTOL_s \left| v_s(t,x) \right|} \right)^2 dx}, \tag{4.31}$$

where t is the current time, and $ATOL_s$ and $RTOL_s$ denote the user-specified absolute error tolerance for $v_s(t,x)$ and the user-specified relative error tolerance for $v_s(t,x)$, respectively. As $\bar{v}_s(t,x)$ has an error of $O(h^{p+2})$ and $v_s(t,x)$ has an error of $O(h^{p+1})$, considering their difference will give an estimate of the error of $v_s(t,x)$.

BACOL is not only able to compute error estimates for its approximate solutions, but is able to adaptively modify the spatial mesh and recompute approximate solutions until the corresponding error estimates lie within the absolute and relative tolerances as specified by the user. If $E(t) = \max(E_s(t)) \geq 1$, BACOL will perform a remeshing and will restart from the beginning of the time step in an attempt to obtain an approximate solution with a smaller error estimate. Further, if the normalized error estimates associated with each subinterval defined by:

$$\hat{E}_i = \sqrt{\sum_s \int_{x_{i-1}}^{x_i} \left(\frac{v_s(t,x) - \bar{v}_s(t,x)}{ATOL_s + RTOL_s \, |v_s(t,x)|} \right)^2 dx}, \quad i = 1, \ldots, N, \tag{4.32}$$

are not all roughly the same size, BACOL will initiate a remeshing and will restart from the beginning of the time-step. To perform a remeshing, BACOL will add and redistribute spatial mesh points so that the normalized error estimates are quasi-equidistributed and so that the overall error estimate should lie within the user-specified tolerances. BACOL is able to iterate this process until the current time-step is successful. At that point, it will then proceed to the next time-step. At the end of the computation, BACOL will return $\boldsymbol{v}(t,x)$ to the user. This is referred to as standard (ST) error control, ensuring that there is an error estimate for the entire returned computed solution that falls within a given user-defined tolerance.

While the ST error control scheme is the only error control scheme that is actually implemented within BACOL, there is a second error control method that BACOL could use: Local extrapolation error control (LE). While not implemented in the source code for BACOL, LE error control for BACOL would follow the exact same process as previously outlined for the ST error control scheme but instead would propagate and return the higher-accuracy solution $\bar{\boldsymbol{v}}(t,x)$ based on the error estimates for the lower-accuracy solution $\boldsymbol{v}(t,x)$. Though the $E_s(t)$ are not accurate error estimates for $\bar{\boldsymbol{v}}(t,x)$, as $\bar{\boldsymbol{v}}(t,x)$ is generally more accurate than $\boldsymbol{v}(t,x)$, the same process of error control as previously outlined will generally be effective, resulting in the error associated with $\bar{\boldsymbol{v}}(t,x)$ lying within the user-specified tolerances. Ultimately, when implementing the LE error control scheme, BACOL would return $\bar{\boldsymbol{v}}(t,x)$ to the user together with the error estimate for $\boldsymbol{v}(t,x)$. In essence, the ST error control scheme works by propagating the lower order approximation based on an error estimate for that approximation, while the LE error control scheme works by propagating the higher order approximation based on an error estimate for the lower order approximation. In terms of performance, neither scheme is entirely more efficient than the other [34].

4.3.3 Interpolation-based error control in BACOLI

BACOLI has two built-in methods to compute error estimates and thus implement error control: The SCI error control scheme and the LOI error control scheme. The SCI scheme can be understood as an modification of the ST scheme found in BACOL, whereas the LOI scheme can be understood as an modification of the LE scheme found in BACOL. I begin here by discussing the SCI error control scheme.

The major issue with BACOL lies in the fact that it computes two full solutions to a given system of PDEs, one of degree p and the second of degree $p+1$. The second solution of degree $p+1$ is used exclusively to compute error estimates for the first solution of degree p, thus more than doubling the time-complexity of BACOL to compute error estimates. BACOLI addresses this issue by completely removing the computation of the approximate solution of degree $p+1$. Instead, BACOLI makes use of what are known as *superconvergent points* present in the approximated solution of degree p to obtain its error estimates.

Recall that for a solution of degree p computed by both BACOL and BACOLI, the overall error associated with the solution is $O(h^{p+1})$, where h is the maximum spatial mesh subinterval size [33]. It has been found empirically however, that for any arbitrary computed solution there are a number of special *superconvergent points*, where the error associated with the solution at these specific points is much better, in some cases as high as $O(h^{2(p-1)})$ [2]. As well, it has been found that on each spatial mesh subinterval there are exactly $p-3$ internal superconvergent points where the error associated with these points is $O(h^{p+2})$. Under the SCI error control scheme, BACOLI constructs the SCI, a piecewise polynomial interpolant with C^1-continuity (implemented as a Hermite-Birkhoff interpolant [33]). Over each mesh subinterval $[x_i, x_{i+1}]$, $i = 1, \ldots, N$, the SCI is built to interpolate the solution values at the $p-3$ internal superconvergent points with error $O(h^{p+2})$, the solution values at the subinterval endpoints, x_i and x_{i+1}, along with their first derivatives, where the error is $O(h^{2(p-1)})$, and the solution values at the closest superconvergent points from each of the adjacent mesh subintervals, where the error of the solution values is $O(h^{p+2})$. This results in a total of $p+3$ interpolation points, and thus the interpolation error associated with the SCI is $O(h^{p+3})$, where h is the maximum spatial mesh subinterval size. As the errors associated with each of the interpolation points is at worst $O(h^{p+2})$, the error in the SCI is dominated by the error of the superconvergent values used as interpolation points and the overall error of the SCI is $O(h^{p+2})$.

To implement the SCI error control scheme, BACOLI uses the SCI in exactly the same way that BACOL uses the degree $p+1$ solution under its ST error control scheme. BACOLI considers the continuous L2 norm between $\boldsymbol{v}(t,x)$ (with error of $O(h^{p+1})$) and the SCI (with error of $O(h^{p+2})$) to obtain an error estimate for $\boldsymbol{v}(t,x)$. BACOLI is then able to adapt and refine the spatial mesh in exactly the same manner as BACOL, to implement error control. In many ways BACOLI works almost identically to BACOL, the only difference being that the degree $p+1$ solution computed in BACOL is replaced by the SCI in BACOLI. This is why the SCI error control scheme in BACOLI can be considered as a modification of the ST error control scheme present in BACOL.

In much the same way, the LOI error control scheme in BACOLI can be considered as a modification of the LE error control scheme present in BACOL, in this case replacing the lower-accuracy solution with an interpolant. Recall that the high accuracy solution has an error that is $O(h^{p+2})$. To implement the LOI error control scheme, BACOLI begins by constructing the LOI, a piecewise polynomial interpolant with C^1-continuity. Over each mesh subinterval $[x_i, x_{i+1}]$, $i = 1, \ldots, N$ the LOI is built to interpolate solution values at $p-3$ internal points and the solution values at the subinterval endpoints x_i and x_{i+1} along with their first derivatives. The internal interpolation points are chosen so that the interpolation error of the LOI agrees asymptotically with the error of the collocation solution of order $p+1$. Since it interpolates $p+1$ values, the LOI has an overall interpolation error of $O(h^{p+1})$. As the errors associated with the solution values are at worst $O(h^{p+2})$, the error in the LOI is dominated by the interpolation error, hence the overall error of the LOI is $O(h^{p+1})$.

Recall from the discussion of the LE error control scheme, that BACOL would generate an error estimate for the lower-accuracy solution then return the higher-accuracy solution. This is valid as the error estimate for the lower-accuracy solution gives an upper bound for the error of the higher-accuracy solution, thus being a valid error estimate for the higher-accuracy solution. The LOI error control scheme implemented in BACOLI follows this template: The LOI approximates the exact solution to the given PDE system with an overall error of $O(h^{p+1})$. Likewise, the approximate solution generated by BACOLI using the collocation method approximates the exact solution to the given PDE system with an overall error of $O(h^{p+2})$. By using the continuous L2 norm as previously discussed, BACOLI creates an error estimate for the LOI, which in turn is a valid error estimate for the approximate solution obtained using collocation. BACOLI is then able to proceed with the process of error control as previously discussed using the LOI. This is the LOI error control scheme as implemented in BACOLI, a modification of the LE error control scheme as implemented in BACOL.

4.4 Results and Discussion of Computed Solutions to the Continuous Model of Criminal Behaviour

Using BACOLI, we are able to compute error controlled numerical solutions to the continuous model of criminal behaviour as in introduced in Section 4.2. Note first however, that BACOLI is only able to compute approximate error-controlled solution to PDEs in one spatial dimension and one time dimension, whereas the PDE model of residential burglary is given in two spatial dimensions and one time dimension. There are two ways to address this incongruity: The first approach is to consider the PDE model of residential burglary in only one spatial dimension, and extrapolate insights from the behaviour of this more simplified model into two spatial dimensions. The second approach is to discretize the two-dimensional PDE model of residential burglary into a large system of one-dimensional PDEs, approximating the entire 2D system. Unfortunately, my attempts to perform a spatial discretization of the two-dimensional PDE system have not been compatable with the BACOLI solver, hence only the first approach will be discussed in detail here.

Recall from Section 4.2 that the continuous model of residential burglary as proposed by Short et al. in [44] is:

$$\frac{\partial A}{\partial t} = \frac{\eta D}{z}\nabla^2 A - \omega\left(A - A^0\right) + \epsilon D\rho A, \tag{4.16}$$

$$\frac{\partial \rho}{\partial t} = \frac{D}{z}\nabla\left[\nabla\rho - \frac{2\rho}{A}\nabla A\right] - \rho A + \gamma, \tag{4.19}$$

where (noting that we are considering only the one-dimensional case here), $A(x,t)$ represents the dynamic component of criminal attractiveness at a given point in time and space, $\rho(x,t)$ represents the overall criminal density at a given point in time and space, and no-flux Neumann boundary conditions are imposed. Instead of using this statement of the continuous model however, I will opt here to instead to use an equivalent restatement of the model provided by Tse and Ward in [46], which is as follows:

$$\frac{\partial A}{\partial t} = \epsilon^2 \nabla A - A + \rho A + \alpha, \tag{4.33}$$

$$\frac{\partial \rho}{\partial t} = D\nabla \left[\nabla \rho - \frac{2\rho}{A}\nabla A\right] - \rho A + \gamma - \alpha \tag{4.34}$$

where $A(x, t)$ and $\rho(x, t)$ are as defined previously, ϵ^2 and D represent the diffusion coefficients of the criminal density, ρ, and criminal attractiveness, A, respectively, α represents the static component of criminal attractiveness (assumed here to be constant over the entire domain), and $\gamma - \alpha$ represents the rate at which criminals are re-introduced into the domain. My reasoning for using this restatement of the continuous model of residential burglary is twofold: Firstly, the statement of the model from equations (4.16) and (4.19) contains a number of extraneous parameters which can be condensed into the dimensionless form presented in equations (4.33) and (4.34). Secondly, Tse and Ward give a much more detailed account of choosing parameter values for equations (4.33) and (4.34) in [46] than Short et al. do for equations (4.16) and (4.19) in [44]. It is for these reasons that we prefer the model provided by Tse and Ward rather than the version of the model provided by Short et al. Regardless, even though there are different parameters, the underlying models are the same [46].

It is not so straightforward however, to map parameter choices from simulations using the discrete model of criminal behaviour as discussed in the previous chapter to simulations using the continuous model of criminal behaviour expressed in equations (4.33) and (4.34). Instead of attempting to correlate parameters between the two models, I instead used insight from Tse and Ward in [46] for choosing optimal parameters designed to illustrate a wide range of possible behaviours of the solutions to the continuous model of residential burglary. As with the simulations from the previous chapter, I performed four separate trials of the continuous model of residential burglary, varying parameters ϵ^2, γ, and D. I opted to fix $\alpha = 1$ both based on simulations from [46] and to fix $A^0 = 1/30$ from simulations of the

Trial	ϵ^2	γ	D
A	10^{-1}	5	0.5
B	10^{-1}	1.15	5
C	10^{-2}	5	0.5
D	10^{-2}	1.15	5

Table 4.3: Parameter choices obtained based off of those in [46] for each of the four simulations using the continuous model of residential burglary as shown in Equations 4.33 and 4.34.

discrete model of residential burglary. Choices of ϵ^2, γ, and D for each of the four trials A,B,C, and D are contained in Table 4.3.

Each of the four trials were evaluated over the spatial domain $x \in [0,1]$ and time domain $t \in [0,50]$, with no-flux Neumann boundary conditions imposed at each of the endpoints $x = 0$ and $x = 1$. Further, each trial was run with the following initial conditions:

$$A(x,0) = \gamma + 0.01 \cdot e^{(-5000\cdot(x-0.5)^2)} + 0.01 \cdot e^{(-5000\cdot(x-0.8)^2)} + 0.01 \cdot e^{(-5000\cdot(x-0.2)^2)}, \quad (4.35)$$

$$\rho(x,0) = 1 - \alpha/\gamma. \quad (4.36)$$

The initial conditions given in equations (4.35) and (4.36) are chosen to be small perturbations to a homogeneous steady state solution to the PDE system, $A(x,t) = \gamma$ and $\rho(x,t) = 1 - \alpha/\gamma$ as discussed in [46]. Plots of the dynamic component of criminal attractiveness $A(x,t)$ associated with each of the four trials are included in Figures 4.2, 4.3, 4.4, and 4.5, respectively. The results of each of these four plots will be discussed in the following chapter.

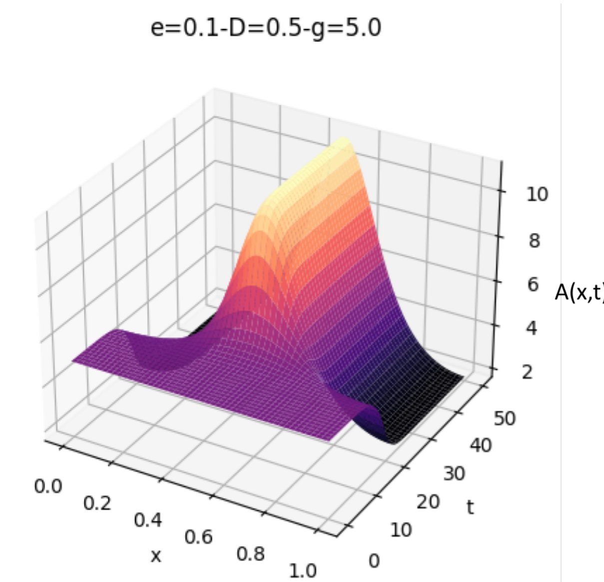

Figure 4.2: Approximate solution $A(x, t)$ computed using BACOLI from the continuous model of residential burglary [46], obtained using the parameters given for Trial A from Table 4.3.

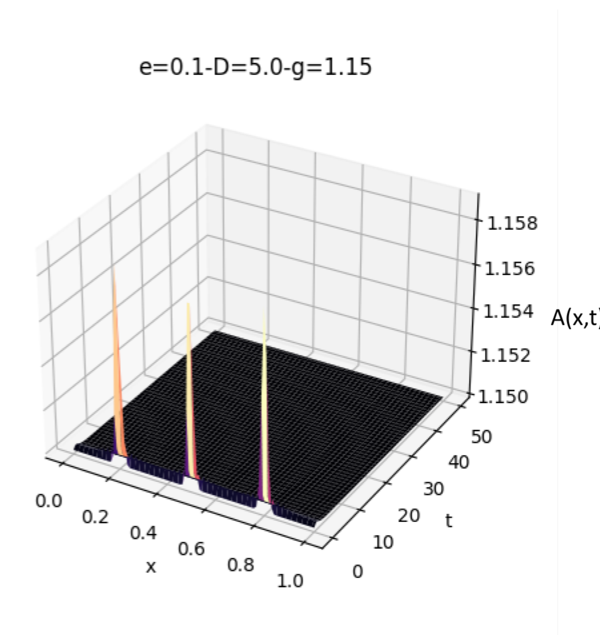

Figure 4.3: Approximate solution $A(x, t)$ computed using BACOLI from the continuous model of residential burglary [46], obtained using the parameters given for Trial B from Table 4.3.

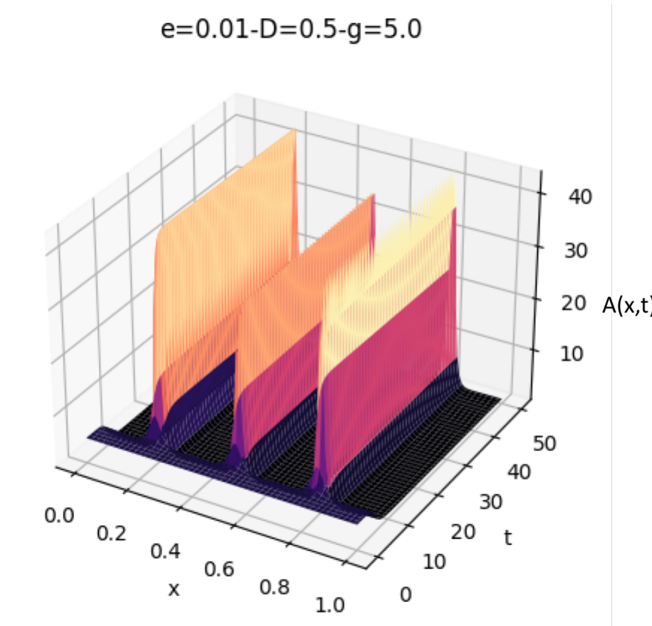

Figure 4.4: Approximate solution $A(x,t)$ computed using BACOLI from the continuous model of residential burglary [46], obtained using the parameters given for Trial C from Table 4.3.

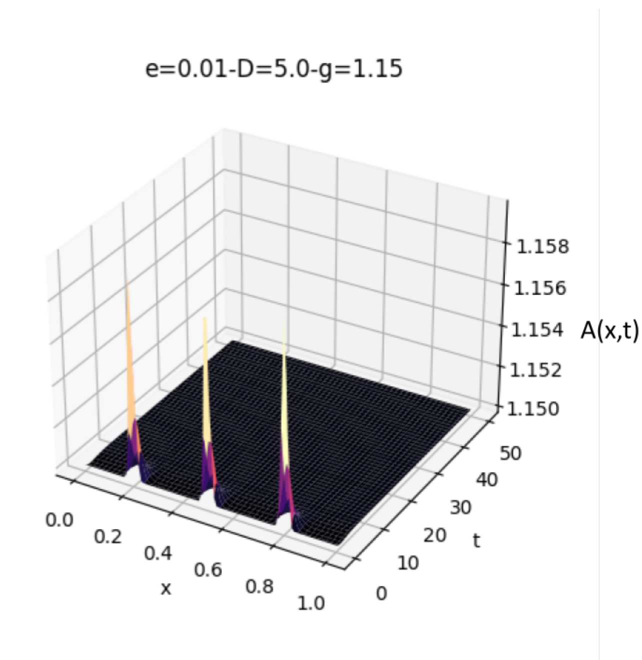

Figure 4.5: Approximate solution $A(x,t)$ computed using BACOLI from the continuous model of residential burglary [46], obtained using the parameters given for Trial D from Table 4.3.

Chapter 5

Discussion of Results From the Discrete and Continuous Models of Criminal Behaviour

With the motivations for and the underlying mathematics of both the discrete and continuous models of criminal behaviour introduced in the previous three chapters, we are now able to pivot to analysing the solutions of these models. Recall the key question motivating this entire project and introduced in Chapter 1: Do predictive policing tools accurately work to predict future crime? Now that we have solutions for both of the models, we are able to begin discussing the answer to this question. What types of behaviour are these models able to produce, and do these behaviours agree with real-world crime dynamics? Further, how are we able to rigorously assess whether either of these models are able to be used in the real world, with a reasonable amount of accuracy?

Within this chapter, I will review the numerical results from both the discrete and continuous models of criminal behaviour, to comment on their general behaviour. I compare the results from both models against existing literature on these models to assess whether my computed results agree with previously published solutions to these models. I then contrast both models against each other to understand their ranges of behaviour. I follow this with a return to the work of David Weisburd [49] to comment on whether the numerical results from these models are realistic with respect to existing theories in empirical criminology.

5.1 An Examination of Results from the Discrete and Continuous Models of Criminal Behaviour

5.1.1 The Discrete Model

To begin an analysis of the discrete model of criminal behaviour, first recall the four simulations performed using the discrete model of criminal behaviour shown originally in Figures 3.2, 3.3, 3.4, and 3.5. Each of the four simulations were performed using this discrete model with parameter choices obtained from [44]. Further, no criminals are present at time $t = 0$ and the criminal attractiveness is homogeneous throughout the entire domain at time $t = 0$. For each of the four trials, snapshots of the overall criminal attractivenesses taken from early in the simulation, at time $t = 200$, are displayed in Figure 5.1. The final snapshots of the overall criminal attractivenesses from each of these four trials, each taken at time $t = 900$, are displayed in Figure 5.2. Note that both Figures 5.1 and 5.2 are simply generated from Figures 3.2, 3.3, 3.4, and 3.5, to aid here in discussing the behaviour of the discrete model of criminal behaviour. Further, to align with the lettering used in Chapter 3 the four simulations performed using the discrete model are labelled A,B,C, and D, displayed in order from top-left to bottom-right in both Figures 5.1 and 5.2.

There are a number of patterns present in the discrete model of criminal behaviour that can be seen in the results shown in Figures 5.1 and 5.2. Using the parameters obtained from [44] there are two general behavioural regimes apparent from these results: In both trials A and B (see the upper two plots in Figures 5.1 and 5.2) the criminal attractiveness initially displays intense small hotspots, that evolve into larger, amorphous hotspots. As the simulations progress, these larger hotspots slowly combine and decrease in intensity, until a homogeneous steady-state is reached. Contrasting this, in both trials C and D (see the lower two plots in Figures 5.1 and 5.2) the criminal attractiveness initially displays intense small hotspots that remain throughout the entire duration of the simulations. These hotspots grow in number and in intensity, but remain absolutely static in space and almost absolutely static

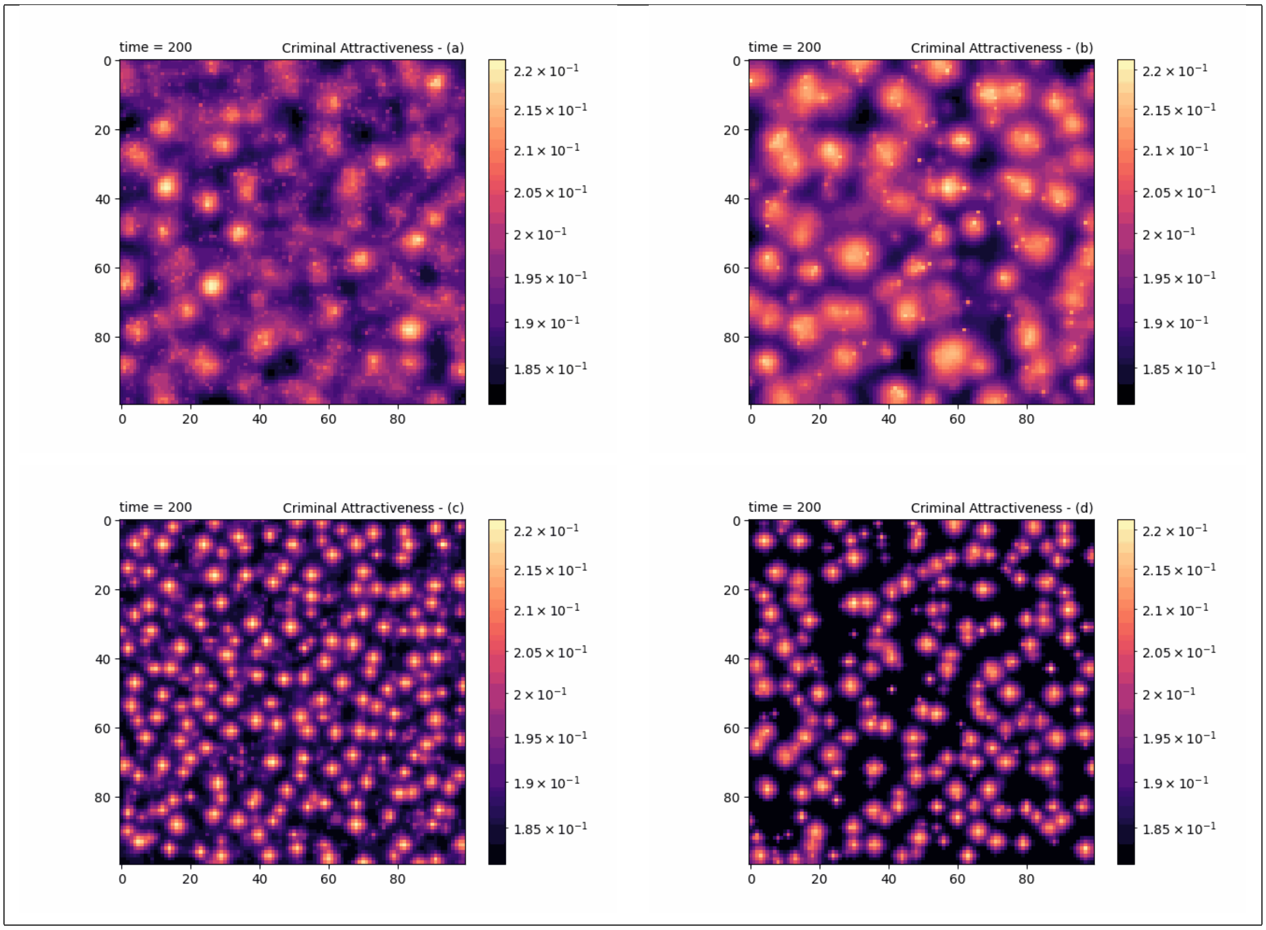

Figure 5.1: Snapshots at $t = 200$ from each of the four separate simulations performed using the discrete model of criminal behaviour first introduced and discussed in Chapter 3. Parameter choices for each of the four trials (labelled A,B,C, and D from top-left to bottom-right) are discussed in Table 3.2 in Chapter 3.3.

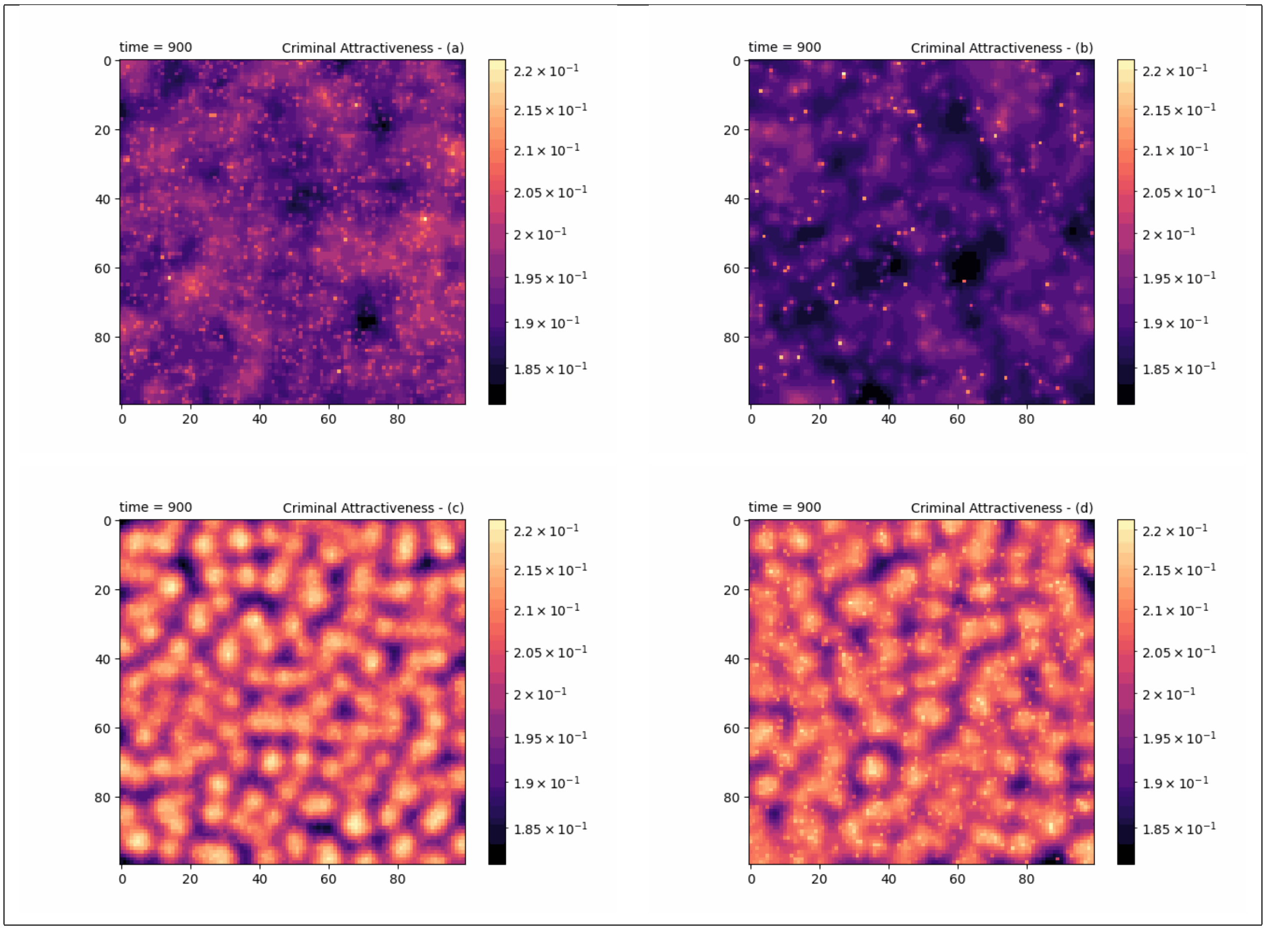

Figure 5.2: Snapshots at $t = 900$ from each of the four separate simulations performed using the discrete model of criminal behaviour first introduced and discussed in Chapter 3. Parameter choices for each of the four trials (labelled A,B,C, and D from top-left to bottom-right) are discussed in Table 3.2 in Chapter 3.3.

in size. A number of the smaller hotspots are amalgamated with nearby larger hotspots in both trials C and D as the simulations evolve, but regardless, the two trials do not evolve towards a homogeneous steady state: Rather, they evolve towards a steady state of intense, localized static hotspots.

While these trials represent a very small choice of potential parameters, their behaviour is still notable. First, note that the behaviour of trials C and D aligns with an observation documented in a number of empirical criminological studies discussed in Chapter 2: That crime in a general sense is seemingly arranged into tight spatial hotspots that remain static over longer periods of time. Note that while trials A and B do not display tight, static hotspots of crime that last throughout the entirety of the simulation, the fact that hotspots form at all in the short-term is notable. Further, while the hotspots in each of the four simulations are all of different size and shape, it can be seen that in all four trials, the hotspots are relatively static in space. In no trial are there profoundly dynamic hotspots that move throughout the entire lattice.

Of interest, is that Short et al. performed similar simulations using the same discrete model and the same choice of parameters in [44] - the results of which are presented in Figure 5.3. The results of the simulations performed here generally do not agree with the results from Short et al. in the short term, but the long term behaviours appear similar: Examining the simulated results from Trial A, Short et al. find that the criminal attractiveness does not display any hotspot behaviour whatsoever, where I found larger amorphous hotspots forming near the beginning of the simulation.

Both simulations agree in the long term however, where a homogeneous steady state is reached. In Trial B, Short et al. find that the criminal attractiveness is arranged into large amorphous *dynamic hotspots*. While in my simulation the criminal attractiveness did

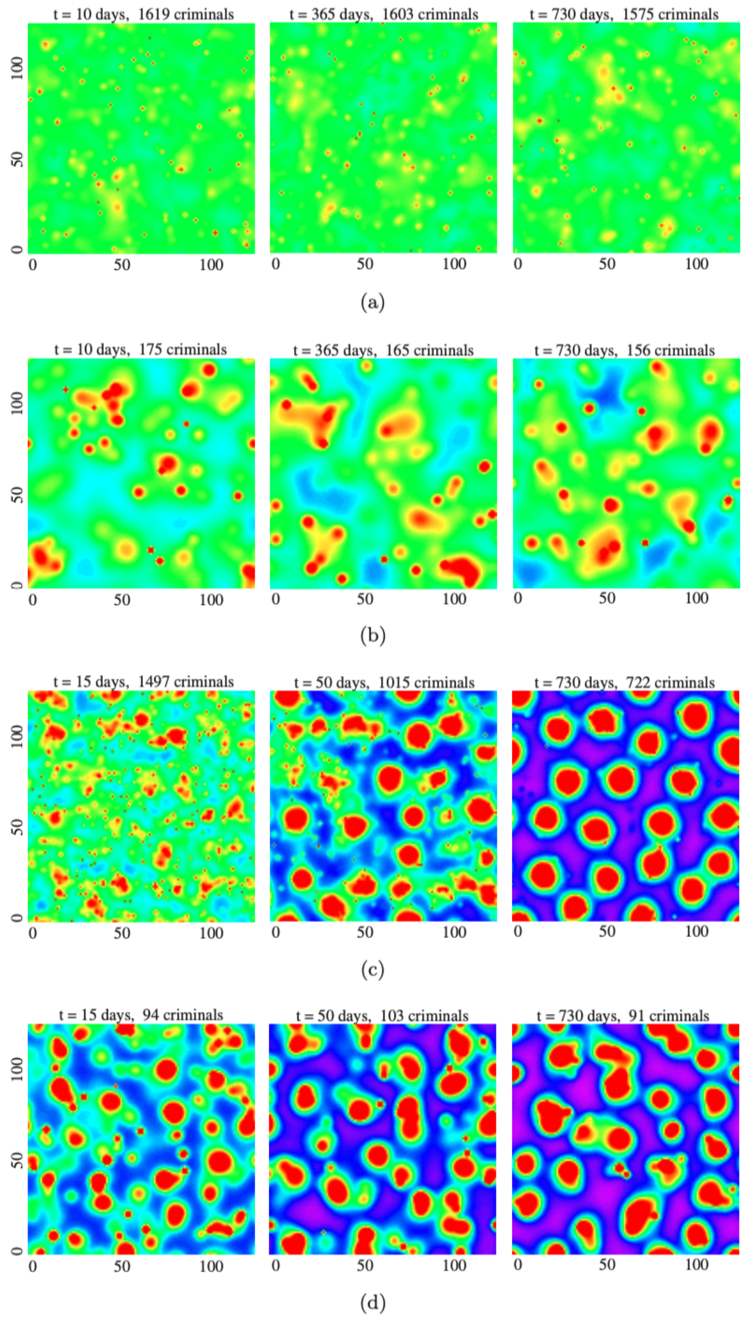

Figure 5.3: Simulated results using the discrete model from Short et al., using the same choice of parameters from Table 3.2 for each trial. Criminal attractiveness is plotted from lower values displayed in violet to higher values displayed in red. [44, Figure 3]

arrange into similarly shaped hotspots, these hotspots were not dynamic in space. In contrast to the results from Trial A, whereas in Trial B Short et al. found that these dynamic hotspots remained constant throughout the duration of the simulation, I found here that a homogeneous steady state was eventually reached. In Trials C and D, Short et al. find that the criminal attractiveness was arranged into large well-defined hotspots, with those in Trial C remaining static in space whereas those in Trial D were dynamic in space. These results generally disagree with my simulated results again: Not only did my simulations display much smaller hotspots than those from Short et al., but in both Trials C and D I found static hotspots where Short et al. found static hotspots forming in Trial C and dynamic hotspots forming in Trial D. Between both sets of simulations however, both hotspot and steady state behaviour is present.

5.1.2 The Continuous Model

Pivoting now to the continuous model of criminal behaviour, recall the four numerical solutions obtained from this model for different parameter choices, originally shown in Figures 4.2, 4.3, 4.4, and 4.5. Of these four numerical solutions, three different distinct behavioural regimes are present, and are reintroduced here in Figure 5.4, to aid in discussing this continuous model. As was done previously with the discussion of the discrete model of criminal behaviour, to align with the lettering used in Chapter 4, the three distinct solutions to the continuous model of criminal behaviour are labelled A, B, and C, and are displayed in order from left to right in Figure 5.4.

Before discussing the plots shown in Figure 5.4, it is important to recall the initial

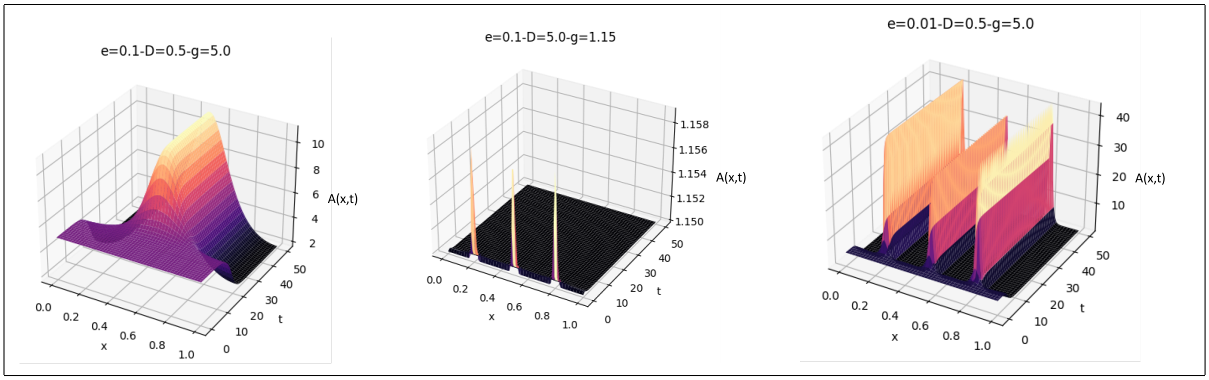

Figure 5.4: Separate numerical solutions to the continuous model of criminal behaviour first introduced and discussed in Chapter 4, chosen to illustrate the three potential behaviours of the model. Parameter choices for each of the three trials (labelled A,B, and C from left to right) are discussed in Table 4.3 in Chapter 4.4.

condition for the criminal attractiveness as used in each of three Trials:

$$A(x,0) = \gamma + 0.01 \cdot e^{(-5000 \cdot (x-0.5)^2)} + 0.01 \cdot e^{(-5000 \cdot (x-0.8)^2)} + 0.01 \cdot e^{(-5000 \cdot (x-0.2)^2)}. \quad (4.35)$$

What this initial condition implies is that, essentially, each numerical solution was computed with a uniform initial condition, save for three small non-uniformities at $x = 0.2$, $x = 0.5$, and $x = 0.8$. Looking now at Figure 5.4, the three different behavioural regimes of solutions to the continuous model of criminal behaviour are as follows: In Trial A (the leftmost plot in Figure 5.4) the criminal attractiveness forms a large, static hotspot in the center of the spatial domain. In Trial B (the center plot in Figure 5.4), the spikes in criminal attractiveness immediately decay to a homogeneous steady state. In Trial C (the rightmost plot in Figure 5.4), the initial spikes in criminal attractiveness persist and multiple strong static hotspots form. Regardless of the trial, no dynamic hotspot behaviour is present.

Of further note is the fact that the computed solutions to the continuous model of residential burglary created and discussed here agree well with computed solutions to the same model as discussed both by Short et al. in [44] and by Tse and Ward in [46]. Results from

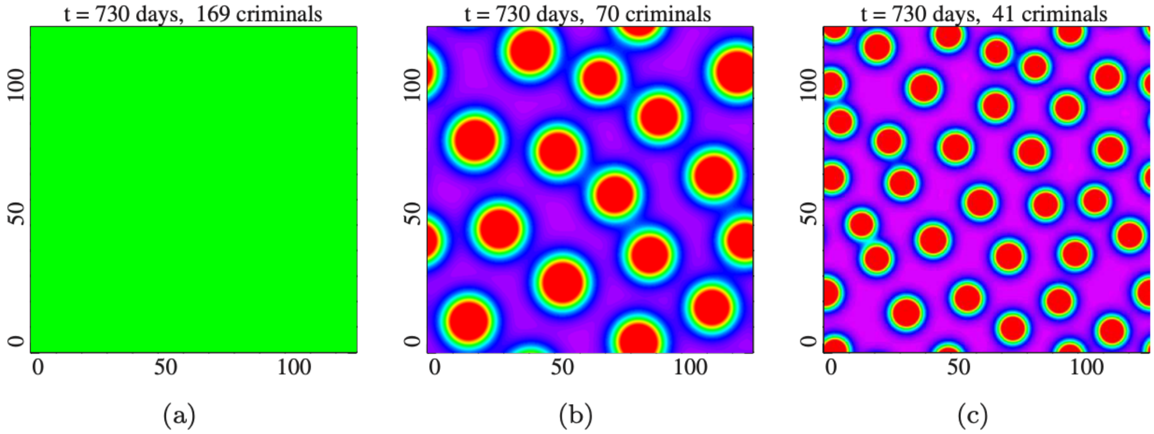

Figure 5.5: Three separate numerical solutions to the two-dimensional continuous model of residential burglary from Short et al. Areas of red indicate higher criminal attractiveness, areas of purple indicade lower criminal attractiveness, areas of green indicate average criminal attractiveness [44, Figure 4].

simulations by Short et al. using the two-dimensional version of the continuous model of residential burglary are included in Figure 5.5, where they find two different behavioural regimes: The first behaviour displayed is that for certain choices of parameters no hotspots form, and the entire spatial domain displays a uniform criminal attractiveness. The second behaviour displayed is that for certain choices of parameters multiple small static hotspots form. Both behavioural regimes align with those found by my computed solutions to the same models. Results from a simulation by Tse and Ward using the one-dimensional version of the continuous model of residential burglary are included in Figure 5.6, where they find behaviour similar to that of Trial A shown in Figure 5.4 here. The computed solution presented by Tse and Ward shows one large static hotspot forming in the center of the spatial domain, agreeing with results from my simulations.

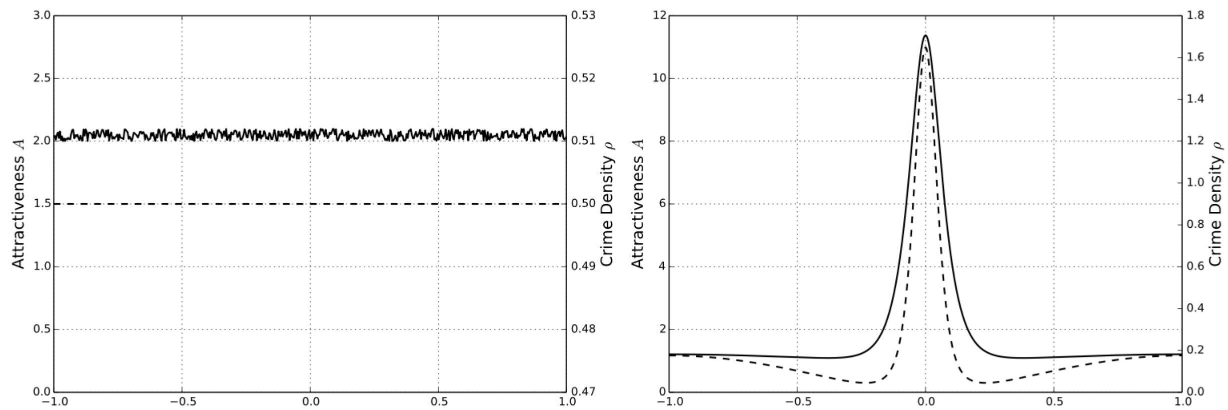

Figure 5.6: Numerical solution to the one-dimensional continuous model of residential burglary from Tse and Ward. Leftmost plot indicates initial condition, and rightmost plot indicates steady state behaviour of the computed solution. Criminal attractiveness is plotted with a solid curve, where criminal density is plotted with a dashed curve [46, Figure 1].

5.1.3 Summary of Numerical Results

In summation, the results of both the discrete and continuous models of criminal behaviour appear to qualitatively agree with each other. For different choices of parameters, both models display both a homogeneous steady state behaviour and a static spatial hotspot behaviour. Of note, whereas Trials A and B of the discrete model achieved a homogeneous steady state near the end of the overall simulation, around time $t = 900$, Trial B of the continuous model achieved a homogeneous steady state almost instantly. Furthermore, Trials C and D of the discrete model slowly developed intense spatial hotspots of heightened criminal attractiveness, whereas Trials A and C of the continuous model very quickly developed strong hotspots. In examining both cases, while the models show similar long-term behaviour, the speed in which they achieve steady state differs greatly. As well, for the chosen parameters, neither model displayed any type of dynamic hotspot behaviour.

Having used both the discrete and continuous models of criminal behaviour to generate simulated crime data, what remains is to assess the applicability of these results. That is to say, how well do the results from either of these models agree with real-world crime dynamics? Although both of the models are rooted in empirical criminology, and due to the use of BACOLI, the error-controlled PDE solver, the computed numerical solutions to the continuous model represent accurate solutions to the continuous model, how authentic are the models themselves? Effectively answering this question is difficult, but not impossible.

5.2 Assessing the Applicability of Results from the Discrete and Continuous Models of Criminal Behaviour

It is a complicated challenge to assess whether the results from either the discrete or continuous models agree with crime trends from the real world. From the work discussed in Chapter 2, it is clear through an examination of empirical crime data that regardless of city, crime as a phenomenon appears to self-arrange into tight spatial hotspots of high activity [49]. Criminal psychology gives further insight into this general trend, motivating this behaviour in both rational choice theory and in the population heterogeneity model of space [51]. Without losing generality and looking into specific cities however, it is effectively impossible to say anything more specific about these crime hotspots or their dynamics. When examining either of the two mathematical models of criminal behaviour discussed in this project, for certain parameter choices, they may describe criminal behaviour with high accuracy in certain environments, while simultaneously being completely inaccurate in other environments. Pulling back even further, there is no clear mechanism by which to tune model parameters or chose realistic initial conditions for either of the two models in an effort to predict crime in a real city (in their paper introducing these models, Short et al. do not discuss in any detail how in which they obtain their parameter choices for their simulations [44]). How possible is it then, to attend to the clear question at the heart of this project: Can predictive models of criminal behaviour fundamentally work to predict future criminal behaviour in the real world?

In an effort to comment on whether either the discrete or continuous models of criminal behaviour are remotely effective at replicating criminal behaviour, recall the "physical law of crime concentrations", as proposed by Weisburd in [49], first introduced in Chapter 2. Though (outside of a general hotspot behaviour) crime dynamics between cities are radically different and context-dependant, through an extensive review of empirical crime data Weisburd has proposed a general law governing crime organization: "For a defined measure of crime at a specific microgeographic unit, the concentration of crime will fall within a narrow bandwidth of percentages for a defined cumulative proportion of crime. For 50 percent concentration, that bandwidth is about 4 percent (from 2.1% to 6%), and for 25 percent concentration, that bandwidth is less than 1.5 percent (from 0.4% to 1.6%)" [49]. Weisburd's physical law of crime concentrations serves as a concrete base-level test of the discrete and continuous models of criminal behaviour. Should either model accurately replicate criminal behaviour in the real world, then for realistic choices of parameters, these models should agree with the physical law of crime concentrations. For the case of the discrete model of criminal behaviour, consider Figures 5.7 and 5.8, each detailing the percentages of grid-points at which 100%, 75%, 50%, and 25% of all crimes occur, for Trials A and D of the discrete model respectively (due to the overall similar behaviour between Trials A & B, and Trials C & D, this analysis is performed only on Trials A and D).

Upon examining Figures 5.7 and 5.8, it is immediately clear that Trials A and D of the discrete model of criminal behaviour do not align with David Weisburd's proposed law of crime concentrations. In the case of Trial A, 50% of crime occurs at 34.64% of all possible grid-points, and 25% of crime occurs at 13.56% of all possible grid-points. Were Trial A to agree with Weisburd's physical law, 50% of crime should occur *at most* at 6% of all possible grid-points, and 25% of crime should occur *at most* at 1.6% of all grid-points. What this ultimately means, is that the simulation of crime generated in Trial A of the discrete model

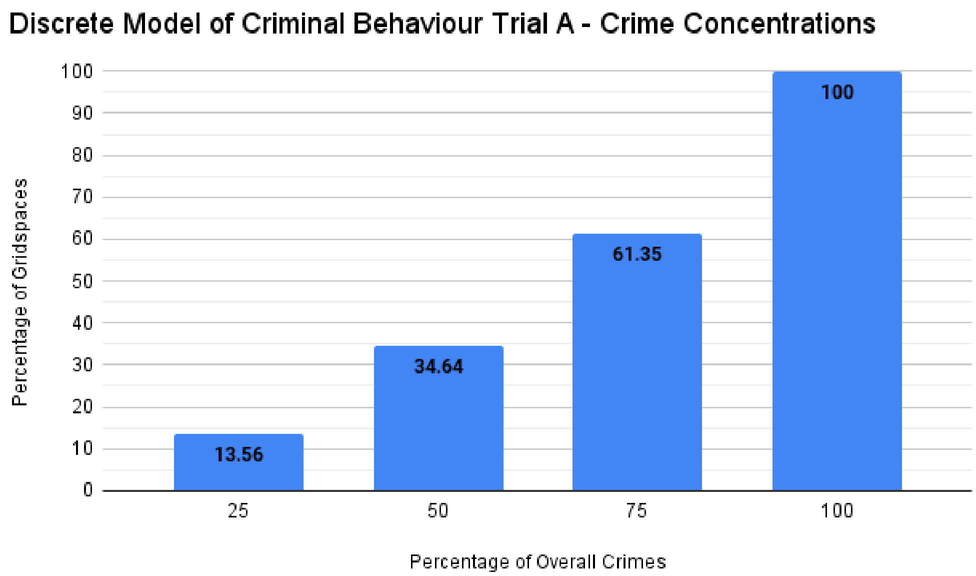

Figure 5.7: For Trial A of the discrete model of criminal behaviour, a plot of what overall proportion of grid-points, 100%, 75%, 50%, and 25% of all crimes occur at, over the course of the entire simulation.

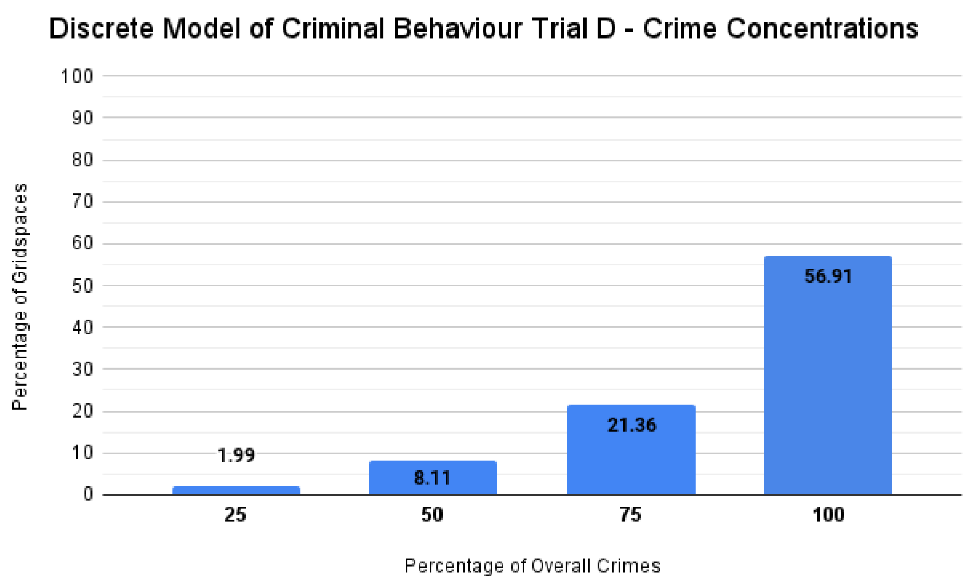

Figure 5.8: For Trial D of the discrete model of criminal behaviour, a plot of what overall proportion of grid-points 100%, 75%, 50%, and 25% of all crimes occur at, over the course of the entire simulation.

does not give crime concentrations remotely close to what has been observed in the real world. Trial A of the discrete model of criminal behaviour displays fundamentally different behaviour from all empirical observations of real-world crime.

Though more concentrated, the results from Trial D of the discrete model are only somewhat more representative of real-world crime data. In the case of Trial D, 50% of crime occurs at 8.11% of all possible grid-points, and 25% of crime occurs at just 1.99% of all possible grid-points. These results are much more in line with what has been observed in the real world (6% and 1.6% respectively), though still these concentrations do not quite fall within Weisburd's proposed bandwidth of percentages. These plots lend doubt to the feasibility of this model of criminal behaviour being used to replicate real-world criminal behaviour, given they show that the computed solutions to the discrete model fundamentally do not share notable characteristics with what would be expected of real-world crime.

Unfortunately, although it would be of substantial interest and use to compare the continuous model of criminal behaviour against the physical law of crime concentrations in the same way as was done with the discrete model, there appears to be a major issue: While with the discrete model it is possible to monitor and record crime occurrence throughout the simulation, there is no discernible way to extrapolate crime occurrence from the continuous model of criminal behaviour. The continuous model consists of two PDEs, one governing how criminal attractiveness evolves through time and space, the second governing how criminal concentration evolves through time and space. The same concrete analysis can thus not easily be performed on the continuous model, due to there being no clear way to divine actual crime occurrence from criminal concentrations. It is thus unclear, whether the continuous model of criminal behaviour accurately reproduces Weisburd's physical law of crime concentrations, or not.

Although these models are able to generate intense hotspots simulating the general hotspot nature of crime, it is unclear whether these models are actually capturing the subtle dynamics of real-world criminal behaviour or are just plausibly painting similar pictures. The comparisons against David Weisburd's physical law of crime concentrations lend some doubt to the effectiveness of these models, though this is still not an entirely conclusive test. It is thus fundamentally inconclusive, whether either the discrete or continuous models of criminal behaviour as proposed by Short et al. in [44] could actually be used to effectively predict crime. Aside from a large-scale controlled field test of these models (something to be discussed further in the following chapter), there does not appear to be any encompassing way to assess whether these predictive models are truly able to accurately predict crime. Potentially though, the focus of this project thus far has been misplaced. In a sense, questions of whether predictive crime models are mechanically effective at crime prediction do not get to the core problems of predictive crime modelling. When discussing any problem within an applied mathematics context, it is important to understand the greater context in which one is working. Pulling back from the mechanics of crime prediction reveals an array of issues completely isolated from the mathematics of generating crime predictions in the first place. In a sense, whether the underlying math of predictive policing works or not is irrelevant: What truly matters are the issues endemic to how predictive policing tools have actually manifested in the real world.

Chapter 6

Reconsidering Predictive Policing

It is at this point that a serious question must be addressed: What exactly is predictive policing? Although it may seem as though this is a simple question, the deeper one digs the more confusing and sprawling the answer becomes. Until this point I have used a somewhat vague working definition: That "predictive policing" is the collection of frameworks and tools concerned with predicting crime occurrence within a given environment and advising on-the-beat police officers on how to best address this forecasted crime. While this understanding of predictive policing was sufficient to discuss both the discrete and continuous models of residential burglary as proposed by Short et al. in [44] and as discussed in Chapters 3, 4, and 5, upon a closer examination, this definition begins to break down. Observe that this definition is extremely broad: Under this definition the study of environmental criminology as a whole must be a part of "predictive policing", as it attempts to situate how crime is organized with respect to space and time, hence grounding how to forecast crime occurrence in concrete theory. Further, both proprietary software sold to police departments marketed to forecast crime and existing tools and techniques of crime modelling used by police departments must be a part of "predictive policing", as their primary goal is to forecast and pre-empt crime occurrence. Even further, the underlying mathematical models that these tools use, e.g., both the discrete and continuous models of residential burglary proposed by Short et al. in [44] must be a part of "predictive policing", given that they too are concerned with forecasting and thus advising police officers on the future occurrence of crime. The argument could even be made that under this definition of "predictive policing" the lived experiences of on-the-beat police officers constitute "predictive policing", as any given officer who has worked the same beat for a long period of time will have some intuitive understanding of where and when crime will occur throughout their patrol. If all of these are manifestations of some nebulous "predictive policing", then are we able to study and critique predictive policing in the first place given its seemingly endless scope?

Ontological spiralling aside, there is remarkably little consensus towards a definition of predictive policing within peer-reviewed academic literature. In their 2019 literature review of to-date academic work on predictive policing [28], Meijer and Wessels comment that while there is no absolute definition of predictive policing accepted by a majority of scholars, many agree on some of the key properties of predictive policing. Based on the work of applied mathematicians Camacho-Collados and Liberatore in [5], Meijer and Wessels find in accordance with my working definition proposed earlier that "predictive policing entails the application of quantitative techniques to forecast where criminal activities might occur in the (near) future" [28]. Based on the work of criminologists Chan and Moses in [6], Meijher and Wessels proceed to argue that another key property of predictive policing is that "predictions based on these analytic tools can guide the decision-making of law enforcement agencies, especially with the deployment of its personnel" [28]. The work of Uchida of the United States National Institute of Justice offers additional insight into the impetus for predictive policing: "predictive policing is a concept that is built on the premise that it is possible to predict when and where crimes will occur again in the future by using sophisticated computer analysis of information about previously committed crimes" [47]. Based on their review, Meijher and Wessels ultimately settle on the following more precise definition of predictive policing:

"Predictive policing is the collection and analysis of data about previous crimes for identification and statistical prediction of individuals or geo-spatial areas with an increased probability of criminal activity to help developing policing intervention and prevention strategies and tactics" [28].

How then do the discrete and continuous models of residential burglary as proposed by Short et al. in [44] and discussed throughout the previous chapters relate to this definition proposed by Meijer and Wessels? It is clear, based on this more formal definition, that these two models *are not* tools of predictive policing, at least not in isolation. Although these models seek to replicate the behaviour of criminals engaging in residential burglary, they do not inherently form predictions using data about previous crimes, nor do they have any mechanism by which to advise police officers or analysts to develop policing intervention and prevention strategies and tactics. It is only when these two models of residential burglary are integrated into larger software packages that are able to generate appropriate parameter choices, initial conditions, and boundary conditions, based on existing policing data, that they are able to make specific predictions about future crime occurence within a given area. Furthermore, it is not without both a significant front-end and back-end that these models are able to be integrated for use by police officers on-the-beat, to advise their behaviour in a meaningful way. It is this detail, that in the real world the two models of residential burglary cannot function in isolation, that is the key focus of this chapter. Although the accuracy of these two models is a key concern within the study of predictive policing more generally, of equal importance is studying the way in which that these models are implemented in real-world software, and the problems that this implementation may pose.

6.1 The Complicated Nature of Crime Data

If their primary goal is to predict future crime occurrence, it is essential that predictive policing tools make use of historical crime data pertaining to a given location. As discussed in the introduction to this chapter, the underlying mathematical models and criminological theory can only perform so much of the work. In the specific cases of the discrete and continuous models of residential burglary as proposed by Short et al. in [44], without realistic choices of model parameters and initial/boundary conditions based on historical crime data there is no chance that these models will be able to replicate crime trends in any specific context. Although we do not know exactly how specific predictive policing tools integrate historic policing data into their underlying numerical models, questioning the validity and nature of this historic policing data in the first place has significant implications on the potential uses of these tools.

For decades, criminologists have understood that significant issues arise when crime data obtained by police departments are conflated with the real occurrence of crime. In 1963 criminologists Kituse and Cicourel [21] published a foundational work critiquing exactly this problem. By examining how official statistics pertaining to the criminal justice system are obtained, Kituse and Cicourel argue that at a base level these statistics are "unreliable because some individuals who manifest devient behaviour are apprehended, classified and duly recorded while others are not" [21]. Kituse and Cicourel argue further that regardless of intention there are many "successive layers of error" within the criminal justice system (certain crimes going unreported by civilians, certain reported crimes not getting proper responses by police, instances where some are implicated for crimes that they did not commit, etc.), which ultimately lend a level of unreliability to the final data as catalogued by police departments. Elliott, then-director of the Center for the Study and Prevention of Violence expanded substantially on this work by analysing the extent to which police departments and independent researchers have assumed that historic policing data is representative of real crime occurrence. In his 1995 report to the American Society of Criminology [12], Elliott first comments on how there is a vast "false-negative" problem within policing records, and how many know of and accept this issue. In the context of the historic policing data for a given policing jurisdiction, a "false-negative" is introduced into a dataset when either an individual commits a crime and is not apprehended, or when an individual has committed a crime in a separate policing jurisdiction but the information pertaining to this previous crime is not shared between jurisdictions [12]. Demonstrating this point, in a 2012 study for the United States Department of Justice, researchers found that more than half (52%) of all violent crimes within the United States each year were not reported to police; an average total of 3,382,200 violent crimes per-year within the United States were not reported to the police in the first place [24]. Although many analysts agree that historic policing data has severe false-negative problems, Elliott further argues that the issue with this already-problematic data lies in its assumed representativeness of all actual crime [12]. While it is clearly the case that historic policing data is not entirely accurate, there has been an inherent assumption that it at least forms a representative sample for actual crime. Researchers examining this issue of representativeness have however found that generally, historic crime data vastly over-implicates certain groups in committing certain crimes, based on a vast history of racism, sexism, and police abuse over the past number of centuries.

There are a number of ways to critique the issue of representativeness within historic policing data. The approach considered by Elliott again in [12], is to contrast the historic arrest data pertaining to an individual as maintained by a police department with their own self-reported assessment of the crimes they have committed. By analysing data obtained by the National Youth Survey (NYS) between the years 1976 and 1990 for those aged 11-17 in 1976, Elliott is able to fundamentally question the supposed representativeness of historic policing data: Foremost, policing data between the years 1976 and 1990 within the United States demonstrates that 10% of males and 1.5% of females (within the specified age range) were charged with at least one criminal felony [12]. In contrast, the self-reported felony data obtained through the NYS indicated that 48% of males and 18% of females (within the specified age range) committed at least one criminal felony [12]. As discussed previously, the fact that the self-reported percentages are so much higher than the official percentages maintained by police records should not be shocking: Many agree that official police records have a strong false-negative issue. Of significance, is that while official police records indicate here that approximately 7 criminal felonies are committed by males for every 1 criminal felony committed by females, the self-reported records obtained by the NYS indicate that only approximately 3 criminal felonies are committed by males for every 1 criminal felony committed by females. Elliott again performs this same analysis but instead in consideration of broad ethnic groups: He finds that unsurprisingly, while official police records indicate that approximately 2.5 criminal felonies are committed by Hispanic peoples for every 1 criminal felony committed by Caucasian peoples, self-reported records obtained by the NYS indicate that in reality this ratio is extremely close to 1:1. Further, while official police records indicate that approximately 4 criminal felonies are committed by Black peoples for every 1 criminal felony committed by Caucasian peoples, self-reported records again indicate that this ratio is extremely close to 1:1. Elliott's findings confirm that the root problem with policing data is not that it is inherently plagued by false-negatives, but that the data that is obtained is not representative of overall crime occurrence.

If historic policing data is generally not accurate to actual crime occurrence, what is the implication for predictive policing? In a critique of predictive policing, researchers Lum and Issac invoke the adage of "garbage in, garbage out" to describe how historical policing data interfaces with mathematical models to predict crime [27]. At their core, predictive mathematical models are designed with the goal of replicating and propagating trends present within existing historical data forward into the future. Predictive models of any type are thus constrained by the accuracy of the historical data that they are required to use: If there is bias within historical data, predictive models will most often replicate that same bias. In the context of predictive policing, given that historical crime data obtained by police is not necessarily representative of actual crime occurrence, any predictive policing software will at best replicate and maintain existing biases present in police arrests. At worst, predictive policing has the potential to entrench existing biases in police arrests by advising police officers to focus even more intensely on areas already disproportionately targeted by police. Given that historic crime data is notably biased against Black and Hispanic men, as described by Elliott in [12], predictive policing software using this data could reasonably be assumed to disproportionately target areas with higher concentrations of Black and Hispanic men, thus reaffirming and entrenching this discrepancy in crime data. Regardless of the efficacy of the underlying mathematical model, if the historic crime data used by predictive policing software is problematic in any way, then the predictive policing software at its core will not accurately predict the *actual* occurrence of crime. It is in this way that predictive policing can be understood not as a set of tools and methodologies to predict and pre-empt crime, but as a set of tools and methodologies to predict and replicate historical on-the-beat police officer behaviour.

Notably, there have been a number of empirical studies comparing the on-the-beat use of predictive policing against traditional crime modelling strategies, attempting to quantify whether these tools are biased against certain minority groups (see [27, 41, 17, 26, 29, 36]). These studies help to enrich the discussion of historic data biased against certain minority groups: While in theory the use of biased data within predictive policing may lead to biased results, these studies attempt to assess whether this bias actually manifests in practice. The conversation surrounding these studies is nuanced however, and a number of these works raise their own concerns within the philosophy of predictive policing. It is for this reason, that these studies will be discussed in greater detail later in this chapter, though there existence should be noted here.

Predictive policing tools making use of flawed historic policing data to entrench existing biases within day-to-day police operation is far from the only potential issue facing predictive policing as it relates to the use of data. Recall the issue of representation within historic policing data: A number of minority groups (such as Black and Hispanic men) are generally overrepresented in this data [12], causing predictive policing tools to potentially over-implicate these minority groups in criminal activity. Critical geographer Jefferson adds additional context to this idea arguing that couching modern policing strategies in the languages of mathematics, algorithms, and computing science, allows for an increased and potentially misplaced trust in these newer strategies [18]. Essentially, although predictive policing is simply a restatement of existing policing strategies and techniques making use of problematic historical policing data, by invoking buzz-words like "artificial intelligence" and "predictive modelling", lends a perceived degree of scientific rigour to predictive policing. In a study of the Chicago Police Department's engagement with predictive policing, Jefferson comments that although predictive policing has essentially served to maintain the status-quo of policing in Chicago (that is, predictive policing tools implicated the same disproportionate percentages of minority groups in crime as were being implicated prior to the trial), that "racialized policing simply has obtained a new, ostensibly scientific justification" [18]. Given it's reliance on historic data, it is arguable that the "best case-scenario" for predictive policing is that it would replicate and maintain existing biases present in police arrests. Jefferson's work offers a more bleak outlook: At best predictive policing will not only replicate and maintain existing biases present in police arrests, but that it will do so while justifying and shielding these existing biases from critique and outside intervention, using the language of "science" and "absolute objectivity".

The issues in how predictive policing tools make use of data are concerning, yet somewhat speculative in nature. Though a number of studies have been performed analysing the extent to which historic policing data exhibits bias, due to the colloquial "blue wall" (that police departments are often very secretive about their internal operations), it is essentially impossible to understand exactly how historic data is integrated into any particular predictive policing tool. This section should therefore not serve as an examination of a known issue inherent to predictive policing, but moreover as a cautionary warning. Predictive policing tools exist and are in use in real policing departments across North America and around the globe: The known issues with historical data may not be conclusive proof that predictive policing is a flawed technology, but these known issues add to the idea that we must be increasingly critical and careful with novel technologies. Building on this point, recall that in the revised definition of predictive policing proposed at the beginning of this chapter, both historical policing data and predictive mathematical models operating in tandem do not form what we commonly understand as "predictive policing". A key piece of the puzzle that has been so-far missing from this discussion is how predictive policing tools are used and interpreted by police officers. In fact, the relationship between predictive policing and police officer is equally as suspect as the predictions built using predictive policing in the first place. Regardless of the quality of the predictive policing software, the bottom line comes in how this software actually influences police behaviour and thus actual arrests, this being a huge issue in-and-of itself.

6.2 The Gap Between Predictive Model and Police Officer

What if we were to perfect predictive policing? Suppose we were able to develop a software tool that addressed the aforementioned concerns with historic policing data, and that contained a "perfect" mathematical framework that was able to predict future crime occurrence with 100% accuracy. It may seem as though this is pointless speculation: Clearly no predictive model, no matter how good, could ever come close to predicting a phenomenon as chaotic as crime with anywhere close to perfect accuracy. Regardless, by considering such an impossible tool, however unlikely, illuminates a second major concern inherent to predictive policing: How is it that future crime occurrence forecast by a predictive policing model is translated and relayed to police officers working on-the-beat? How are these predictions interpreted by police analysts and then enforced by police officers? Although focus is often placed on the accuracy of predictive policing models and their underlying mathematics, comparatively less time has been spent questioning how predictive policing tools actually manifest as crime enforcement when placed in the hands of police officers. When already-problematic tools are folded into the complex web of day-to-day police operation, these existing issues may be exacerbated far beyond their initial scope.

Through a series of interviews with software developers and ride-alongs with police officers, sociologist Lally argues that the largest issue plaguing predictive policing tools does not lie in the accuracy of their underlying models, but in how those models are interpreted and used [22]. In an interview with an unnamed developer who worked on predictive policing software, Lally reports that the developer comments on the issue of implementation, remarking that "it makes almost no difference which algorithm you use, what matters is how the police turn [predictions] into actionable plans" [22]. Building on this argument, another developer speaks to how the focus of both developers and critics alike has been sorely misplaced on the accuracy of predictive models, and that fundamentally most different algorithms used in predictive policing tools perform with a similar accuracy: The big difference between any two predictive tools lies in how forecasts are interpreted for use by police officers, and as a direct consequence in how police officers enforce these predictions. Lally notices that while many software developers seemed to be quite confident in the accuracy of their models, when questioned about how these models were implemented into the real world many of these developers became much more hesitant to comment on their overall efficacy. There is no set of best practices in place governing how predictions are converted into crime enforcement strategies, a glaring and often unaddressed problem facing predictive policing tools.

Where Lally worked to understand how predictive policing forecasts were exercised in the real world through interviews with software developers, those responsible for creating the predictive policing tools to generate these forecasts, policy analysts Sandhu and Fussey look to understand the same issue, only through the eyes of the police officers who must use these forecasts. Through interviews conducted between 2017 and 2019, Sandhu and Fussey attempted to understand how police officers and departments in the United Kingdom are both using predictive policing tools and addressing criticisms of these same tools [40]. While a number of those interviewed throughout the course of the study indicated their hopes that predictive policing could remove the subjectivity and bias from day-to-day police operations and allow for more effective police patrols, many people commented that the predictive policing tools themselves only served to highlight the subjective decision-making of on-the-beat police officers: Sandhu and Fussey comment that "many research interviewees claimed that officers make the subjective decision as to how a particular event will be labelled and categorised in police records and, thus, how it will be understood by a predictive technology" [40]. It is impossible to have an implementation of predictive policing that completely removes human subjectivity, thus undermining the claims to absolute objectivity that prevail in discussions on predictive policing. Further, Sandhu and Fussey find through their interviews that many of those officers using predictive policing technology are either unsurprised by or unconvinced by the predictions generated by these tools. In the case of the former, many officers noted that predictive policing tools were not telling them anything they didn't already know. One of those interviewed by Sandhu and Fussey, the head of a digital policing unit, succinctly commented that:

"Predictive policing has taken off in the last 5 years, some of it is interesting, but some of it just suggests the obvious; it is likely that burglary will continue in [location], that retail theft will be concentrated in shopping malls... Well, no shit!" [40]

Put in this way, predictive policing tools are of interesting theoretical use, but in practice they may not have the ability to improve on the lived experience of police officers who have worked in the same geographic areas throughout their careers. In addition, Sandhu and Fussey find that many of those interviewed were not only unimpressed by the predictive policing tools, but were unconvinced that the forecasts generated by predictive policing tools should be followed absolutely. In another interview, a separate head of a digital policing unit conveyed that:

"I cannot ever envisage a time when we just take that suggestion or answer, or you know, something will pop out of a machine and [we] will just blindly go and do that. There will always be the need for a human being in that process... to be able to say, 'OK, well the computer says this, but actually we need to [do that].' [...] I would be really unhappy with following a line of inquiry that was purely based on what a computer [was] saying." [40]

In summary, there is noticeable scepticism both between software developers working on creating predictive policing tools, and within police departments using these tools in the real world. There is not necessarily scepticism that the models themselves are flawed, but simply that there are "successive layers of error" [21] in how police officers deploy predictive policing tools to address crime. Despite claims from technology firms and police departments investing in these firms, predictive policing software can never completely remove human decision-making from the equation of policing. Regardless of reason or intention, if the police officers using predictive policing tools neglect to listen to the forecasts given to them, then what is truly the point of using these tools in the first place?

6.3 Is There Anything Truly Certain About Predictive Policing?

Recall from earlier in the thesis the mention of a collection of empirical studies attempting to assess the on-the-beat usage of predictive policing software. Throughout this project, the ultimate goal has been to attend to the core question "does predictive policing work to reduce crime occurrence?"One might hope that this collection of work should offer a decisive answer to this key question. While discussing the theoretical background, mathematical results, and sociological concerns pertaining to predictive policing tools does help to understand the nature of these tools and their potential use-cases, controlled field-trials of predictive policing tools should give the ultimate answer to the problem at the core of this thesis. However, the problem with these empirical studies is that they fundamentally do not agree with each-other. That is, there are as many studies in support of predictive policing as there are studies claiming that predictive policing has had no noticeable impact on crime occurrence. This final section looks to discuss these studies on predictive policing, in an attempt to assess why it is that we still seem to know very little about these tools.

A number of authors have made a case against the use of predictive policing, on the grounds that these tools simply do not seem to reduce crime. Though not examining exactly the same technology as discussed in this project, a 2014 study published by the RAND Corporation assessed a predictive policing tool known as "Risk Terrain Modelling" (RTM), through a 29-week study on property crime levels in the city of Shreveport, Louisiana [17]. When comparing officers using RTM versus a control group of officers using traditional crime deterrence techniques, the research team behind the report comments succinctly that "the [RTM] program did not generate a statistically significant reduction in property crime" [17]. The authors of the report acknowledge that to conclusively answer the question of whether predictive policing works or not, studies far deeper and broader in scope must be performed on these tools before strong conclusions can be made about their efficacy:

"What is not clear, however, is whether distributing hot spot maps based on predictive analytics will lead to significant crime reductions compared with distributing hot spot maps based on traditional crime mapping techniques. In the case of the Shreveport predictive policing experiment, we did not see statistical evidence to this effect."

The study by the RAND Corporation is not the only piece of empirical work on predictive policing with inconclusive results. A 2016 study published in the *Journal of Experimental Criminology* found that in examining the Chicago Police Department's use of another separate predictive policing tool designed to predict gun violence, the "Strategic Subjects List" (SSL), the tool was again inconclusive in its results [41]. The authors of this report comment that while the SSL did result in more violent offenders being arrested when compared against control data, the number of violent homicides within Chicago over the course of the study did not decrease due to the SSL. Further, the authors theorize that the SSL only served to put officers into contact with already known violent offenders, that the officers would have been in contact with regardless [41]. Agreeing with the RAND Corporation, the authors of this report comment that more work must be done, as their results are fundamentally inconclusive.

But what of those in support of the predictive policing project? As mentioned prior, there are those authors of empirical studies who argue that predictive policing tools have in fact had noticeable impacts on overall crime rates in certain cities. In contrast to the 2014 study performed by the RAND Corporation on Risk Terrain Modelling (RTM) [17], a 2010 study performed by Piza et al. on the use of RTM to predict violent crimes in Newark, New Jersey, found that RTM drastically reduced the number of violent crimes, by 19% between the years 2006 and 2009 [36]. While this initially seems remarkable, upon closer reading the data in the paper is not controlled for any other factors, nor has the paper been peer-reviewed. Potentially more damming, the authors of this paper are in fact the creators of RTM, a fact that is not mentioned anywhere within [36]. It is for these reasons, that the findings of this report are, at the very least, somewhat suspect.

This report by Piza et al. is not the only empirical study in support of predictive policing. Potentially the most popular study of this type was published by Mohler et al. in 2015 [29]. This work assesses the use of "Epidemic-Type Aftershock Sequence" (ETAS) crime forecasting to predict and pre-empt crime by both the Los Angeles Police Department (LAPD) and the Kent Police Department (KPD). Through controlled experiments, the authors of this work found that by using ETAS crime forecasting, overall crime in both jurisdictions was reduced by 7.4%. While this result seems decisive - that predictive policing does in fact have positive impacts on crime rates - this paper is not without critique. Every co-author of [29] either holds stock in, or is a co-founder of PredPol, the predictive policing firm using ETAS in proprietary software. When examining the collection of empirical studies assessing the efficacy of predictive policing tools, the results are decidedly unclear. Many studies have found that predictive policing has had no real impact on crime, whereas other studies in support of predictive policing have been authored by suspect groups. In many ways, we are nowhere close to understanding predictive policing as a tool used in the real world.

It is at this point that potentially the most fundamental flaw with our current understanding and use of predictive policing must be underlined: there are (at the time of writing) no true answers to the question "Does predictive policing work to reduce crime occurrence?" From their literature review mentioned earlier, sociologists Meijer and Wessels comment that of the papers reviewed, there is not enough peer-reviewed empirical scholarship on predictive policing tools to determine absolutely whether they work or whether they do not: There exists only anecdotal evidence supporting both the perceived benefits and drawbacks of predictive policing [28]. While many have argued both for the benefits and drawbacks of predictive policing, relatively few have run empirical trials examining the use of predictive policing tools in the field to determine their efficacy. Once the vast amount of rhetoric both supporting and opposing predictive policing has been stripped away, there is nothing left. Looking back at the decade in which predictive policing has emerged as an applied policing practice, we are no closer to understanding whether this tool works, or whether it has done remarkably little at preventing crime.

Chapter 7

Conclusion

Predictive policing is a remarkably complex topic. This project has aimed to describe and understand modern predictive crime modelling tools at their most fundamental level, though this work is only scratching the very surface of the greater picture of predictive policing. Assessing the effectiveness of any predictive policing algorithm necessitates understanding the spatial dynamics of crime at a fundamental level.

Crime, regardless of environment, unambiguously displays distinct spatial patterns at micro-geographic resolutions. From the work of empirical criminologists, as discussed in Chapter 2, we know that across cities and countries, crime appears to self-arrange into strong spatial hotspots, that appear remarkably consistent over time periods ranging from weeks to years [49]. Criminologist David Weisburd takes these observations further, finding that across a number of different cities urban crime is similarly concentrated in space [49]. From this work Weisburd proposes a "physical law of crime concentrations" that should hold in all urban environments, promoting complex yet consistent dynamics underlying criminal behaviour. While general, this work provides good motivation for a mathematical model of criminal behaviour, work engaged by Martin B. Short et al. in their initial work that served to motivate this entire project: "A Statistical Model of Criminal Behaviour" [44].

Within [44], Short et al. attempt to distill the patterns in crime, as observed by empirical criminologists and discussed in Chapter 2, into a discrete agent-based mathematical model of criminal behaviour. Simulated criminals move throughout a finite grid of potential addresses to burglarize, deciding whether or not to commit crimes based on the "criminal attractiveness" of a given address, an arbitrary number quantifying how attractive a potential site is to a potential criminal with respect to all other potential sites. From this discrete agent-based model, Short et al. derived a continuous PDE model of criminal behaviour, to model both criminal attractiveness and criminal density throughout a given problem domain. These two models, dubbed the "discrete" and "continuous" models of criminal behaviour and discussed in Chapters 3 and 4, respectively, serve as the foundation of this project. While other models of criminal behaviour exist and are in use in predictive policing software today [38], these models represent one avenue by which to examine and discuss the mechanics of predictive policing, and the ways in which predictive policing tools might fare in the real-world.

On examining simulations based on both the discrete and continuous models of criminal behaviour as discussed by Short et al., we see that the results are somewhat varied. While on one hand results from both models, as discussed in Chapter 5 generally agree with existing results from literature (namely, from [44, 46]), it is questionable at best whether these simulated results agree with real-world crime dynamics as studied by environmental criminologists. Of particular note is that all simulations performed using the discrete model of criminal behaviour do not agree particularly well with Weisburd's physical law of crime concentrations [51], casting doubt on whether these models can accurately represent real-world crime dynamics. All this being said, upon a deeper review of literature concerning the implementation of predictive policing tools within police departments, the issue of model accuracy may be completely superseded by issues concerning the implementation of predictive policing tools.

Upon considering how predictive policing must be integrated into the real world, a number of key challenges arise. Foremost, to function in the real world any predictive policing software must be "tuned" using historic policing data, to replicate patterns present in this historic data. Given that essentially all historic policing data is marred with both false negatives and false positives, thus disproportionately implicating marginalized populations in crime occurrence [12], any predictive algorithm trained on this flawed data can at best replicate these flawed trends. Further, predictive policing tools are essentially only able to generate suggestions on where crime will occur. It is up to the individual police officer to interpret and act on any suggestions from a predictive algorithm. While great emphasis has been placed on the objectivity of predictive policing algorithms, to-date there is little research on how officers use these algorithms in the field [22]. These concerns aside, the small collection of peer-reviewed empirical studies on predictive policing performed to-date have been either inconclusive, or have found that predictive policing has had no noticeable effect on overall crime rates in the jurisdiction of study [28]. Any comments on the mathematics of predictive policing aside, there are a number of serious issues with this technology at the most fundamental level.

7.1 Potential Future Work

Concerning environmental criminology, while a great deal of work has been done to justify Weisburd's physical law of crime concentration [51], additional empirical data could be compiled on crime rates in more diverse contexts. Weisburd motivates his physical law using crime data from a handful of cities in the United States as well as from Tel Aviv, Israel. What remains to be shown is if this law holds in other cities and countries across the globe. Further, a majority of environmental criminology literature on the spatial distribution of crime focuses almost exclusively on the case of urban crime [51]. The spatial dynamics of crime in rural contexts have been left relatively unexplored, and deserve further inquiry.

As well, the models of criminal behaviour discussed within this project are concerned exclusively with residential burglary. The spatial dynamics of other forms of crime such as motor vehicle theft, assault, and commercial burglary are decidedly more complex, and as a result the discrete and continuous models of residential burglary discussed in this project are insufficient to address these other forms of crime. Potentially, models of criminal behaviour tuned to these separate forms of crime could be constructed in much the same way as the models shown here.

Turning specifically to the discrete model of criminal behaviour discussed in this project, the jarring issue of this model disagreeing to some extent with Weisburd's physical law of crime concentrations must be addressed. Potentially, by experimenting with additional choices of parameters outside of those used by Short et al. in [44], or by fundamentally altering the model, this can be achieved. Additionally, it remains unclear if this model could be tuned to historic data to potentially replicate real-world crime occurrence. By experimenting with model parameters obtained from historic crime data, in addition to varying the grid shape to potentially be non-uniform, it may be possible to better simulate the crime dynamics of a real-world city. Regardless, based on the preliminary work here, the discrete model of criminal behaviour does not appear to stand on its own when tasked to accurately model real-world criminal behaviour

Examining now the continuous model of criminal behaviour, the glaring challenge is to compute error-controlled numerical solutions to the two-dimensional version of this model. BACOL and BACOLI are only able to compute error-controlled numerical solutions to systems of PDEs in one spatial dimension and one time dimension: While it is potentially possible to perform a spatial discretization on the two-dimensional variant of this model, leading to a large system of PDEs in one spatial dimension, an implementation of this approach has proved elusive. Potentially, it would be a simpler first step to consider constrained versions of the continuous model of criminal behaviour in two spatial dimensions first, such as assuming radial symmetry throughout the problem domain, though such sub-problems have been left unconsidered here. Another potential problem of interest concerning the continuous model, is whether it is possible to obtain numerical solutions to this model with crime hotspots that are dynamic in space. Potentially by varying model parameters over time, or by adding an extra time-dependant factor to the model this could be achieved to agree with behaviour from the discrete model of criminal behaviour. In short, it is still unclear both the entire range of potential behaviours that the continuous model can exhibit, and whether the continuous model can be fit to historic policing data to replicate real-world crime dynamics.

Considering how predictive policing tools interface with existing police departments, there are two major questions that require more work to fully understand: One, exactly how police officers understand and actualize predictions made by software into their policing routines and as a direct consequence, two, whether predictive policing tools have any substantive effect on crime reduction. Due to the often difficult nature of studying day-to-day police operations and publishing these results, empirical peer-reviewed studies concerning predictive policing's use in the field are remarkably few and far between. The simplest way to truly understand the utility of this technology is by performing further controlled field trials to assess whether predictive policing has any noticeable impact on crime reduction.

7.2 Can We Predict Crime? Should We Predict Crime?

To the potential outsider, predictive policing may appear to solve a number of difficult problems concerning both police departments and police culture today. In the face of mounting criticisms of an overall lack of police accountability, predictive policing has been pitched by police departments as a truly objective crime prevention measure, removing any individual bias present in police officers in order to achieve a more just policing system. While predictive policing appears to have a solid foundation in both empirical criminology and applied mathematics, upon a closer examination, predictive policing tools lack the degree of rigour and specificity that proponents of predictive policing claim. Although crime displays definite trends in both time and space, it appears to be fundamentally too chaotic for us to currently understand.

Throughout the entirety of this project I have been concerned with questioning our ability to predict and understand crime. Potentially however, this line of inquiry is fundamentally at fault. Predictive policing acts as a hopeful solution to the issue of our "natural predisposition to committing mayhem" [39]. Perhaps it is wrong though, to think that by surgically addressing crime before it even occurs is the most effective strategy for "crime reduction". Our time may better be served by addressing the root inequalities that cause people to commit crimes, rather than trying to anticipate criminal behaviour in the first place. Through the work shown here neither predictive policing nor any similar tools appear to be adequate solutions to our modern crime problem. It is perhaps time then, to consider an altogether novel approach to the way that we envisage crime reduction today.

Bibliography

[1] Mark Andrejevic. “To Preempt a Thief”. In: *International Journal of Communication* 11.1 (2017), pp. 879–896.

[2] Tom Arsenault, Tristan Smith, and Paul Muir. “Superconvergent Interpolants for Efficient Spatial Error Estimation in 1D PDE Collocation Solvers”. In: *Canadian Applied Math Quarterly* 17.3 (2009), pp. 409–431.

[3] Harvin Bhathal. *Canada’s Predictive Policing Tech Is Poorly Regulated Under AI Policy*. Ed. by truthout. June 2022. URL: https://truthout.org/articles/canadas-predictive-policing-tech-is-poorly-regulated-under-ai-policy/.

[4] Data for Black Lives. *What is PREDICTIVE POLICING? — Crash Course: Data for Black Lives — Episode 4, Season 1*. Youtube. Jan. 2017. URL: https://www.youtube.com/watch?v=eFt3eCR5pJ4.

[5] M. Camacho-Collados and F. Liberatore. “A Decision Support System for predictive police patrolling”. In: *Decision Support Systems* 75 (2015), pp. 25–37.

[6] Janet Chan and Lyria Bennett Moses. “Is Big Data challenging criminology?” In: *Theoretical Criminology* 20.1 (2015), pp. 1–19.

[7] Derek B. Cornish and Ronald V. Clarke. *The rational choice perspective, in Environmental Criminology and Crime Analysis*. Ed. by Richard Wortley and Michael Townsley. 2nd. 2016, pp. 29–61.

[8] Carl De Boor. “Package for Calculating with B-Splines”. In: *SIAM Journal of Numerical Analysis* 14.3 (1977), pp. 441–472.

[9] Chicago Police Department. *Strategic Subject List - Historical Data*. [Dataset and brief history maintained by the CPD pertaining to the discontinued Strategic Subject List Program.] 2017. URL: https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List-Historical/4aki-r3np/about_data.

[10] Ángel Diaz. *Data-driven policing’s threat to our constitutional rights*. Ed. by Brookings. Sept. 2021. URL: https://www.brookings.edu/articles/data-driven-policings-threat-to-our-constitutional-rights/.

[11] John E. Eck, Jeffrey S. Gersh, and Charlene Taylor. *Finding Crime Hot Spots Through Repeat Address Mapping, in Analyzing Crime Patterns: Frontiers of Practice*. Ed. by Victor Goldsmith et al. Sage Publications, 2000, pp. 49–63.

[12] Delbert S. Elliott. *Lies, Damn Lies and Arrest Statistics*. Tech. rep. CSPV-015. Boulder, Colorado, United States of America: Center for the Study and Prevention of Violence, University of Colorado, 1995.

[13] Peter K. Enns. "The Public's Increasing Punitiveness and Its Influence on Mass Incarceration in the United States". In: *American Journal of Political Science* 58.4 (2014), pp. 857–872.

[14] Michael T Heath. *SCIENTIFIC COMPUTING: An Introductory Survey, Second Edition*. New York: McGraw Hill Higher Education, 2002, pp. 422–436, 447–456.

[15] Will Douglas Heaven. *Predictive policing algorithms are racist. They need to be dismantled.* Ed. by MIT Technology Review. July 2020. URL: `https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/`.

[16] Will Douglas Heaven. *Predictive policing is still racist-whatever data it uses*. Ed. by MIT Technology Review. Feb. 2021. URL: `https://www.technologyreview.com/2021/02/05/1017560/predictive-policing-racist-algorithmic-bias-data-crime-predpol/`.

[17] Priscillia Hunt, Jessica Saunders, and John S. Hollywood. *Evaluation of the Shreveport Predictive Policing Experiment*. Tech. rep. Santa Monica, California, United States of America: RAND Corporation, 2014.

[18] Brian Jordan Jefferson. "Predictable Policing: Predictive Crime Mapping and Geographies of Policing and Race". In: *Annals of the American Association of Geographers* 108.1 (2018), pp. 1–16.

[19] S. D. Johnson, K. Bowers, and A. Hirschfield. "New Insights Into the Spatial and Temporal Distribution of Repeat Victimization". In: *The British Journal of Criminology* 37.2 (1997), pp. 224–241.

[20] Shane D. Johnson et al. "Space-Time Patterns of Risk: A Cross National Assessment of Residential Burglary Victimization". In: *Journal of Quantitative Criminology* 23 (2007), pp. 201–219.

[21] John I. Kituse and Aaron V. Cicourel. "A Note on the Uses of Official Statistics". In: *Social Problems* 11.2 (1963), pp. 131–139.

[22] Nick Lally. ""It makes almost no difference which algorithm you use": on the modularity of predictive policing". In: *Urban Geography* 43.9 (2022), pp. 1437–1455.

[23] Grant Lamond. "What is a Crime?" In: *Oxford Journal of Legal Studies* 27.4 (2007), pp. 609–632.

[24] Lynn Langton et al. *Victimizations Not Reported to the Police, 2006-2010*. U.S. Department of Justice Special Report. Aug. 2012.

[25] *Letters to the Editor: Boycott collaberations with police.*

[26] E. S. Levine et al. "The New York City Police Department's Domain Awareness System". In: *Informs Journal on Applied Analytics* 47.1 (2017), pp. 70–84.

[27] Kristian Lum and William Isaac. "To Predict and Serve?" In: *Significance* 13.5 (2016), pp. 14–19.

[28] Albert Meijer and Martijn Wessels. "Predictive Policing: Review of Benefits and Drawbacks". In: *International Journal of Public Administration* 42.12 (2019), pp. 1031–1039.

[29] G. O. Mohler et al. "Randomized Controlled Field Trials of Predictive Policing". In: *Journal of the American Statistical Association* 110.512 (2015), pp. 1399–1411.

[30] Nathan Munn. *Police Across Canada Are Using Predictive Policing Algorithms, Report Finds*. Ed. by VICE. Sept. 2020. URL: `https://www.vice.com/en/article/k7q55x/police-across-canada-are-using-predictive-policing-algorithms-report-finds`.

[31] Patricia O'Brien. "Crime and Punishment as Historical Problem". In: *Journal of Social History* 11.4 (1978), pp. 508–520.

[32] *A Description of DASSL: A Differential/Algebraic System Solver*. Vol. SAND82-8637. Sandia Report. Montreal, Canada: Sandia National Laboratories, 1982, pp. 1–10.

[33] Jack Pew, Zhi Li, and Paul Muir. "Algorithm 962: BACOLI: B-spline Adaptive Collocation Software for PDEs with Interpolation-Based Spatial Error Control". In: *ACM Transactions on Mathematical Software* 42.3 (2016), 25:1–25:17.

[34] Jack Pew et al. *Performance Analysis of Error-Control B-spline Gaussian Collocation Software for PDEs*. Tech. rep. Halifax, Nova Scotia, Canada: Saint Marys University, Apr. 2018.

[35] Glenn L. Pierce, Susan Spaar, and LeBaron R. Briggs. *The Character of Police Work: Strategic and Tactical Implications*. Tech. rep. Norman, Oklahoma, United States of America: Center for Applied Social Research, University of Oklahoma, Nov. 1988.

[36] Eric L. Piza, Leslie W. Kennedy, and Joel M. Caplan. *Risk Clusters, Hotspots, and Spatial Intelligence: Risk Terrain Modeling as an Algorithm for Police Resource Allocation Strategies*. Tech. rep. New Brunswick, New Jersey, United States of America: Rutgers Center on Public Security, June 2010.

[37] PredPol. *Predictive Poliing Technology: The PredPol Algorithm*. Retrieved from: https://web.archive.o Date of archival: May 13, 2020. 2020. URL: `https://www.predpol.com/technology/`.

[38] David Robinson and Logan Koepke. *Stuck in a Pattern: Early evidence on "predictive policing" and civil rights*. Tech. rep. Washington D.C., United States of America: Upturn, Aug. 2016.

[39] Mitchell P. Roth. *An Eye for an Eye: A Global History of Crime and Punishment*. Reaktion Books Ltd., 2014, p. 18.

[40] Ajay Sandhu and Peter Fussey. "The 'uberization of policing'? How police negotiate and operationalise predictive policing technology". In: *Policing and Society* 31.1 (2021), pp. 66–81.

[41] Jessica Saunders, Priscillia Hunt, and John S. Hollywood. "Predictions put into practice: a quasi-experimental evaluation of Chicago's predictive policing pilot". In: *Journal of Experimental Criminology* 12.1 (2016), pp. 347–371.

[42] Robert C. Scott. “Deemphasize Punishment, Reemphasize Crime Prevention”. In: *Federal Sentencing Reporter* 20.5 (2008), pp. 299–303.

[43] Lawernce W. Sherman, Patrick R. Gartin, and Michael E. Buerger. “Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place”. In: *Criminology* 27.1 (1989), pp. 27–55.

[44] M. B. Short et al. “A Statistical Model of Criminal Behavior”. In: *Mathematical Models and Methods in Applied Sciences* 18.Suppl. (2008), pp. 1249–1267.

[45] Matt Stroud. *Heatlisted*. Ed. by The Verge. May 2021. URL: `https://www.theverge.com/c/22444020/chicago-pd-predictive-policing-heat-list`.

[46] W.H. Tse and M. J. Ward. “Hotspot Formation and Dynamics for a Continuum Model of Crime”. In: *European Journal of Applied Mathematics* 27.3 (2016), pp. 583–624.

[47] Craig D. Uchida. “A National Discussion on Predictive Policing: Defining Our Terms and Mapping Successful Implementation Strategies”. In: NCJ 230404. National Institute of Justice. Los Angeles, California, United States of America, Nov. 2009.

[48] R. Wang, P. Keast, and P. Muir. “BACOL: B-Spline Adaptive COLlocation Software for 1-D Parabolic PDEs”. In: *ACM Transactions on Mathematical Software* 30.4 (2004), pp. 454–470.

[49] David Weisburd. “The Law of Crime Concentration and the Criminology of Place”. In: *Criminology* 53.2 (2015), pp. 133–157.

[50] David Weisburd and Shai Amram. “The law of concentrations of crime at place: the case od Tel Aviv-Jaffa”. In: *Police Practice and Research: An International Journal* 15.2 (2014), pp. 101–114.

[51] David Weisburd et al. *Place Matters: Criminology for the Twenty-First Century*. Cambridge University Press, 2016.

[52] Richard Wortley and Michael Townsley. *Environmental Criminology and Crime Analysis*. 2nd. Routledge, 2016.

Appendix A

Python Script Implementing Short et al.'s Discrete Model of Residential Burglary

```
"""
Discrete Residential Burglary Simulation

This is an implementation of the discrete model of residential burglary as
introduced in 'A Statistical Model of Criminal Behavior' (M.B. Short et al.,
2008).

Model parameters can be set in the 'varInit' method, plot parameters can be set
in the 'genPlot' method. The overall simulation is saved as a .gif file to the
working directory by default.
"""

import numpy as np
import numpy.random as random
import matplotlib.pyplot as plt
from matplotlib.animation import FuncAnimation

class CrimeSite:
    """
    A class used to represent an individual crime site (s) in the overall model,
    containing all necessary values associated with the crime site (s)

    Attributes
    ----------
    loc : tuple
        The coordinates of the site in the overall model's grid
    A0 : double
        The value of the static component of criminal attractiveness associated
```

```
    with the site
Bst : double
    The value of the dynamic component of criminal attractiveness associated
    with the site at the current time-step (initialized by default to 0)
BstNext : double
    A buffer value, holding the value of the dynamic component of criminal
    attractiveness associated with the site for the next time step
    (initialized to 0)
Nst : int
    The number of criminals residing at the site at the beginning of the
    current time-step (initialized by default to 0)
NstNext : int
    A buffer value, holding the number of criminals to reside at the site
    at the beginning of the next time-step (initialized to 0)
Est : int
    The number of criminal events that occured at the site over the duration
    of the current time-step (initialized to 0)

Methods
-------
setBstNext(setVal)
    Sets the value of BstNext
decNst()
    Decrements the value of Nst by 1
incNstNext()
    Increments the value of NstNext by 1
incEst()
    Increments the value of Est by 1
getLoc()
    Returns the value of loc
getAst()
    Returns the overall criminal attractiveness Ast=A0+Bst
getBst()
    Returns the value of Bst
getBstNext()
    Returns the value of BstNext
getNst()
    Returns the value of Nst
getNstNext()
    Returns the value of NstNext
getEst()
    Returns the value of Est
copyBuffer()
    Copies the values from BstNest and NstNext to Bst and Nst respectively,
    then zeros BstNext, NstNext and Est
```

```
"""

def __init__(self, loc, A0, Bst=0, Nst=0):
    """
    Parameters
    ----------
    loc : tuple
        The coordinates of the site in the overall model’s grid
    A0 : double
        The value of the static component of criminal attractiveness
        associated with the site
    Bst : double, optional
        The value of the dynamic component of criminal attractiveness
        associated with the site at the current time-step (initialized by
        default to 0)
    Nst : int, optional
        The number of criminals residing at the site at the beginning of the
        current time-step (initialized by default to 0)
    """

    self.loc = loc
    self.A0 = A0
    self.Bst = Bst
    self.BstNext = 0
    self.Nst = Nst
    self.NstNext = 0
    self.Est = 0

def setBstNext(self, setVal):
    """
    Sets the value of BstNext.

    Parameters
    ----------
    setVal : double
        the value to set to BstNext
    """
    self.BstNext = setVal

def decNst(self):
    """
    Decrements the value of Nst by 1.
    """
    self.Nst -= 1
```

```
def incNstNext(self):
    """
    Increments the value of NstNext by 1.
    """
    self.NstNext += 1

def incEst(self):
    """
    Increments the value of Est by 1.
    """
    self.Est += 1

def getLoc(self):
    """
    Returns the value of loc.
    """
    return self.loc

def getAst(self):
    """
    Returns the overall criminal attractiveness Ast=A0+Bst.
    """
    Ast = self.A0 + self.Bst
    return Ast

def getBst(self):
    """
    Returns the value of Bst.
    """
    return self.Bst

def getBstNext(self):
    """
    Returns the value of BstNext.
    """
    return self.BstNext

def getNst(self):
    """
    Returns the value of Nst.
    """
    return self.Nst

def getNstNext(self):
    """
```

```
        Returns the value of NstNext.
        """
        return self.NstNext

    def getEst(self):
        """
        Returns the value of Est.
        """
        return self.Est

    def copyBuffer(self):
        """
        Copies the values from BstNest and NstNext to Bst and Nst respectively,
        then zeros BstNext, NstNext and Est.
        """
        self.Bst = self.BstNext
        self.BstNext = 0
        self.Nst = self.NstNext
        self.NstNext = 0
        self.Est = 0

class CrimeGrid:
    """
    A class used to represent the overall grid of crime sites in the model.
    Each crime site is represented as an instance of the 'CrimeSite' class,
    stored in a numpy ndarray representing the grid of lattice points.

    Attributes
    ----------
    shape : tuple
        The shape of the overall grid of crime sites
    modelGrid : numpy.narray
        A matrix with entries of a specified shape filled with datatype
        'CrimeSite', representing the grid of lattice points in the model

    Methods
    -------
    runSiteCrimes(pos)
        Simulates n potential criminal events at a given CrimeSite indicated by
        pos, where n is the number of criminals at the given CrimeSite (or, the
        value of Nst at the given crimeSite) and pos is a tuple corresponding to
        the loc value of the given crimeSite. The probability of a criminal
        event being sucessful is given in 'A Statistical Model of Criminal
        Behavior' (M.B. Short et al., 2008). For each successful criminal
        event, decrement the value of Nst and increment the value of Est at the
```

```
    given CrimeSite. For each unsuccessful criminal event, decrement the
    value of Nst at the given CrimeSite and increment the value of NstNext
    at an adjacent crime site, determined using a random walk formula given
    in (M.B. Short et al., 2008).
runAllCrimes()
    Runs the method 'runSiteCrimes(pos)' on all CrimeSites in modelGrid.
compBstNext(pos)
    Computes the value of the dynamic component of criminal attractiveness
    for the next time-step (BstNext) for a given CrimeSite indicated by pos,
    a tuple corresponding to the loc value of the given CrimeSite. This
    value BstNext is computed using a formula given in 'A Statistical Model
    of Criminal Behavior' (M.B. Short et al., 2008), and once computed is
    stored in the given CrimeSite.
compAllBstNext()
    Runs the method 'compBstNext(pos)' on all CrimeSites in modelGrid.
genCriminals(pos)
    Generates a new criminal at a given CrimeSite indicated by pos with
    probability Gamma, where pos is a tuple corresponding to the loc value
    of the given CrimeSite and Gamma is a global model parameter specified
    in the 'varInit' method. To simulate the generation of a criminal,
    increment the NstNext parameter of the given CrimeSite.
genAllCriminals()
    Runs the method 'genCriminals(pos)' on all CrimeSites in modelGrid.
copyAllBuffers()
    For all CrimeSites in modelGrid, runs the method 'copyBuffer()' to
    initialize the necessary variables for the next time-step.
returnStepData()
    Returns an narray containing the current dynamic component of each
    site's overall attractiveness.
printGridInfo(printShape=False, printAst=False, printBst=False,
              printBstNext=False, printNst=False, printNstNext=False,
              printEst=False)
    A helper method to print out  information pertaining to the grid's
    current state. Parameters represent the possible pieces of information
    to print, and are all set to 'False' by default.
"""

def __init__(self, shape, A0, Bst):
    """
    Parameters
    ----------
    shape : tuple
        The shape of the overall model grid
    A0 : double
        The value of the static component of criminal attractiveness
```

```
            associated with the grid. Though A0 can vary between grid sites, in
            this implementation it is fixed for all grid sites
        Bst : Double
            The initial value of the dynamic component of criminal
            attractiveness associated with each site. Though this initial value
            can vary between grid sites, in this implementation it is fixed for
            all grid sites
        """

        self.shape = shape
        self.grid = np.empty(shape=shape, dtype=CrimeSite)
        for i in range(shape):
            self.grid[i] = CrimeSite(i, A0, Bst=Bst)

    def runSiteCrimes(self, pos):
        """
        Simulates n potential criminal events at a given CrimeSite indicated by
        pos, where n is the number of criminals at the given CrimeSite (or, the
        value of Nst at the given crimeSite) and pos is a tuple corresponding to
        the loc value of the given crimeSite. The probability of a criminal
        event being sucessful is given in 'A Statistical Model of Criminal
        Behavior' (M.B. Short et al., 2008). For each successful criminal
        event, decrement the value of Nst and increment the value of Est at the
        given CrimeSite. For each unsuccessful criminal event, decrement the
        value of Nst at the given CrimeSite and increment the value of NstNext
        at an adjacent crime site, determined using a random walk formula given
        in (M.B. Short et al., 2008).

        Parameters
        ----------
        pos : tuple
            The coordinates corresponding to the given CrimeSite with which to
            simulate n potential criminal events
        """

        while self.grid[pos].getNst() > 0:

            # Decrementing the total number of criminals at site
            self.grid[pos].decNst()

            # Generating a random number to compare against the probability of
            # a criminal event
            rand = random.rand()
            crimeProb = 1 - np.exp((-1)*(self.grid[pos].getAst())*(dt))
```

```
        if rand < crimeProb:
            # Successful criminal event. Increment the number of criminal
            # events at site, Est
            self.grid[pos].incEst()

        else:
            # Unsucessful criminal event. Begin by generating an array of
            # probabilities representing how likely the criminal agent will
            # be to move to any adjacent site
            adjSites = getAdjVals(self.grid, pos)
            AstSum = 0
            for site in adjSites:
                AstSum += site.getAst()

            moveProbArray = []
            for site in adjSites:
                moveProb = site.getAst()
                moveProb = moveProb / AstSum
                moveProbArray.append(moveProb)

            # Based on the array of movement probabilties, move the criminal
            # agent to an adjacent site
            moveChoice = random.choice(adjSites, p=moveProbArray).getLoc()
            self.grid[moveChoice].incNstNext()

def runAllCrimes(self):
    """
    Runs the method 'runSiteCrimes(pos)' on all CrimeSites in modelGrid.
    """

    for site in self.grid:
        self.runSiteCrimes(site.getLoc())

def compBstNext(self, pos):
    """
    Computes the value of the dynamic component of criminal attractiveness
    for the next time-step (BstNext) for a given CrimeSite indicated by pos,
    a tuple corresponding to the loc value of the given CrimeSite. This
    value BstNext is computed using a formula given in 'A Statistical Model
    of Criminal Behavior' (M.B. Short et al., 2008), and once computed is
    stored in the given CrimeSite.

    Parameters
    ----------
    pos : tuple
```

```
            The coordinates corresponding to the given CrimeSite with which to
            compute BstNext
        """

        # Computing the sum of the adjacent Bst values
        adjSites = getAdjVals(self.grid, pos)
        adjBstSum = 0
        for site in adjSites:
            adjBstSum += site.getBst()

        # Computing BstNext
        BstNext = ( (1-eta)*self.grid[pos].getBst() + (eta/len(adjSites))* \
                  adjBstSum ) * (1-omega*dt) + theta*self.grid[pos].getEst()

        self.grid[pos].setBstNext(BstNext)

    def compAllBstNext(self):
        """
        Runs the method 'compBstNext(pos)' on all CrimeSites in modelGrid.
        """

        for site in self.grid:
            self.compBstNext(site.getLoc())

    def genCriminals(self, pos):
        """
        Generates a new criminal at a given CrimeSite indicated by pos with
        probability Gamma, where pos is a tuple corresponding to the loc value
        of the given CrimeSite and Gamma is a global model parameter specified
        in the 'varInit' method. To simulate the generation of a criminal,
        increment the NstNext parameter of the given CrimeSite.

        Parameters
        ----------
        pos : tuple
            The coordinates corresponding to the given CrimeSite with which to
            generate criminals at
        """

        rand = random.rand()
        if (rand < Gamma):
            self.grid[pos].incNstNext()

    def genAllCriminals(self):
        """
```

```
        Runs the method 'genCriminals(pos)' on all CrimeSites in modelGrid.
        """

        for site in self.grid:
            self.genCriminals(site.getLoc())

    def copyAllBuffers(self):
        """
        For all CrimeSites in modelGrid, runs the method 'copyBuffer()' to
        initialize the necessary variables for the next time-step.
        """

        for i in range(self.shape):
            self.grid[i].copyBuffer()

    def returnStepData(self):
        """
        Returns an narray containing the current dynamic component of each
        site's overall attractiveness.
        """

        stepData = np.empty(self.shape)
        for i in range(self.shape):
            stepData[i] = self.grid[i].getBst()

        return stepData

    def printGridInfo(self, printShape=False, printAst=False, printBst=False,
                      printBstNext=False, printNst=False, printNstNext=False,
                      printEst=False):
        """
        A helper method to print out  information pertaining to the grid's
        current state. Parameters represent the possible pieces of information
        to print, and are all set to 'False' by default.

        Parameters
        ----------
        printShape : boolean, optional
            Flag that indicates whether the shape of the model's grid should be
            printed or not. Set to false by default
        printAst : boolean, optional
            Flag that indicates whether the Ast values should be printed or not.
            Set to 'False' by default
        printBst : boolean, optional
            Flag that indicates whether the Bst values should be printed or not.
```

```
        Set to 'False' by default
    printBstNext : boolean, optional
        Flag that indicates whether the BstNext values should be printed or
        not. Set to 'False' by default
    printNst : boolean, optional
        Flag that indicates whether the Nst values should be printed or not.
        Set to 'False' by default
    printNstNext : boolean, optional
        Flag that indicates whether the NstNext values should be printed or
        not. Set to 'False' by default
    printEst : boolean, optional
        Flag that indicates whether the Est values should be printed or not.
        Set to 'False' by default
    """

    if (printShape == True):
        print("Grid shape:", self.shape)

    if (printAst == True):
        AstVals = []
        for site in self.grid:
            AstVals.append(site.getAst())
        print("Ast Values:", AstVals)

    if (printBst == True):
        BstVals = []
        for site in self.grid:
            BstVals.append(site.getBst())
        print("Bst Values:", BstVals)

    if (printBstNext == True):
        BstNextVals = []
        for site in self.grid:
            BstNextVals.append(site.getBstNext())
        print("BstNext Values:", BstNextVals)

    if (printNst == True):
        NstVals = []
        for site in self.grid:
            NstVals.append(site.getNst())
        print("Nst Values:", NstVals)

    if (printNstNext == True):
        NstNextVals = []
        for site in self.grid:
```

```
                NstNextVals.append(site.getNstNext())
            print("NstNext Values:", NstNextVals)

        if (printEst == True):
            EstVals = []
            for site in self.grid:
                EstVals.append(site.getEst())
            print("Est Values:", EstVals)

        print('---------------------------------------------------------------')

def varInit():
    """
    Method to initialize the necessary model parameters.
    - l denotes the length between adjacent CrimeSites
    - dt denotes the duration between time-steps
    - omega denotes the rate at which the dynamic component of criminal
      attractiveness decreases
    - eta denotes the measure of neighborhood effects on the dynamic component
      of criminal attractiveness
    - theta denotes the increase in attractiveness at a site due to a
      successful criminal event
    - Gamma denotes the probability of burglar generation at each site
    - A0 denotes the static component of criminal attractiveness at each site
    """
    global l, dt, omega, eta, theta, Gamma, A0
    l = 1.0
    dt = (1.0 / 100.0)
    omega = (1.0 / 15.0)
    eta = 0.03
    theta = 5.6
    Gamma = 0.002
    A0 = (1/30.0)

def getAdjVals(array, pos):
    """
    Method to return the adjacent values of a numpy array.

    Parameters
    ----------
    array : narray
        Array to chose adjacent values from
    pos : tuple
        Location in array to find adjacent values of
    """
```

```
    adjValues = []
    if (pos > 0 and pos < array.size):
        adjValues.append(array[pos-1])
    if (pos > -1 and pos+1 < array.size):
        adjValues.append(array[pos+1])

    return adjValues

def buildAnimation(data, gridSize, numFrames):
"""
Method to create the final simulation .gif from simulated crime data.

Parameters
----------
data : narray
Simulated crime data to create .gif file from
gridSize : tuple
Dimensions of the simulation grid
numFrames : int
The number of time-steps throughout the entire simulation
"""

    fig = plt.figure()
    ax = plt.axes(xlim=(0, gridSize), ylim=(0, 30))
    line, = ax.plot([], [], lw=3)

    def init():
        line.set_data([], [])
        return line,

    def animate(i):
        x = np.linspace(0, gridSize, gridSize)
        y = data[:,i]
        line.set_data(x, y)
        return line,

    anim = FuncAnimation(fig, animate, init_func=init, frames=numFrames, interval=20, bl


    anim.save('1DCrimeSimD.gif')

varInit()
gridSize = 100
numTimeSteps = 5000
```

```
modelGrid = CrimeGrid(gridSize, A0=A0, Bst=0.01)
outputData = np.empty(shape=(gridSize, numTimeSteps))

for t in range(numTimeSteps):

    modelGrid.runAllCrimes()
    modelGrid.compAllBstNext()
    modelGrid.genAllCriminals()
    modelGrid.copyAllBuffers()
    outputData[:,t] = modelGrid.returnStepData()

buildAnimation(outputData, gridSize=gridSize, numFrames=numTimeSteps)
```

Appendix A

Python Script Implementing BACOLI_PY to Solve the 1D Burgers' Equation

```
# This program defines the one-dimensional Burgers' equation along with
# initial and boundary conditions. It then calls bacoli_py to compute an
# approximate solution from t=0 to t=2. Finally, it plots the computed
# solution on the XYZ plane.

# Imports
import bacoli_py
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt
from matplotlib import cm

# Defining problem-dependent parameters
npde = 1

# Defining the the PDE system
def f(t, x, u, ux, uxx, fval):
    fval[0] = -1*u[0]*ux[0] + uxx[0]

    return fval

# Defining the left spatial boundary condition
def bndxa(t, u, ux, bval):
    bval[0] = u[0] - (1 / (1 - (t/4)))

    return bval

# Defining the right spatial boundary condition
```

```
def bndxb(t, u, ux, bval):
    bval[0] = u[0] - (1 / (1 + np.exp(1/2) - (t/4)))

    return bval

# Defining the initial condition
def uinit(x, u):
    u[0] = 1 / (1 + np.exp(x/2))

    return u

# Initializing a bacoli_py Solver object
solver = bacoli_py.Solver()

# Packing the functions as defined in the problem definition file and the
# number of PDE's into a bacoli_py ProblemDefinition object
problem_definition = bacoli_py.ProblemDefinition(npde, f=f, bndxa=bndxa,
                                                  bndxb=bndxb, uinit=uinit)

# Specifying the initial time, initial mesh, output points and output times
spa_pnt_final = 1
t_res = 100
x_res = 100
initial_time = 0.0
final_time = 2.0
initial_mesh = [0, spa_pnt_final]
t_outputs = np.linspace(0.001, final_time, t_res)
x_outputs = np.linspace(0.0, spa_pnt_final, x_res)

# Calling the solver to compute an approximate solution to the PDE
evaluation = solver.solve(problem_definition, initial_time, initial_mesh,
                          t_outputs, x_outputs, atol=1e-6, rtol=1e-6,
                          dirichlet=True)

# Assigning the solution component of evaluation to an narray
solution = evaluation.u

# Building an x,y mesh to plot the solution values on and assigning
# the solution values to zMesh
xRange = x_outputs
yRange = t_outputs
xMesh, yMesh = np.meshgrid(xRange, yRange)
zMesh = solution[0,:,:]

# Building and saving the finalized graph:
```

```
fig = plt.figure()
ax = plt.axes(projection='3d')

plot = ax.plot_surface(xMesh, yMesh, zMesh, cmap=cm.magma)

ax.set_title('Approx. Sol\'n. to 1D Burgers\' Equation')
ax.set_xlabel('x')
ax.set_ylabel('t')

plt.show()
```