Modeling Society Following a Nuclear Weapon of Mass Destruction (WMD) Event: An Agent-based Modeling Approach

William G. Kennedy
Department of Computational and Data Sciences
Center for Social Complexity
George Mason University
Fairfax, VA 22030
wkennedy@gmu.edu

Andrew Crooks
Department of Geography and RENEW Institute
University at Buffalo
Buffalo, NY 14261
atcrooks@buffalo.edu

Annetta Burger
Department of Computational and Data Sciences
Center for Social Complexity
George Mason University
Fairfax, VA 22030
aburger2@gmu.edu

Talha Oz
Department of Computational and Data Sciences
Center for Social Complexity
George Mason University
Fairfax, VA 22030
toz@gmu.edu

Xiaoyi Yuan
Department of Computational and Data Sciences
Center for Social Complexity
George Mason University
Fairfax, VA 22030
xyuan5@gmu.edu

Na Jiang
Department of Computational and Data Sciences
Center for Social Complexity
George Mason University
Fairfax, VA 22030
njiang8@gmu.edu

January 5, 2021

Abstract

While nuclear weapons of mass destruction exist, thankfully they have only been used in anger twice. Therefore, there is little know about how people will react to them. As a consequence of this unknown, we synthesized a hundred years of disaster research to build a model to explore this gap in our understanding of the social effects of a nuclear weapon of mass destruction (NWMD). By reviewing disaster literature, we argue that disasters, including a NWMD, should be viewed as a complex system of three parts (i.e., the physical, social and individual). These three parts inform an agent-based model on how society might react following a nuclear weapon of mass destruction. Specifically, the agent-based model captures the main properties of complex adaptive systems such as heterogeneity, webs of connections (i.e., social networks), relationships and interactions, and adaptations arising from individual actions and decisions. Our NWMD model represents the road network and weapon effects as part of the physical environment. It also includes synthesized individuals and their social environment through agents' social networks and emergent group dynamics after the event. This NWMD model supports the exploration of the effects of different agent behavior in times of disaster. In the base model, we characterized the response of victims of a nuclear WMD, first responders, and the rest of the population not directly impacted by the weapon. Such a model of the New York mega-city is poised to support additional studies of social effects of a nuclear WMD or disasters more generally.

***K*eywords** Synthetic Populations · Agent-Based Modeling, · New York · Geographical Information Systems · Social Networks · Weapon of Mass Destruction

Table of Contents

List of Figures

List of Tables

"*All models are wrong, but some are useful.*" George Box [1]
"*Think how hard physics would be if particles could think.*" Murray Gell-Mann [2]

1 Introduction

The Manhattan Project produced the first and only nuclear weapons ever used in warfare. The project was an scientific and engineering marvel. For the next several decades, a program of nuclear weapons testing advanced our understanding of these weapons and their effects. With the end of nuclear weapons testing in 1992, the science, engineering, and confidence in the reliability of these weapons has been certified based on extremely detailed, sub-atomic level modeling of the components of the weapons systems [3]. On the other hand, how people react to a nuclear weapon in not as well known in a scientific or in a practical sense comparable to the developed knowledge of the physical effects. The study of this "soft" side, the effects on people and their reaction, started immediately after the two weapons were detonated in August 1945 (e.g., [4]). The development of the understanding of the affected population's reaction to nuclear weapons, of course, can not be the subject of a nuclear testing program anything like the study of the physical effects due to ethical considerations.

The Defense Nuclear Threat Reduction Agency (DTRA, `https://www.dtra.mil`), a successor to the Manhattan Project, is the organization for the U.S. government responsible for understanding these weapons. The agency began to develop an understanding of the "soft side" of nuclear weapons through computer simulation (modeling) within the last decade (see, as an example, [5, 6]). To contribute to the understanding, with this report we document our project characterizing the reaction of a mega-city's population to a nuclear weapon. The objective of the research described here was to characterize the reaction of the population of a mega-city and surrounding region to a nuclear weapon of mass destruction (WMD) event. The characterization became a quantitative description of the numbers of people projected to react in a variety of ways during the first minutes to hours following a "small" nuclear WMD event (as will be shown in Section 4). This was accomplished through the development of an agent-based model (ABM) representing the physical environment, the population details, their social networks, their use of the infrastructure at an individual level, their routine behaviors before the event, and their reactive behaviors after the event (see Section 3).

The scope of this project was the development of a one-to-one, geographically accurate model of heterogeneous individuals' response to a nuclear WMD event. By building such a model, we can study the societal dynamics and the complex cascading behavior that emerge across time and space from the *bottom up*. This agent-based modeling methodology lets us develop a more complete understanding of societal consequences of a potential nuclear WMD event. We modeled individual behavior based not simply on statistical descriptions of "big data" but based on established social science theory. Four core ideas underlie our approach: people are not random; people have heterogeneous individual characteristics; people's behavior is driven by a hierarchy of goals and their observable environment; and societal behavior emerges from individual behaviors. We will elaborate on these ideas in the following paragraphs.

Beyond People as Random Variables: Since human behavior is still not well understood despite thousands of years of observation, there is a temptation to simulate behavior as a random variable. We believe that this technique is ill-advised [7]. Further, human decision-making is not solely the product of rational optimization, as modeled by rational choice theory. Instead, it is driven by the interactions of humans' bounded rationality [8] with their emotional state, social context, physical state and physiology, individual characteristics, and their history. We captured this by modeling people as individual agents, based on well-established theories of rational, emotional, socially-influenced, goal-directed behavior and with diverse, non-uniform characteristics and goals (e.g., [9] [10], [11], and [12] as will be discussed more in Section 3.3.5

Hierarchy of Goals: Humans do have many things in common, and one of the longest-surviving theories of behavior is Maslow's hierarchy of needs [12]. While the hierarchy is not rigid, the ordering has generally survived intact since originally written. People's top priorities are their physiological requirements; i.e., their immediate survival needs. If their immediate physical needs are met, they can consider near-term safety and security needs. Only when these are met can they consider general social needs. These first three levels – immediate survival needs, near-term safety, and near-term security needs - will be the primary and shared drivers of human behavior in responding in the short term to a nuclear WMD event. Although these are shared drivers, people's responses can be diverse and will be modeled as such, based on theories and data on individual differences.

Individual characteristics: Humans are diverse, and it is vital to capture this diversity in any model of human behavior. For example, different individuals will have different tolerances for risk, levels of altruism, and individual personality traits that are likely to affect their behavior in times of crisis. Individuals' social contexts will also drive their decision-making: a person without dependents may be more likely to evacuate alone by the fastest possible means, while parents may first attempt to reunite with their children before evacuating. Finally, broad demographic differences may also drive behavior: for example, in a society with ethnic tensions, people may be more likely to seek shelter with members of their own ethnic group. Our model will capture such individual characteristics, and how they drive behavior and decision-making.

Social Psychology and Behavior Within Society: People are social, influenced by and influencing others continuously. They seem to usually employ simple rules in their daily activities and would likely apply similar rules even in emergencies. Such simple rules can explain complex behaviors and have been applied in large-scale social simulations. Adaptive behavior by individuals in a community's population can take on a number of forms, from short-term adaptations to longer-term ones. For example, Crooks and Wise [13] showed how dramatic changes to the road network configuration after the 2010 Haiti earthquake affected the distribution of post-disaster aid, which led to changes in how individuals traveled to find aid based on what they knew about the network. The model that we built is based on sound social psychology theory, social behavioral data, and where possible, empirical data of similar types of events (See Sections 2, 3.3.5, and 4.1).

This report expands on these core ideas by synthesising material from various publications over the course of the project (see [14, 15, 16, 17, 18, 19, 20]). In the remainder of this report, we discuss how disaster science has grown over the last century and how this informs our modeling of society following a nuclear weapon of mass destruction (WMD) event (Section 3). After presenting the formal model we then discuss some results from our model in Section 4. Finally, we provide a discussion of this research including what we have learned and identify areas of further research (Section 5). Finally, for interested readers we provide a summery of research outputs and links to the code and data products resulting from this project Appendix A.

2 Background

While the focus of this research was on how society might respond from a nuclear WMD event, we would be reminisced if we did not first discuss disaster research and the role of complexity. Our rationale being that there has been century of disaster research which has been used to gain a greater understanding of how individuals, families, and social systems operate under extreme stress, how individuals and societies respond to disrupted social systems, and what can be done to aid those harmed by disasters. Studies of behavior in these disasters has provided insights into individual and community coping mechanisms (e.g., [21]). Over the recent decades, insights from research on disasters[1] have been institutionalized and applied in new strategies and tactics for mitigation, preparation and warning, emergency response and aid, and recovery (e.g., [27]) at local, national, and global levels to reduce the risk of disruption and harm from disasters. The increasing frequency of disasters caused by climate-change [28] has added a sense of urgency to disaster research while at the same time researchers in several fields, notably economics (e.g., [29]), socio-ecology (e.g., [30]), and human-coupled systems (e.g., [31]), have recognized the importance of framing human systems as complex adaptive systems and have begun to study their implications for global markets, climate change, and the organization of cities. The application of complexity theory and complex adaptive systems has led to new conceptual tools for explanation in sustainability and urban studies research such as adaptive capacity (e.g., [32, 33, 34]) and spatial clustering of socioeconomic groups (e.g., [35]), and a few notable works in disaster research have begun to explore its application to disasters both in discourse (e.g., [36]) and in practice (e.g., [37]).

Our project proposed an approach for exploring disasters using the lens of complex adaptive systems that we argue can organize theories and provide explanation of the interactions and adaptations observed in disasters. Complex adaptive systems are nonlinear dynamic systems in which the interactions between individual elements and actors lead to emergent behavioral patterns and adaptation which one might witness after a WMD event [38, 39, 40]. Nonlinear systems are those in which inputs and outputs of the system are not proportional to each other. In the family of complex systems, complex adaptive systems are distinguishable by processes of adaptation including learning carried out by actors who respond to changes both inside and outside of the system's boundaries. System properties such as learning and adaptation in general lead to dynamics that include emergent behavior, flows of information, and system shifts between stability and instability (i.e., in and out of equilibrium). The emergent behavior results from interactions between individual components or subsystems, feedback loops, and self-organization. Popular examples of complex adaptive systems include cities in urban studies, ecosystems in ecology, and ant colonies in biology. The view through lenses of complex adaptive systems in these areas have led to discoveries such as the patterns of power laws and scaling in cities [41], the importance of heterogeneity for resilience of ecosystems [40], adaptive cycles in ecology and societies [42], and the role of self-organization in evolutionary biology [43].

The theme of complexity was already evident in disaster research (e.g., [44, 45, 46, 47]), but complexity theory has only been directly applied in a few rare examples (e.g., [48, 49, 50]). Generally, theory, as used in disaster research, is built on case studies of disaster events and on statistics, both drawn from decades of data collection, not unifying theories. In fact, there are no unifying theories. Rather, we would argue that the field is dominated by middle-ground theories such as uniformities of societal patterns following sequence patterns [51], therapeutic adjustments in disaster [22], patterns of pre-disaster growth and decline continuing after a disaster [52], and social networks shrinking as disaster victims prioritize resources and energy [53]. The importance of a unifying general theory lies in its explanation of human behavior in disasters and the potential for prediction and knowledge of areas that could be affected to improve societal well-being. Overlaps of current theory in disasters and complex adaptive systems could point to the application of a new set of theories for application in disaster mitigation, preparation, management, and recovery as well as new methodological tools that being applied in the field of complexity science.

To determine whether applying theories of complexity and complex adaptive systems could support explanation of human behavior, within this report we will explore three complex adaptive systems as evident from theories in disaster studies. These systems are: 1) the physical system (Section 2.1.1), consisting of geological, biological, meteorological, ecological, and human-built systems (including nuclear weapons effects in Section 3.3.4); 2) the social system (Section 2.1.2), consisting of formal and informal socio-cultural structures and collective behavior; and 3) the individual system (Section 2.1.3), consisting of the actor agents and their cognition.

Before we do, some terms and concepts need explanation. Collective behavior can be understood as the aggregated behavior of individuals in informal groups, families, or formal organizations. Individual actors in socio-ecological systems (i.e., physical systems) and collective behavior (i.e., social systems) have their own bio-physical and cognitive systems that process information and emotion before any identifiable behaviors and actions. Individual, goal-driven

[1]Readers should note that there are a multitude of definitions of what constitutes a disaster, and it is not the purpose of this report to discuss that issue, as it would distract from the main topic. However, readers are referred to example definitions such as those covered in [22, 23, 24, 25, 26].

behavior drives actions and feeds bottom-up processes that reshape both the physical and social systems [8]. The social system is composed of individuals whose collective actions are aggregated into formal structures such as organizations and governments and informal behavior such as everyday actions evident in commuting, migration, and market demand. These social and cultural actions aggregate into larger forces that shape human and natural systems—for example, creating neighborhoods and setting aside wildlife sanctuaries. Physical dynamics in the geological, ecological, and human-built systems affect the conditions under which individual actors and groups behave, and the systems of collective behavior respond to and influence the shape of physical systems and the individual cognition of actors [54].

Each of these systems changes and adapts in response to internal processes and each other as part of a complex adaptive system of systems. The interactions and adaptations of these systems produce nonlinear relationships with properties of aggregation, feedback, self-organization, emergence, diversity or heterogeneity, and flows of information, resources, and energy. According to Holland [38], the decision-making behavior and adaptation of the individual entities in the system together build on and add to the complexity of the overall system, ultimately creating a "*whole that is greater than the sum of its parts*." Typically studied in separate disciplines, the integration of these systems into a complex adaptive system of systems may improve explanation of the phenomena and dynamics in disaster (such as a WMD event) that interact and cut across systems and suggest new theory and sources for data that support explanation of human behavior.

A systematic review of theory in the disaster literature as presented in this report will demonstrate the properties and dynamics of complex adaptive systems and how complexity theory is integral to understanding human behavior in disasters by addressing the interactions across systems (which we will demonstrate in Sections 3 and 4). For the purpose of this research, the main properties of complex adaptive systems are narrowed to the following: heterogeneity, webs of connections (such as social networks), relationships and interactions, and adaptations arising from individual actions, decisions, and learning (which relate back to our four core ideas first introduced in Section 1). As noted above, we will explore disaster theories in three intersecting complex adaptive systems: the physical system, the social system, and the individual actor's system using evidence from the disaster literature. The remainder of this background section is organized as follows. First we identify disaster theories that align with the three systems and test whether properties of complex adaptive systems exist (Section 2.1). We will then discuss the characteristics of a complex adaptive system and how complexity theory can be applied to disaster research in Section 2.2. Finally, we will summarize the findings and explore implications for future disaster research (Section 2.2.4) and how this leads to our model exploring how people might react in a WMD event (which is presented in Section 3).

2.1 Organization of Disaster Research as Three Systems

To find evidence that disasters can be viewed as complex adaptive systems, we focus our discussion on social science theories, frameworks, and models that explain how people on the ground behave before, during, and after a disaster as it relates to the physical, social, and individual systems. This is particularly important as it allows us to draw analogies to how people may react during a WMD event which unlike other disasters (e.g., earthquakes, fires, hurricanes) is not that common. Works specifically relating to emergency response management, risk management, and public communications in disasters were avoided because our research is targeting the underlying, bottom-up effects of human behavior in disasters rather than hierarchical, top-down, organized responses that eventually come to play. Also, we do not explicitly look at the interactions between human and technological systems as we find that this, in a sense, is often implied in the literature, rather than studied directly (e.g., [55, 56, 57, 58]). For those interested in research and theory on technological and organizational disasters and the interactions of individuals on the built system (i.e., infrastructure, etc.), we refer the reader to the works of such as [59, 60, 61, 62, 63, 64, 65]. In addition, there is a large body of work on current emergency and disaster management that includes literature reviews and curricula for the emergency management community (e.g., [66, 67, 68, 69]). The intent of this review is to demonstrate how disaster theories and frameworks have evolved so that we may understand the processes within disasters with respect to physical, social, and individual systems. Our goal is to lay the foundation for thinking of disasters in a complex adaptive systems framework (Section 2.2) and incorporate this into our framework and model of how society might respond to a WMD event (Section 3).

2.1.1 Physical Systems

The physical system, as defined in this report, consists of geological, biological, meteorological, ecological, and human-built systems, and the dynamics within this system have always been present in disaster research. Early disaster and catastrophe studies centered on extreme events that caused injury or the loss of life and property within the social and physical systems of cities and townships. These events were conceptualized as unforeseen, rare, and extreme interruptions in everyday life and social activity, and after a period of recovery and social change, activities would stabilize and return to normal (see [51, 70, 71, 72]). Research during this period can be categorized by studies on specific forms of disaster (e.g., bombing, explosion, earthquake, tornado, hurricane, fire, or flood) and their social and psychological impact on individuals and communities. Although work between the 1930s and the 1950s was in its early stages, Wallace [72] collected a body of disaster studies and completed an interdisciplinary survey with contributions from the fields of psychiatry, general medicine, psychology, sociology, anthropology, economics, and political science. During this early period of research, disasters were often understood in the context of a trigger event within separate, external processes such as natural geophysical and meteorological phenomena—e.g., earthquakes, hurricanes, and tornados [73]—manmade events such as famine, pestilence, war, or revolution [74].

Breaking away from this approach, White [75] argued that disasters resulted from interactions between acts of "nature: and acts of man." For example, White's [75] study of flooding disasters explored natural hazard and flood plain environmental features and their human occupation, corresponding social and economic policies, and behavioral adjustments to the flood plain. In a broader study of disruption from environmental extremes, Burton et al. [76] found disaster events resulted from a combination of physical and social processes in which communities adjusted their behavior based on perceptions of the environmental hazards. This was followed by the systemic framework of a hazardousness of place in a regional ecology [77], and a socio-ecological system model of an interaction process between man and nature [78]. Subsequent studies have shown that disasters occur at individual, group, and societal levels, and individuals collectively adjust and shape their environment at local, national and global scales (e.g., [79, 52, 80]) based on their perceived risk and in response to natural extremes. Barkun's [81] study of the systemic issues of temporal and spatial scales in disasters revealed that modern disasters were not constrained by spatial and temporal boundaries, but rather occurred across scales. For example, Typhoon Haiyun in 2013 triggered a top-down, organized international response to a regional disaster in the Philippines [82]. Local, bottom-up responses had significant roles in the case of the 2010 and 2012 earthquakes in Canterbury, New Zealand, even though they were found insufficient without external aid [83].

By the 1980s, conceptions of climatic vulnerability and resilience entered into the vernacular, and the understanding of disasters shifted from single events or type of hazard to ongoing processes and relationships (e.g., [84]). Natural disasters were coming to be considered the outcome of extreme geophysical processes and the failures of human systems to appropriately manage ongoing relationships with their habitats. A longitudinal review of large-scale disasters by O'Keefe et al. [85] revealed that geological changes could not explain the increasing costs and loss of life unless the population's vulnerabilities and socio-economic factors were also assessed. Timmerman [47] introduced the concept of risk and risk assessment, and he addressed how hazards stress the socio-ecological system, inducing adjustments and adaptations in the social system depending on periodicity. Social systems under continuous and periodic stress make permanent and temporary adjustments to continue functioning, such as in the case of annual flooding, while those systems that experience periodic stress only adjust to disasters of greater magnitude, such as in the case of building improvements to resist earthquakes. In a collection of studies, Hewitt [84] showed that disasters were dependent on how social systems assess and adapt, avoid, or reduce the risk from hazards.

The significance of human contribution to losses of life and property in disasters was widely recognized by the 1990s, and studies on vulnerability accounted for social, cultural, economic, and political processes as well as the ongoing geophysical and biological processes that trigger natural disaster events. One such example is the Pressure and Release Model and the Access Model proposed by Blaikie et al. [86]. In these models, the focus was on the social vulnerabilities within a community rather than the hazard itself. These two models integrated top-down natural and social forces with bottom-up individual decision-making of a population of actors, and they incorporated macro- and micro- mechanisms into the larger socio-ecological system. By studying the interconnections of natural and social systems, researchers began to analyze the interrelated and interdependent elements in the ecological research fields with respect to human sociocultural systems and ecological networks. The integration of ecological and social systems revealed interdependent relationships and adaptive strategies that evolved through selective forces to reduce vulnerability in disasters [87]. Cannon [88], using a similar systems approach, showed how the risks and opportunities in the environmental system are unevenly distributed throughout the population based on social power structures and can be analyzed through vulnerability maps. Mileti [27] used the human-coupled systems (or socio-ecological systems) approach to understand the complex interactions between the environment and human perceptions, actions, and organizations to introduce the concept of sustainable hazard mitigation with the objective of using sustainable community planning to reduce disaster losses.

New models like Blaikie's et al.[86] and Mileti's [27] conceptualizations of "sustainability" reflected a shift in disaster issues from a paradigm of hazards and emergency to that of risk reduction and mitigation. Disaster events were no longer viewed as unusual and infrequent, but part of larger socio-ecological processes deeply rooted in local communities, and these processes were embedded in complex adaptive systems (e.g., [89, 90, 91, 92]). Such work also emphasized that when studying local- and regional-scale interactions one should account for local memory and learning processes to address sustainability, risk reduction, and the adaptive capacity of these socio-ecological systems [93]. The work was subsequently extended to account for dynamically linked systems with structures, processes, feedback, nonlinearities, uncertainty, resilience, and entropy that signal a complex adaptive system's ability to self-organize and build capacity and to learn and adapt to recurrent disturbances and change (e.g., [94, 95, 96, 97, 98, 99, 100]). In the context of disaster research, these effects have been explored as a community's adaptive capacity (e.g., [101, 102, 103]) and adaptive resilience (e.g., [104, 105, 106, 107, 100]). As a result of these works, the application of system of systems and properties of complexity has increasingly begun to appear in the disaster literature (e.g., [108, 109, 48, 96, 50, 110, 111, 112, 113, 100]). Work addressing complexity has spawned new areas for interdisciplinary theory such as relating critical infrastructure to society's feedback loops in cascading disasters [114]. However, it has not been fully applied the lens of complex adaptive systems.

In summarizing this section, the general body of disaster research on areas within the physical system has grown from treating each disaster as an isolated, unique, extreme event to the study of disaster events as part of larger socio-ecological and human-coupled system processes. As more data was collected over time, researchers were able to define elements of the social and the ecological systems and incorporate them into the context of a complex adaptive system (e.g., [90, 115]), and this categorization has provided the language and metrics to find evidence of causation in the midst of multi-scale dynamics, non-linearity, and uncertainty. The interactions, inter-dependencies, processes, feedback, and learning of the physical system's elements have led to the conceptualizations of adaptive capacity and adaptive resilience. With this physical subsystem of disasters discussed, we now turn to the social side of disasters.

2.1.2 Social Systems

At the core of disaster research is the study of group behavior (which we integrate into out model in Section 3.3 and it results in Section 4.4). Here we focus on the social system and the behavior of groups, families, organizations, and communities underlying disaster response. Prince [71] was the first to document the disintegration of the social system, particularly governance and behavioral norms, in response to a disaster. His work established a central theory of disaster: with the crisis comes social chaos and then the transition of organizations into new forms of collective behavior, social relationships, and compositions [71, p. 67] (see also: [116, 117]). Studies following Prince's [71] work differentiated behavior based on the disaster's causation, whether "man-made" events, such as war and technological accidents, or, on the other hand, "natural," such as disease epidemic, earthquake, or flood. The natural events were viewed as infrequent occurrences that could be mitigated with better preparation and response [73, 74, 72]. Similar to work on the physical system cited in the last section (2.1.1), this early research on the social system was generally descriptive and concentrated on documenting basic observed behaviors (e.g., victim trauma, convergence of aid responders to the area [73, 71, 72]). The importance of this work during this period was to dispel disaster myths by establishing that panic was an infrequent behavior and required specific conditions and that emergency warnings could significantly affect behavior [73, 118].

By the 1960s, a growing body of work led to the development of a number of disaster theories. One such theory was Fritz's theory of therapeutic adjustments [22] in which the situational characteristics of the disaster along with community adjustments lead to a shared experience that provides physical and emotional support. Fritz found that human behavior differentiated in relation to the disaster's spatial zones, time periods, type of involvement (e.g., victim, national guard, medical professional), and prior preparation and conditioning. Leeds [119] proposed that the cultural norm of unilateral giving replaces that of reciprocity in response to the social vacuums that arise from non-routine situations. Anderson's [120] study of a 1964 Ohio River Valley flood found that repeated community adaptations created a sub-culture of learned organizational responses in norms, values, knowledge, and technology to cope with the physical system. In this specific case, community leaders in the Cincinnati area developed a set of emergency standby mechanisms and complex inter-organizational disaster plans to combat floods. Drabek and Boggs' work [121] identified that family ties had a significant effect on responses to warnings and the decision to evacuate, and Turner [122] proposed that mechanical and organic solidarity are enacted and used by community residents to provide emotional support and overcome disaster trauma. By the end of the decade, Barton [21] collated a comprehensive volume summarizing disaster theories of individual and collective behavior in response to extreme stress. Most notably, Barton proposed a detailed model of the therapeutic community response that included the activation of a communication system, the willingness of victims to communicate the extent of deprivation, sympathetic identification with the victims, relative deprivation, blaming of the victims, a normative mechanism, and situational and motivational determinants of helping [21]. With a relatively robust and growing body of studies, researchers were able to move away from simplistic explanation and models for behavior in disasters. New descriptive and explanatory theories were able to differentiate behavior found within the types and stages of disasters, and disaster research on collective behavior in the following decades built on these findings of behavioral differentiation.

The growing body of empirical data on human behavior in disasters ultimately led to a challenge of the predominate conception of therapeutic community culture in disasters. Although it has been challenged in later research (e.g., [123, 124]), Erikson's study of the 1972 Buffalo Creek flood [125] suggested that the disruption of social networks and neighborhoods could result in a collective trauma of fear, apathy, and demoralization. Further, longitudinal studies of disasters by Quarantelli and Dynes [126] showed that community cooperation frequently occurred in the early emergency stages of the disaster, but that conflict arose in the later stages due to variations in socio-cultural conditions. Oliver-Smith [127] explained this with in-group/out-group dynamics and varying patterns of social identification and interaction in the face of evolving problems during the long processes of recovery and reconstruction. The characteristics of long, slow disaster processes, and long-term disruption and stress, such as those in technical disasters with chronic community stress, prevented the emergence of a therapeutic community [128, 129]. In such instances, it was found that communities experienced a corrosive community process, characterized by blame assignment and evasive, unresponsive authorities [130]. Rather than creating therapeutic processes through cohesion and support, emergent groups become non-responsive, competitive, and hostile.

To uncover the therapeutic and corrosive processes of social relationships, studies of informal relationships and social network effects entered broadly into disaster research around the 1980s. Up to this time, research on the social system had focused primarily on organizations and to a lesser extent on families. Drabek et al. [131] completed an in-depth study of kinship and friendship relationships that revealed exchanges in these relationships supported disaster recovery. These networks of relationships were shown to operate as parallel structures to formalized organizations. Later it was shown that families made decisions, determined disaster activities, and mediated the flow of information as a unit [132], and social networks were crucial to the early formation of emergent citizen groups [133]. Bolin [134] developed a preliminary model of family recovery based on levels of embeddedness in kin and institutional networks. A comparative case study of disasters by Bolin and Bolton [135] along the dimensions of disaster agents, ethnic groups, patterns of destruction, aid utilization, and victim recovery revealed complexities and variations in the process of disaster recovery, but also found that at all the disaster sites kin relationships provided morale and emotional support. In another social model, Bates [136] conceptualized modern society as a complex network of social systems which were later shown as linked through social mechanisms [137]. These enduring social relationships were again found to be the determinants of collective behavior in an application of emergent norm theory to evacuation behavior [138]. All of the studies discussed in this section so far indicate the significant role that social relationships have played as both potential vulnerabilities and opportunities for support and adaptation in disasters.

Building on theories of therapeutic and chronic processes and effects of social relationships on human behavior in disasters, the 1990s can be characterized as a period of discovery and differentiation in which new details of human behavior were uncovered rather than a period of major advances in theoretical understanding. Researchers found that disaster phases were not necessarily sequential and did not uniformly affect an area [139]. Other work found that disaster vulnerability and response was situational, and analysis of vulnerability variables, gender, age, ethnicity, and disability revealed that individual hazard perception and choice of behavior were constrained by existing relationships and power in social structures (e.g., [140, 141]). Race, education, and age [142], gender [143, 144, 145], age and income [146, 147], and ethnicity [148] have also been shown to deferentially affect the experience of and recovery from disasters.

Disaster research in the 21st century has brought a surge of social theories and models that integrate underlying social, economic, and political processes, the interconnectedness of individuals and communities in resilience, and the complexity of these dynamics in disasters. Perhaps the most significant advance is the re-conceptualization of citizens as resources rather than simply victims [149, 150]. Numerous studies have argued that community resilience and local capacities are neglected in disaster planning and response (e.g., [151, 152, 153]). Dynes [154, 155] incorporated social capital into the conceptualizations of communities in disasters. Nakagawa and Shaw [156] and Shaw and Goda [157] subsequently found that higher levels of social capital and collective action were associated with faster disaster recovery from the Kobe earthquake of 1995. Micro social networks have also been demonstrated to be important for disaster recovery and the evolution of institutions to solve post-disaster collective action problems [158], and varying forms of social capital in bonding, bridging, and linking social ties could alter the effects of disaster resilience and recovery mechanisms [159].

The local capacity of communities to prepare for, respond to, and recover from disasters is now embedded in new models of community resilience including Tobin and Whiteford's structural-cognitive model [160], Rose's economic model of inherent and adaptive resilience [107, 99], Maguire and Hagan's social resilience model [161], and Cutter et al.'s widely adopted Disaster Resilience of Place (DROP) model [101]. In such work resilience can be broadly understood as the ability to withstand stressors and return to normal activities. Norris et al. [106] found that community resilience emerges from four primary sets of adaptive capacities: economic development, social capital, information and communication, and community competence. Their work noted this requires intangible community capabilities such as flexibility, decision-making skills, and trust, while others have shown that resilient communities are those that effectively activate formal and latent social connections for self-organization and local leadership (e.g., [162]). To account for the complexities in disasters, Pelling [93] developed a participatory framework of vulnerability and risk assessment that enables disaster risk reduction and management to cross scales from the global to the local. This framework allows for adaptive learning using local knowledge that empowers the local community in times of disasters. We see this in the theory of adaptive governance [163], in which key persons self-organize into social networks to develop common understandings and policies for ecosystem-based management, that has been applied to disaster resilience and risk reduction [164, 165]. In an analysis of disaster risk reduction, Wisner et al. [166] confirmed the importance of these properties in disasters: multi-scale, top down and bottom-up dynamics, outside specialist knowledge from many disciplines, and local knowledge. These theories, frameworks, and models integrate new conceptualizations of group behavior that include social capital and resilience, networks and learning, open and adaptive systems, adaptive capacity, and complexity.

Research on the social system in disasters has revealed complex processes of social interactions, at times cohesive and at other times divisive. Early work documented the basic behaviors in different types and stages of disasters (e.g., [22, 72]), and these were later attributed to social norms of reciprocity, culture, family, and solidarity in crisis. Building on a body of empirical studies, models of a therapeutic community response [21] and a corrosive community process [128] showed how collective behavior could become either cohesive or divisive. These processes could be present in the same disaster depending on the stage and duration of the disaster, socio-cultural conditions, and the quality of social relationships. Ultimately researchers developed complex theories that incorporated underlying social, economic, and political processes such as community resilience, adaptive governance, and social capital (e.g., [160, 163, 101]). All of these current theories incorporate individuals in webs of relationships acting within complex, adaptive social and physical processes.

2.1.3 Individual Systems

The actors underlying the social system are individuals whose cognitive systems determine behavior and interactions with others and their environment, the physical and social systems (this notion will be brought out in Section 3.3 when we introduce our agents). Early work on disaster theories of individual cognition and psychology arose from studies of population reactions to stresses in war and later to extreme weather events. For example, Wallace [72] called the dominant individual reaction a disaster syndrome, in which those affected by the event were described words such as "shock," "dazed," "stupor," "apathy," "stunned," and "numbed" as a result of cognitive dysfunction that arose from disruption of their culture and routine behaviors. Killian [167] theorized that this led to conflict as individuals struggled to sustain the behaviors required for membership within social groups after a disaster. Empirical studies provided evidence of a mix of individual reactions in disasters; specifically, Tyhurst [168] found that approximately 75% of individuals displayed symptoms of a stunned and bewildered lack of awareness or restricted field of attention, while 10-25% were confused, paralyzed, hysterical, or screaming, and 12-25% were cool and collected. Panic was also found to be an unlikely response to disasters [73], but rather manifested only under specific conditions [118]. To explain some of the variation, Glass [169] proposed individual psychological states at each stage of a disaster: pre-impact (denial, adopts fatalistic concept, apathy, and training), warning (over-activity and flight), recoil (under-activity, apathy, disaster syndrome, or fatigue), and post-impact (grief, understanding of personal loss, anger, or resentment). In the post-impact stage, scapegoating was found to rise from a complex mix of frustration, fear, guilt, and latent hostility [170].

Moving beyond the stages of disaster, Fritz [22] theorized that when disasters strike and social patterns and cultural norms are disrupted, individuals are forced to make critical choices within very short time-spans. Issues of survival, subsistence, shelter, and health take precedent over social order and meaning, and individual reactions to the perceived context differentiates their behavior in relation to location, time, involvement in the disaster, and preparation and conditioning. Crawshaw [171] provided empirical validation of these differentiated individual reactions and attributed them to the needs of individuals in specific age groups and family make-up. During this time period, Lazarus [172] proposed a psychological stress theory in which individuals engage in threat appraisal rather than anxiety arousal before engaging in coping mechanisms of actions to strengthen resources from harm, attack, avoidance, or defense. The theory was later modified to include cognitive appraisal with assessments of the person-environment transactions and problem- and emotion-focused forms of coping under stress [173, 174]. Adding to the complexity of individual behavior and decision-making, Drabek and Boggs [121] found that individual behaviors and choices were heavily influenced by the warnings and evacuations of relatives and by their familial roles as parent, child, elder and younger family members.

In his summary review of disaster studies, Barton [21] attributed individual behavior to personal emotions and preferences and to role behavior in informal and formal organizations, and deeper psychological and social norm explanations followed. Perry and Lindell [175] developed a conceptual model of inter-related factors that have individual psychological consequences along three dimensions: the characteristics of the disaster, the characteristics of the social system, and the pre-impact characteristics of the individual. These factors included community preparedness, forewarning, scope and duration of impact, destruction of kin and friendship networks, extent of property damage, pre-impact psychological stability, grief reactions, disaster subculture, and existence of a therapeutic community and institutional rehabilitation. Mawson [176] proposed a theoretical model of social attachment to explain self-preservation and the lack of behavior in disasters and crises. In response to a threat or disaster, the typical individual seeks the proximity of familiar persons and places, and thus individuals do not flee from a disaster, but rather flee to social attachments. In the post-impact stage of disasters, a significant part of individual trauma was found to be the loss of these family members, community, and other social attachments [125, 73, 168].

Individual stress and behavior in disasters was not only attributed to psychological explanations of loss or change of social attachments, but also to the disruption of social norms and corresponding rules of behavior. Emergent norm theory [177] attempted to explain that when individuals encounter new situations, new norms can "emerge" spontaneously from ongoing social processes and events without reflection of existing social structure. However, societal norms continue to constrain individual behavior in times of stress [178], and individuals maintain these norms and extend their social roles to address the needs of a crisis [178]. Individual self-categorization [177] and social identity have been found to be significant in disaster behavior, explaining emergent groups and affecting group solidarity [179] and provisions of aid [180]. Similar to the pattern of research studies in the social system, individual behavior research in the 1990s provided few new theories, but it did produce findings of both heterogeneous and homogeneous behavior at different times of the disaster. A meta-analysis of studies on psychopathology, psychological problems, and pathologies or impairments suffered by post-disaster victims [181] provided evidence of significant heterogeneity in post-disaster responses. Individual responses varied by victim and disaster characteristics depending on the death rate/loss of social attachments, time elapsed from impact, and degree of human responsibility. A contrasting study by Goltz et al. [182] found that the rapid onset of disasters elicited more homogeneous responses with individual behavior motivated by fear and influenced by the presence of others. At the onset individuals engaged in rational self-protective activities to prevent injury during an earthquake. These were survival-oriented, learned and adaptive responses from past experiences. Adaptive behavior also affected the social attachments of individual victims as they pruned their social networks in times of disaster to optimize energy and resources [53].

The significance of social attachments as an explanation of individual behavior was further reinforced in the 2000s. Hobfoll's [183, 184] process-based theory, Conservation of Resources (COR), predicts that resource loss is the principal ingredient in the stress process and that self and individual stress is derived from primary social attachments within families and intimate social groups. These attachments, as reflected by social embeddedness, the size, level of activity, and closeness of a social network, were also found to protect individuals from psychological distress [185]. Mawson [186] revised his theory of social attachments and proposed that individuals balance the need to be close to affiliative attachments and to be far from physical threats. This social attachment theory was elaborated by Mawson [187] using a biopsychosocial approach based on stimulation-seeking and stimulation-avoidance behavior. "Panic," including flight, aggression, and other forms of intense agitation, is a result of intense stimulation-seeking behavior, activities that facilitate contact between an organism's sensory receptors and external objects, arising from a high level of arousal. Individual resilience and forms of capital, including social, are the latent measures of capacities and resources in the Resilience Activation Framework [56] which can be used to test how access to social resources promotes adaptations and coping mechanisms in crisis and disasters.

Disaster research in the 2000s also introduced new cognitive science approaches for explanation of individual behavior that focus on decision-making within the broader context of survival, loss, and social norms. In a socio-cultural model, Paton [188] used multiple dimensions of risk assessment and preparation based on motivation and intention variables to provide explanation for disaster preparation behavior, and Rosenstein [189] proposed an assessment for Decision-Making Capacity (DMC). Van Fenema [190] proposed the concept of collaborative elasticity, a collective capability to manage the unexpected in crisis, that leverages theories of individual cognition, distributed cognition, and the collective mind, and Ripley [191] emphasized cognitive responses to disasters and decision-making in a survival arc of denial, deliberation, and decision before action. Leveraging work on decision-making and game theory, Eiser et al. [192] proposed a conceptual framework in which individuals make decisions based on perceived risks in conditions of uncertainty, and Espina and Teng-Calleja [193] have recently applied social cognitive theory [194] to show how individual and environmental factors influence disaster preparedness. Individual interpretations of risk and actions in uncertainty are shaped by experience, personal feelings and values, beliefs, and interpersonal and social dynamics. In social cognitive theory personal agency is regulated between direct personal agency, proxy agency (relying on others to act in one's interests), and collective agency (social coordination and interdependency). Extending social cognitive theory, Benight and Bandura [195] found that human agency and perceived coping self-efficacy affected an individual's recovery from trauma.

Compared with the research on the social systems, work on theories of individual cognition and behavior in disasters outside of emergency management and organizational theory studies is not well-integrated into the disaster literature. However, as occurred in the physical and social systems there has been a pattern of initial observation and descriptive theory, discovery of underlying explanation, and research that gradually leads to more complex theories. These theories have been posited from the perspective of roles and social norms (e.g., [177]) and psychology (e.g., [172]). Later theories integrated both of these areas in multi-dimensional analysis of inter-related factors (e.g. [188, 175]) or delved into biopsychosocial approaches (e.g., [187, 181]). The recent cognitive approaches to individual behavior bring all of these approaches into a decision-making framework that accounts for the effects of roles and social norms, social attachments, and biopsychosocial processes in disasters (e.g., [192, 191]). It also provides explanation for individual behavior as an output of the human cognitive system, a complex adaptive system that incorporates, among other features, decision-making, perception, reasoning, memory, emotion, and biology. Now having reviewed the complex adaptive nature of the physical, social, and individual systems of disasters, we can directly consider disaster as a phenomenon within a complex adaptive system framework for the remainder of this report.

2.2 The Role of Complex Adaptive Systems in Disaster Research

In the previous sections we reviewed disaster theories within three systems: the physical (Section 2.1.1), social (Section 2.1.2), and individual (Section 2.1.3). We specifically highlighted specific properties of complex adaptive systems: heterogeneity, webs of connections, relationships and interactions, and adaptations arising from individual actions, decisions, and learning. Along with these properties, a complex adaptive system contains dynamics that include feedback loops, patterns of self-organization, flows of information and resources, and system shifts between stability and instability (in and out of equilibrium). By definition, a disaster is a disruption of the social system after which components and actors of the system must adapt and readjust in order to return to some form of equilibrium. Our review found relatively few explanations of the interactions, processes, and feedback that cut across the three systems, other than that of Gunderson and Holling [90]. We would argue that this gap can be addressed with explicit study of the interactions between subsystems through the lens of complex adaptive systems. Complexity science identifies a complex adaptive system as a system within which the interactions between individual elements and actors lead to emergent behavioral patterns and adaptation [38, 39, 40]. The properties and dynamics of complex adaptive systems are found in the physical system, the social system, and the individual system. To test whether disaster and complexity theory are integral to understanding human behavior in disasters, a systematic literature review of disaster theories was organized and discussed in Section 2.1. Our organization was then used to explore theories related to the physical, social, and individual systems. The following section provides a brief argument for how each of the systems can be linked (Section 2.2.1) building upon what was discussed in Section 2.1. This leads us to how the integration of these systems (i.e., physical, social and individual) exposes their interactions, and we introduce our framework of the intersecting complex adaptive systems of disaster (Section 2.2.2). Lastly, we discuss how concepts of complex adaptive systems and complexity science can be applied in disaster research (Section 2.2.3).

2.2.1 Linkages between the Physical, Social, and Individual Systems

In the context of the physical systems, disasters are caused by a combination of physical and social processes. Early theories identified periods of stability, system disruption, and a return to stability as discussed earlier (Section 2.1.1). By the mid-1900s, the physical and social processes of disasters were recognized as being shaped by both individuals and society (e.g., [76]). Disasters were later recognized as events occurring across multiple scales (e.g., [81]). Interdependencies between the ecological (i.e., physical) and human (i.e., social) systems and adaptive strategies led to evolutionary change [87] such as seen in agroforestry processes in the Amazon (e.g. [196, 197, 198]). The adaptation and adjustments of a social system created varying hazards and risk profiles, and models of disaster illustrated how top-down and bottom-up processes affected disaster outcomes. The understanding of dynamically linked systems has led to current socio-ecological models in which disaster events occur within a complex adaptive system that learns and adapts in response to ongoing interactions. In this perspective the effects of individual and social system behavior in groups and organizations are critical factors for explaining behavioral response and resilience to disasters (e.g., [109, 34]).

Collective behavior occurs as part of a social system (Section 2.1.2), whether it consists of informal groups, families, organizations, or communities, and these collective behaviors are the aggregate behavior of individuals who respond to disaster events. Behavioral responses are the result of both the physical effects of the disaster and the interactions in formal and informal social relationships. Modern disaster theories integrate these understandings of complex social, economic, and political processes aggregated from the interactions between community members in therapeutic and corrosive processes. They are now used to measure the capacity of communities to survive and recover from disaster events (e.g., [101, 106]), and these theories account for processes of adaptation, learning, and decision-making that are core properties of a complex adaptive system (e.g., [166]).

As separate components of the social system, individuals are the drivers of bottom-up processes, and the collective action arising from these individual systems are driven by their cognition (Section 2.1.3). Empirical data has shown how varied individual responses are, and that these responses were largely rational and adaptive, albeit heavily influenced by social connections, identity, experience, norms, and roles. Just as the individual system influences the physical and social systems, past experience and the current context of the physical and social systems shape decision-making in the individual system (e.g., [194, 184]). The social system as embodied in familiar connections has been found to be a significant factor in what decisions are made and how well individuals survive disasters (e.g., [121, 186, 187]). Recent research leverages cognitive science approaches that utilize decision-making theory to explain individual behavior as shaped by complex interacting variables including the environmental context, emotions, experience, social norms, and identity (e.g., [192, 188]).

2.2.2 Integrating the Physical, Social, and Individual Systems

Within physical, social, and individual systems, the properties and dynamics of a complex adaptive system are evident in the sense that interactions between individual elements lead to emergent behavioral patterns and adaptation, and heterogeneity is present in disaster impacts, collective behavior, and individual experience. The webs of connections, relationships, and interactions within the systems lead to adaptations as system elements learn and respond to new experiences. More significant, the system effects in the physical system affect both individuals and society; societal dynamics impact both the physical and individual systems; and individual actors affect their physical and social systems. The dynamics in these three interacting systems as hypothesized could create their own sets of adaptations and emergent behaviors; however, disaster research studies tend to be focused on one particular subset of human behavior in the systems rather than on how the interactions between systems create feedback and aggregate effects.

To account for these greater systems interactions, we propose a model with three intersecting systems within a complex adaptive system: the physical, social, and the individual systems, as shown in Figure 1. Interactions between the physical, social, and individual complex adaptive systems aggregate to create larger effects from their properties and dynamics. Heterogeneity can be found in the variations of disaster impacts on populations and geography. Flows exist with the migration of populations, individuals sharing information, the physical force of the disaster, and subsystems interacting. For example, when significant rainfall leads to river flooding, charitable organizations cooperate with federal agencies, and individuals self-organize and apply occupational skills to save neighbors. In a disaster, all three systems are thrown out of equilibrium and go through periods of adjustment to return to some form of stability. The return to equilibrium internal to each system and externally between systems is accomplished in processes of emergence and adaptation.

At the center of a complex adaptive system are the heterogenous actors interacting in processes that create feedback, shifting the system in and out of equilibrium at some tipping point or critical threshold when a smaller change triggers a set of unstoppable processes such as bank runs [199]. In disasters, the tipping point is when conditions have built up to a point at which society is seriously harmed and can no longer operate its essential functions. The time of impact is an example of the tipping point for a tornado, whereas the point at which a river crests over its levee would be the tipping point for a slow-onset flooding disaster. These tipping points occur when feedback mechanisms in the system are out of balance. In the case of a natural disaster such as a wildfire, they can occur when positive feedback (adding energy into the system and amplifying change) in the form of dry, hot air is not balanced with weather systems bringing negative feedback (removing energy from the system and decreasing change) in the form of rain moving into the area. A nuclear power plant accident also provides an example of feedback in a man-made disaster. For example, cutbacks in funding for well-trained, qualified technicians could create positive feedback that leads to a failure to identify minor operating problems and implement the appropriate safety protocols. Numerous feedback mechanisms can be found operating in any particular disaster, adding to the complexity of the system and creating both added risk and opportunities for mitigation.

A complex adaptive system's internal interactions that lead to emergence, adaptation, and, at times, disasters are evident in many real-world examples. For instance, volcanoes provide an example of how physical forces build up pressure inside the earth until they reach a critical threshold and are released in an eruption. From the physical system, discussed in Section 2.1.1, we can observe how a greater frequency of volcanic eruptions creates a negative feedback signal that signals settlements to move farther from the volcano and away from hazards. The fertility of volcanic soil creates positive feedback motivating people to settle closer to the volcano, thus increasing the hazard. The negative and positive feedback lead to settlement patterns such as found by Small and Naumann [200]. In this scenario, past individual experiences of volcanic eruptions and livelihoods, such as taking care of livestock, also influence how people understand and respond to these events, creating variations in the perceptions of the hazard and evacuation rates [201, 202]. Cultural and social factors create social forces and norms that affect how communities organize and communicate to prepare, mitigate, and respond to volcanic eruptions [203], and complexity theory suggests these forces can be identified and measured to find patterns in system behavior.

The case of volcanic eruptions does not illustrate how social (Section 2.2.4) and individual (Section 2.1.3) systems affect the physical system (Section 2.1.1), but rather how the systems adjust their patterns of behavior to physical forces. In another case of a natural hazard, flooding is affected by both the physical and the social systems in a set of feedback mechanisms. The physical flow of excess water from rain runoff and snowmelt in seasonal weather patterns creates forces that carve natural drainage basins collecting water and funneling it into shared outlets. When organizations, governments, or industries set aside land areas and build infrastructure such as dams, levees, or housing, the social

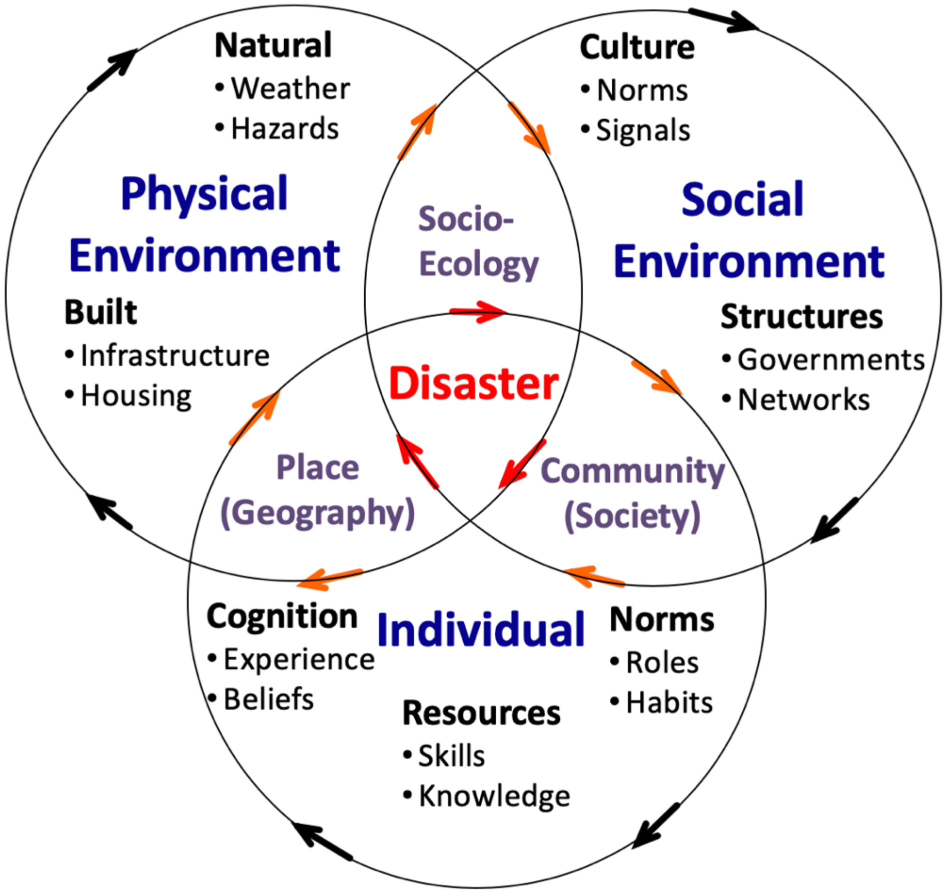


Figure 1: Intersecting Complex Adaptive Systems of Disaster

system reshapes the flow of water and creates its own structural forces that affect the physical system, developing new spaces for flooding in catchment areas. The relationship between physical and social actions that reshape topographies in the environment and create new waterflow patterns can be understood as feedback between forces in the physical and social systems.

In the social system (Section 2.1.2), forces that affect flooding include structures built by governments and landowners, building codes, and flood management policies in the form of regulation, insurance, and land management policies implemented to reduce the risk and costs of flooding (e.g., [204, 205]). The adjustments between the physical (Section 2.1.1) and social systems (Section 2.1.2) create cycles of flooding and implementation of new flood protections [206]. These interactions reduce the frequency of flooding events, but also increase the risk of catastrophic floods if flood protection measures fail [207]. As theorized in the Panarchy model [90] and complex adaptive system theory in complexity science [30], flows in the two complex adaptive systems, the physical and social, are continuously interacting and readjusting. The feedback between flows ultimately leads to a tipping point or critical threshold, when the systems are thrown out of equilibrium due to failures in the drainage system and catastrophic flooding.

The individual system, discussed in Section 2.1.3, also creates forces that lead to feedback between complex adaptive systems. They form from bottom-up processes such as when public opinion builds to a point of revolution or when an individual in technological systems triggers extreme damage through human error or implementation of weapons of mass destruction such as planes or bombs. The Exxon Valdez oil spill provides one such example of the impact of individual actors interacting with physical and social systems. At the disaster's tipping point, the Exxon Valdez tankship grounded on Bligh Reef in Prince William Sound due to multiple factors [208]. The ship's master (captain) and the third mate played individual roles with errors in judgment related to alcohol, fatigue, and work overload. Expanding out from the individual to the social system, the Exxon Shipping Company was found to have inadequate manning procedures, insufficient chemical dependency monitoring programs, and to have manipulated shipboard reporting of crew overtime. The Vessel Traffic Service was found to be without adequate equipment, manning levels personnel training, and management oversight. Preventative and mitigation measures in the form of regulations, equipment, and response plans to guard against the environmental risk and respond to oil spills were determined to be inadequate relative to the scale of the oil spill [209, 208, 210]. Among the findings, oil spills were an acceptable, inevitable risk given the vital importance of the oil industry, but environmental safeguards were needed to reduce the risks inherent with oil production and transportation [210]. In the aftermath, post-disaster clean-up and litigation efforts were just as harmful as the oil spill itself [211, 212], continuing environmental and societal damage long after the disaster [213].

Economic forces, social structures, and technological complexity as represented by the Exxon Shipping Company, the oil industry, and the government provided positive feedback in the complex adaptive systems. Negative feedback was limited in this technological disaster with few processes that could lead to reductions in risk and mitigation, and those that existed were poorly linked to those responsible for the positive feedback [214]. The continuation of unabated positive feedback leads to punctuated entropy which is a permanent decline in the adaptive capacity of the system due to cumulative disaster impacts [215]. Further complicating system action and adaptation, perceptions of the incident were manipulated in media coverage [216, 217]. In these cases of high-impact, low- probability events, the complexity of complex adaptive system dynamics as well as society's acceptance of often poorly understood risks makes such technological disasters inevitable.

The analysis of disasters from the perspective of complex adaptive systems provides insight into the forces that shape and lead to disaster as discussed here and in Section 2.2.1. By identifying the connections between the physical, social, and individual systems and examining their relationships, researchers can measure the flows that push complex adaptive systems in and out of equilibrium. In the case of volcanic eruptions, equilibrium is established from long-term flows that lead to a balance of population size relative to the distance from the volcano. Changes in flooding hazards show how the social system continuously adapts with the physical system, reshaping flood plains. Analysis of feedback within the social system involved in the Exxon Valdez oil spill shows that the positive feedback from economic and social forces were not balanced with corresponding negative feedback. The oil spill was an inevitable outcome of a social system out of equilibrium, creating positive feedback that tipped over into the physical system and throwing it also out of balance. In the context of disaster, the lens of complex adaptive systems provides new avenues for exploring causation and identifying forces and solutions that can lead to negative feedback and maintain equilibrium between systems.

2.2.3 Applications and Implications of Complex Adaptive Systems

Beyond the fundamental analysis of feedback and system equilibrium, the application of complex adaptive systems and complexity science to disasters has a number of implications for disaster research. Theories of self-organization, emergence, and interacting processes are central to both complexity and disaster, and these properties, as understood in complex adaptive systems, are already being applied in some areas, such as with the model of the adaptive cycle in Panarchy [90]. More important, the properties highlighted in this review—heterogeneity, webs of connections, relationships and interactions, and adaptations arising from individual actions, decisions, and learning—give rise to non-linear dynamics and high levels of uncertainty. The nonlinear dynamics indicate the potential presence of power laws and, thus, proportional relative changes in the system that vary as a power of some attribute. Power laws are absent of any "average," and events in the systems described by power laws occur as "many small ones, a few larger ones, and occasionally extremely large ones" [218]. The negative and positive feedback mechanisms often made visible in data distributed by power laws are continuously driving the system into a critical phase [219], functioning at ever greater efficiencies and toward the edge of chaos as described in Kauffman [43] and Lansing [39].

In complexity theory, the feedback created by multiple interacting subsystems creates observable patterns as the system shifts in and out of equilibrium. One example of these patterns is the self-similarity evident in the social organizations represented by networks. Self-similarity occurs when one part of an object displays the same pattern as its whole, such as the leaves of a fern. Fractals are self-similar geometric objects or patterns, and by applying analysis to identify these patterns, fractal network researchers have found network patterns to be a function of natural optimization processes [220, 221]. Family groups self-organize themselves into nested hierarchies and social systems proportionally sized in relation to available flows of food, material resources, and other cultural information [220]. Another societal pattern can be found in urban growth. As societal cultures change the rates of innovation, their wealth creation, patterns of consumption, and behavior follow scaling relationships [222]. The measurement of these relationships maps behavioral patterns or signatures in one set of cities that could serve as indicators or patterns of properties in others. Signatures are the distinctive characteristic patterns that can signal the presence of particular variables or interactions. Family network structures and urban growth both illustrate how the dynamics of a complex adaptive system create patterns with the potential for short-term prediction of self-organization and optimized scaling.

The complexity of the systems and their nonlinear dynamical nature preclude the possibility of traditional event prediction. Unlike classical Newton approximations that produce a single-point solution, such as point estimates of parameters in linear regression, a complex adaptive system cannot be approximated with linear equations. Instead, researchers must look for a variety and range of bounded solutions. These systems are also sensitive to initial conditions as described in Lorenz's [223] "butterfly effect." As complex adaptive systems we can expect that the pre-existing conditions of the physical and social systems will have significant effect on how well social systems prepare for, respond to, and recover from disasters, and this sensitivity to initial conditions will prevent the guarantee of any event prediction. However, although more research is needed, there is evidence from simple models of complex systems that early-warning signals could be detected in systems' behavioral patterns before some tipping point and a shift in the system occurs [224]. Just as patterns in weather systems can provide short-term predictability, patterns in disasters could be bounded in probabilistic outcomes. Practitioners will not find one-off solutions or policies for disaster preparation, response, and recovery; rather, changes in behavior and policy could mitigate some harmful effects while preventing others. As a result, the effectiveness of any one policy will vary over multiple events and hazards and will require adjustments given existing conditions, and modelers will need to create simulations that represent the probabilities of intervention strategies [225, 226]. Researchers will need to develop simple and complex models that specifically study the mechanisms and feedback relevant in disasters (e.g., [227]). The goal of scientific study in this area will require a shift from requiring definitive prediction to determining probable outcomes.

Understanding the dynamics of a complex adaptive system requires the exploration of the latent capabilities and vulnerabilities in the particular system; i.e., those unobservable variables that are found to be significant in disaster outcomes as a result of bottom-up and feedback processes [228, 229, 83, 230]. These processes can be partially attributed to the self-organization that occurs as individuals and organizations exploit existing assets or weaknesses. Existing techniques to analyze latent variables include those from statistics (e.g., Regression Analysis, Latent Dirichlet Allocation), machine learning (e.g., Latent Semantic Analysis, Factor Analysis, Hidden Markov Models), and the tangential field of network analysis such as the measurement of social network capacity. The application of these and other innovative statistical and machine learning techniques for latent variable analysis tailored for disaster could improve research and practice in disaster management (e.g., [231].

Researchers and practitioners continue to contend with the multitude of interacting variables and adaptations in the physical (Section 2.1.1), social (Section 2.1.2), and individual (Section 2.1.3) systems, and conceptual models of complex adaptive systems such as discussed in this report (Section 2.2.2) are needed to study and test these interactions in their system of systems. Research in the area of complex adaptive systems must explore the multiple interactions of system components explained by multiple theories and visible in nonlinear dynamics that cannot be studied using traditional qualitative and mathematical models [41]. Social network analysis and geographical information systems (GIS) and other computational methods in the expanding field of computational social science provide new forms of data that can more precisely measure the processes in physical, social, and individual complex adaptive systems (e.g., [232, 233, 16]). Techniques in the computational social sciences using power law analysis and agent-based models are necessary for the analysis of systems that cannot be reduced to single elements, actors, or processes. In experimentation with complex adaptive systems, computer simulations provide the necessary repeated measurements of accumulated data that magnify small differences, making them observable. They also enable collaborative processes that develop shared-risk models for community stakeholders and policy-makers (e.g., [234, 235]), transforming data into information and then knowledge. Analysis of complex adaptive systems in disaster requires a different approach and new methodological tools that can manage large, heterogenous datasets; identify and calculate power laws; run the high sample sizes of events in simulations needed to generate a range of expected outcomes; simulate events and implement theories that cannot be tested with available field data (some of which we tackle in Section 3); identify and explore dynamics with multi-level dependencies; and apply machine learning and other computational techniques for latent variable analysis.

The implications discussed above and corresponding suggestions for methodological approaches are not intended as a comprehensive set of approaches and techniques to address the disaster research questions in complex adaptive systems; rather, they suggest potential areas for exploration. Dynamics in complex adaptive systems produce detectable patterns and potential signatures for particular interactions, and although problems in complex adaptive systems do not produce optimal, single-point solutions, possible outcomes can be computed. The lens of complex adaptive systems presents a new paradigm for disasters that leads to new lines of inquiry. What data is necessary to observe and measure the key feedback processes present in disaster? What are the repeatable patterns observed in disasters? What do they signify? How should computer models of a complex adaptive system in disaster be designed to improve understanding of system interactions? Further work is necessary to establish whether complex adaptive systems can provide any level of prediction, as was the case in weather forecast modeling; however, the process of studying disasters from the perspective of theories of complexity can provide insight into the interactions of individual, social, and physical systems behavior.

2.2.4 Summary of Disasters Viewed through the Lens of Complex Adaptive Systems

A century of research on disasters has evolved from the study of single discrete events that were largely addressed with top-down responses for emergency rescue and management to the study of continuous, repetitive events with complex interactions between systems and across scales. Originally these discrete events were not seen as interacting systems or as a system of systems; however, this has changed over the last few decades. To demonstrate this development within this report we have reviewed and organised theories for disaster study as three complex adaptive systems, the physical (Section 2.1.1), social (Section 2.1.2), and individual (Section 2.1.3). Furthermore, we showed that these systems are interconnected (Section 2.2.1) and described how these systems are integrated through webs of connections and characterized by all the traits of complex adaptive systems (i.e., heterogeneity, interacting subsystems, emergence, adaptation, and learning as discussed in our introduction (Section 1)). The lens of complex adaptive systems enabled us to introduce a new conceptual framework of physical, social, and individual systems that interact across scales (Section 2.2.2). This conceptualization lays down a foundation for disaster science that explicitly studies disaster events as parts of ongoing interactions and processes of subsystems rather than addressing them as individual systems. The recognition that disasters arise within complex adaptive systems offers us a deeper understanding of these events and the theories and tools available in complexity science as applied in Section 2.2.3. As evident in this background into the development of our model, the explicit study of the interactions between these three systems is standard in today's literature (Sections 2.2 and 2.2.1); however, further exploration of their feedback are needed to improve understanding of the nonlinear dynamics that dominate disaster phenomena (Section 2.2.3). The major contribution of this synthesis of material is a framework (Section 2.2.2) that can be used as a conceptual device to integrate disaster theories into a large-scale system that balances the interacting dynamics of multiple subsystems which we will utilize in Section 3. The framework does not invalidate older theories; rather, it creates a space for these theories to intersect and interact, providing stronger explanation for human and environment behavior. It also furthers the conception of complex adaptive systems in disasters and underscores its relevance by directly recognizing and addressing the inherent complexity of disasters. With this perspective, researchers can better take advantage of available computational techniques (such as agent-based modeling as will be introduced in Section 3.3) for studying complexity and more fully explore the dynamics that take place at the intersections of the physical, social, and individual systems.

3 Methodology

3.1 Introduction and Data Sources

To study the reaction of the population of a mega-city to the effects of a nuclear WMD (weapons of mass destruction), we first needed to select a mega-city. There are only two cities in the U.S. that meet the definition of a mega-city, population over 10 million. They are New York and Los Angles. New York is characterized by well defined boundaries and is built vertically, i.e., with its skyscrapers, it has a high population density. Los Angles is a basically spread out. We selected New York as the focus of our research. Our methodology was to build an agent-based model of the New York City commuting area at the individual level, model them at work, and we would model a relatively small nuclear weapon to be imaginable as a terrorist attack, not a super power weapon. We then developed code to implement our understanding of how people have reacted to the two nuclear weapons of World War II, how people reacted to the events of September 11, 2001, and other non-notice events. The goal was to characterize the reaction of the population of a mega-city to a nuclear WMD event. We characterized their reaction as counts of what they are doing on a minute-by-minute basis after the detonation.

The commuting area for NYC includes all of Connecticut and parts of New Jersey, New York state and Pennsylvania, as shown in Figure 2. The area covers a 262 x 234 km area, where 23,004,272 people were living as the census recorded for 2010. Our study area is slightly larger than the official metropolitan statistical area to include both high-density urban and suburban area in the study area, which may provide more heterogeneous population reference for our synthetic population. Besides, using this larger area to represent a mega-city and its surrounding area allows us to capture more dynamics, such as, migration and traffic flows in and out the urban center. With the study area decided, we identified sources of information necessary for our project.

To create a simulation of NYC, our ABM includes both empirical data and data synthesized from empirical sources that are integrated into the complex system of systems described in Figure 3. These data are intended to capture the locations and activities that occupy a population for most of their daily life. Data representing their daily movements can used to derive information on the physical, small-world networks that dominate real-life, social interactions [236]). Specifically, we focus on the daily patterns of where people live and work and capture their movement between these locations as they commute. This section discusses the data sources we used to create both the synthetic population and the model that explores the populations reaction to a nuclear WMD event.

The information about roads, population, schools, workplaces, and commutes used to describe out synthetic population and their behavior in our ABM was obtained from a number of U.S. Government websites. Table 1 showed the reference data for specific model representations. The Tiger Shape files were simplified to create a road network for the agents to travel along in routine commuting behavior, and the US Census Longitudinal Employer Household Dynamics Origin Destination Employment Statistics were used to identify in what counties the ABM's agents commuted. The ABM's agent population was derived from the 2010 US Census Tracts, providing both the characteristics of the agents' gender, age, work status, and home locations and the composition of their households. Further details on the use of these data for the population synthesis are discussed in Section 3.2 and an early paper on this work [14]. Because these data are sourced from the government not only are they reliable, but they are also more likely to be used or adopted by policymakers in the case of using decision-making tools.

While there is data to describe the physical environment and the population, unfortunately, in addition to the challenge of a dearth of data that can characterize a disaster caused by a NWMD detonation, the collection, experimentation, and analysis of dynamic social networks that are not geocoded in social media remain a logistical challenge. Our solution was to develop an ABM that create virtual spaces in which agents are modeled interacting with their social and physical systems. The ABM developed in this research was designed to include location and social networks in the agent's decision-making process in a NWMD detonation and is described in Section 4.4. We demonstrate how an agent-based model that integrates social networks with a spatially explicit environment improves the realism of an emergency response simulation. The results of the model are presented in Section 4.5 and show how emergent networks changing over time are gathered in the model as well as providing data that characterize the response of agents to a NWMD detonation in NYC.

During emergencies and disasters people turn to family and friends for material and emotional support, and the structure and composition of these networks have real effect on how well a community responds and recovers from these events [237]. In the response phase of an emergency or disaster, social networks are used for information and physical support as individuals, groups, and families decide to evacuate, shelter, or find and give aid [237, 238] also discussed in Sections 3.3.6 and 3.3.7. Social networks in this phase of an emergency are very dynamic as people scramble to find safety for themselves and loved ones. Often *ad hoc* emergent groups form temporarily with short-term goals to find shelter or to provide aid and rescue [133, 239]. Research in this area is rarely available because the collection of social network data during the response phase is prohibitive and after-the-fact accounts can be unreliable due to trauma [240]. Agent-based

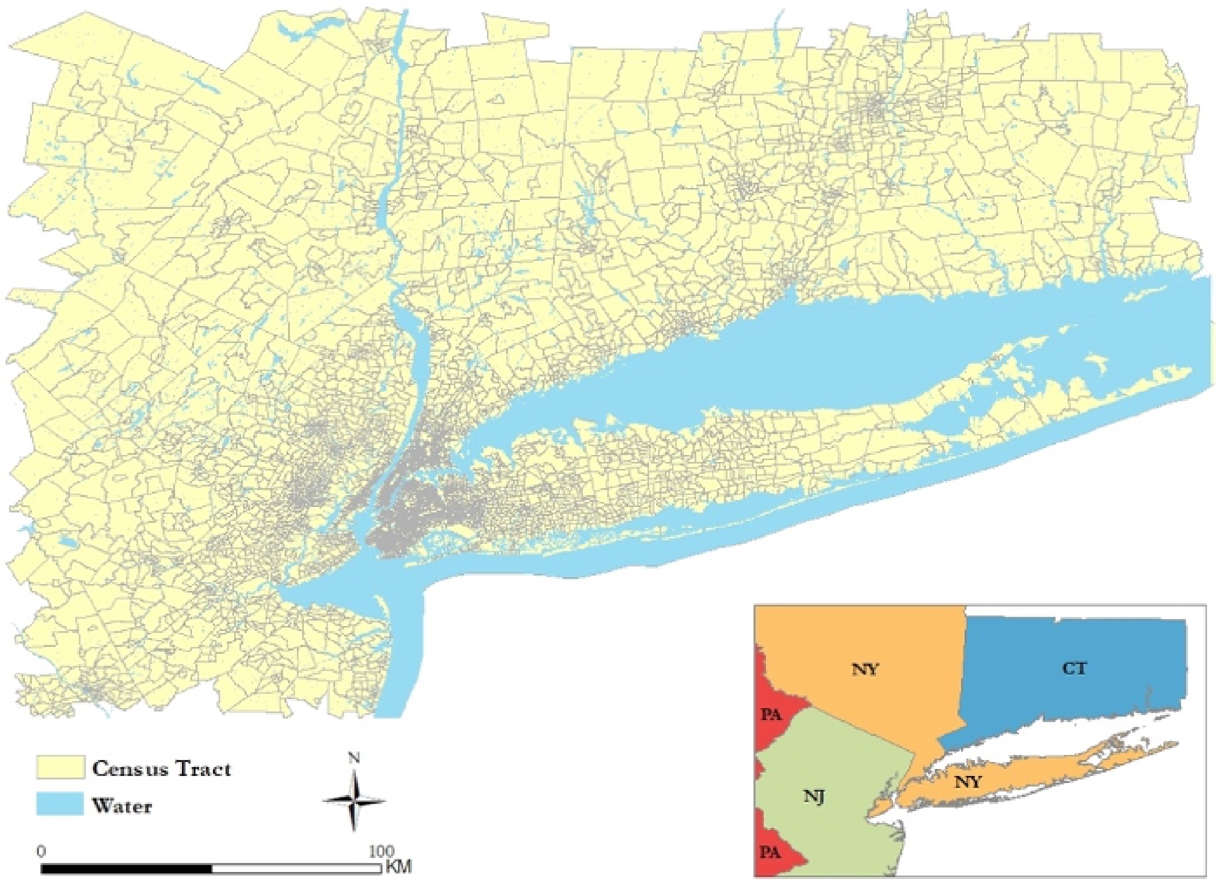


Figure 2: Study Area Showing Census Tracts.

modeling can provide some insight into the dynamics of social networks in the, response phase of a disaster when paired with empirical demographic and geographic data. In Section 2, we discussed how more than a century of disaster research has evolved into understandings of CASs framed as three CAS, the physical, social, and individual, interact in spaces that produce disasters. The social interaction of individuals in groups in disaster provide observation data of human behavior in disasters. These data are collected in qualitative research and measured and experimented with in computational social science techniques such as social network analysis and agent-based models.

Unfortunately, research on social networks in NWMD events do not exist. Without empirical data on social networks, analysis in this area is not possible (this is one reason we discussed in detail in Section 2 what 100 years of disaster research has given us). Instead, evidence on social networks using social network analysis (SNA) must rely on a proxy event, such as the 9/11 World Trade Towers attack and the Boston Marathon bombing evacuation and emergency response. Researchers used data on organizational networks to reveal the coordinated and emergent responses of existing inter-organizational networks in the case of the 9/11 World Trade Towers attack [241] and the effectiveness of organizational emergency planning in the Boston Marathon bombing [242].

The rest of this section describes the model we developed to explore and characterize the reaction of the population of NYC to a nuclear WMD event. We refer to the model as the NWMD model and we designed it building on the concepts introduced in Section 2, specially viewing a disasters through the lens of complex adaptive systems. The model captures the main properties of physical, social, and individual complex adaptive systems, specifically heterogeneity, webs of connections (i.e., social networks), relationships and interactions, and adaptations arising from individual actions and decisions in the modeled physical environment, including the nuclear weapon.

The description of our methodology is in two major parts, the creation of a synthetic population for use in the model and the development of a model implementing agents representing individuals of the population and how they are expected

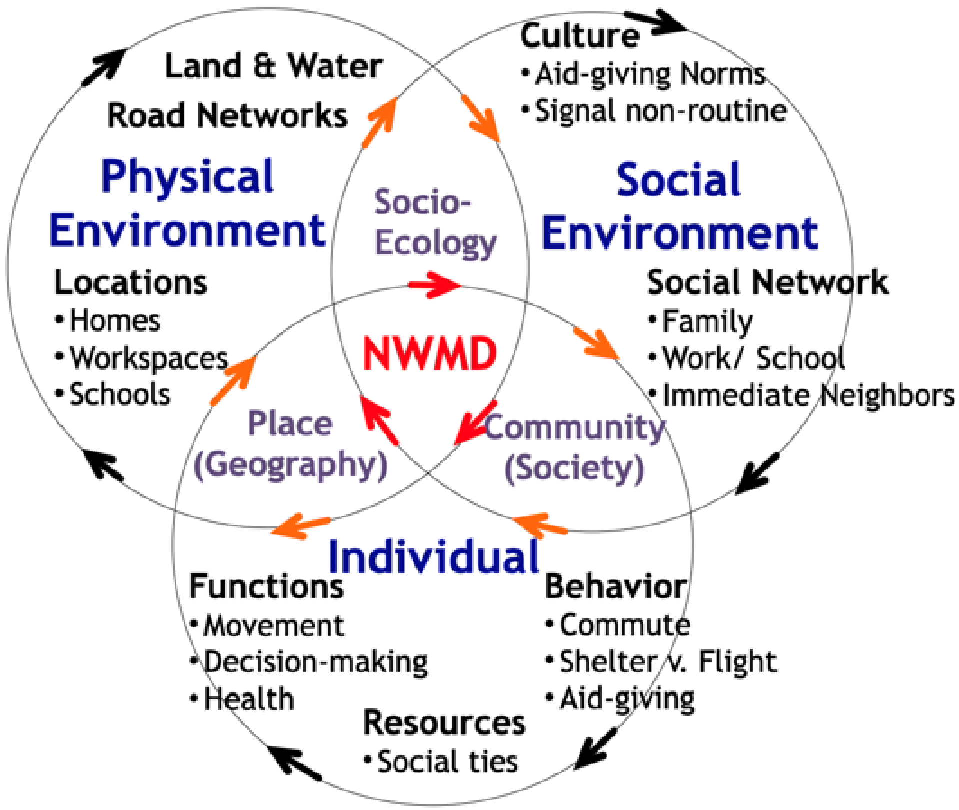


Figure 3: Complex Adaptive Systems (CAS) of the Nuclear WMD (NWMD) ABM Model.

Table 1: Multiple Data Sources Used in Our Study

Data Type	Characterization	Dataset	Granularity
Space	Road networks	2010 U.S. Census TIGER Shape files	Primary, secondary and local road
Population	Household demographics	2010 U.S. Census tracts	Census Tracts
School	School	US Environmental Protection Agency (EPA) Office of Environmental Information (OEI)	Geolocations Coordinates
Workplace	Establishment sizes	2010 U.S. Census Bureau's County Business Patterns	County
Commute	Commuting flow by tract	U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES)	Census blocks aggregated to census tract

to behave routinely and in reaction to a nuclear WMD event. We created our heterogeneous population (Section 3.2) which also gives a brief overview of synthetic populations and how we advanced these efforts. Next we turn to how we used a computational technique called agent-based modeling which allows us to capture individual interactions and decision making at the micro-level and how this leads to more macro patterns emerging (Section 3.3). Similar to the synthetic population we give a brief overview of agent-based models that have been used to study disasters and discuss our innovations here as well. The discussion will demonstrate that the physical, social, and individual systems were not created independently, but to work together as an integrated model with compatible levels of modeling detail.

3.2 Synthetic Population Generation

3.2.1 Synthetic Population Generation Background

Although, we have entered the era of big data and now have resources to acquire data from a variety of applications (e.g., cell phones, social media), there is still one area where we have limited ground truth data, such as data related to individual people. There are multiple reasons for this lack of data, such as the privacy regulation, the code of ethics and the expense to collect such a dataset [243]. Due to the lack of individual level data and increasing demands of this type of data, methods for creating artificial or synthetic population are now widely being discussed by scholars (e.g., [244, 245, 246, 247]). In addition to academic research, several research organizations (e.g., US Census, RTI international) are placing considerable efforts to create synthetic populations to support various research efforts including that of computational modeling.

However, many synthetic population datasets do not include social networks, even though several studies have indicated that social networks play an important role in how societies interact [237] especially in times of disasters as was discussed in Section 2.2. Moreover, social networks is also playing a more important role in field agent-based modeling as they allow us to study the connections between different actors (which we will revisit in Section 3.3). Hence, along with the generation of synthetic population we would argue that synthetic social networks should be generated at the same time.

As we can see, scholars and organization have filled the gap to some extent, however, the motivations for them to create synthetic population datasets are different. As synthetic population datasets are often generated to fit in specific research purpose and more than often such dataset often has restricted uses (e.g., traffic simulation [248] and disease spread [247]). When different research questions are brought up, it's hard to reuse those synthetic population datasets. As a result, one challenge remains, that of creating and sharing realistic synthetic populations which incorporate social networks. Therefore, this project introduces a mixed method that creates a reusable synthetic population dataset incorporating social networks, which aims to overcome this challenge.

Current population synthesis methods in agent-based models originate from microsimulation techniques [249] and involve a two-step process of fitting a population to a set of relevant attributes and constraints and then generating individual units on the fitted population [250, 245, 246]. Traditionally, population synthesis methods can be broken into two: 1) synthetic reconstruction (SR); 2) combinatorial optimization (CO) or re-weighting [251, 252]. As for SR, this method involves obtaining the joint distribution of relevant attributes and using Iterative Proportional Fitting (IPF, [253]) with the sample population used to create a fitted population and generate individual units on that population. While, CO involves creating a population and modifying it with the sample population until it meets a threshold of required constraints [254, 246]. Both methods have their advantages and disadvantages. For example, combinatorial optimization can minimize errors by using constraints of by using constraints extracted dis-aggregate datasets, such as, Public Use Microdata Areas (PUMAs) and Samples of Anonymized Records (SAR) [255, 256, 257]. As mentioned, dis-aggregated level data are required for methods to minimize the error during the synthesis process. As a result, creating synthetic population with limited data resource (e.g., lack of dis-aggregate level data) is remained as a challenge.

Synthetic populations has been applied to study the behaviors of large population such as demography, transportation, ecology, epidemiology and policy analysis, meanwhile, modeling approaches including cellular automata, micro-simulation and agent-based modeling are used as the artificial environment to utilize those synthetic population datasets (e.g., [258, 243, 254, 245, 259, 248]). Within agent-based modeling specifically, synthetic populations and the social networks of the agents are used to explore a wide range of topics including epidemiology [247, 260], power structures [261], diffusion in networks [262, 263], common pool resource governance [264], rumors and riots [265], evacuation [246] and safety-nets in socioeconomics [266]. With the growing demands for synthetic social network in agent-based modeling area (as we needed for this this project), there has been much interest in generating realistic synthetic populations based on data within the agent-based modeling community [250], little attention has been given to how to incorporate realistic social networks for a given population based on actual real world demographic information. Even, there are many agent-based models that utilize social networks (e.g., [267, 268]), most of these networks were grown during the simulation (e.g., [265]), use stylized networks (e.g., [269]), or simply assume adjacent agents are part of

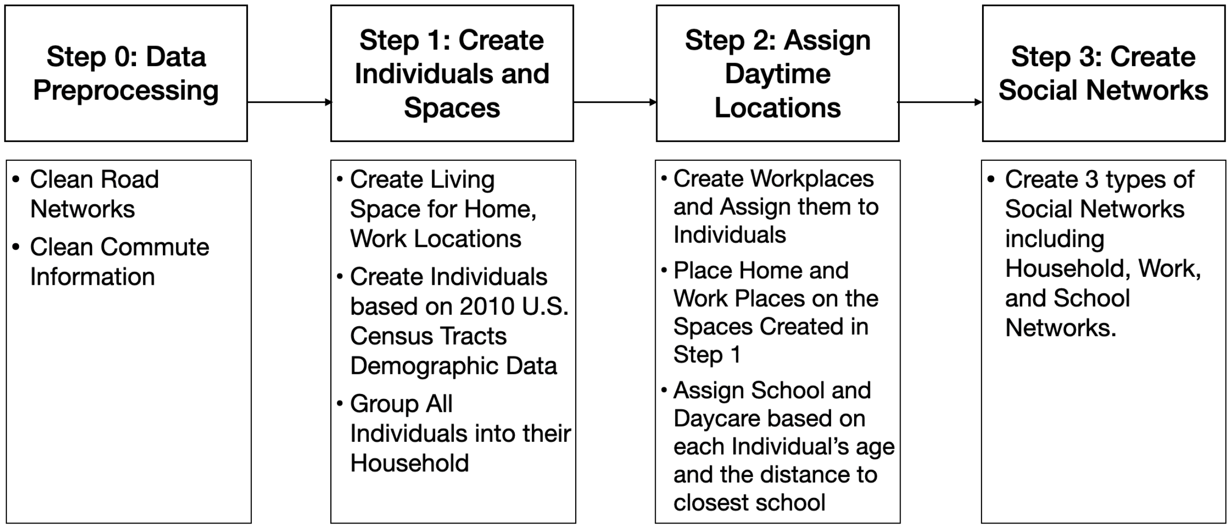


Figure 4: Workflow of process for population and social network synthesis.

the same network (e.g., [264]). Others have created synthetic populations with realistic social networks (e.g., [270]), however, the agents themselves operate within a network and are not geographically explicit. Others have created geographically explicit synthetic populations with social networks but the agents social networks do not evolve over time (e.g., [246]). Hence, another challenges in this area is to create realistic synthetic social network along with the population.

As discussed above, most of the synthetic population datasets and synthetic social network datasets are generated based on specific research purposes, which may obstacle the reusability of existing synthetic population datasets. Although accessible synthetic population datasets like RTI synthetic population [257] can be used to create the social network directly, low accuracies of those datasets are the main reason for us to propose our own method to create synthetic population and the social networks. To summarize, works done by other scholars provide the opportunities to create a method integrating current population synthesis methods and available social network algorithms [271, 246, 251]. By doing so, it's possible to generate an accurate reusable synthetic population dataset that incorporate multi-level information (e.g., individual level, household level) and its synthetic social network. Therefore, a mixed method that creates a reusable synthetic population dataset with social networks is introduced in next.

3.2.2 Synthetic Population Generation Method Detail

This section discusses the details of each steps in our synthetic population generation method which is outlined in Figure 4. This method is derived from the works of Barthelemy and Toint [251] and Wise [246], but adaptions and adjustments are made during the synthesis process shown by Table 2, which we discuss in detail in Step 1 and 2. Overall, we first carry out data prepossessing of the road network to ensure all road segments were connected into one network (Step 0). Next, we created spaces on the road network to place home and business sites based on the road type (e.g., primary and second road). We then created individuals using synthetic reconstruction based on the 2010 census data and grouped them to households. In step 2, Individuals in households are then assigned workplaces consistent with the data from the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES). Furthermore, we also assigned younger household members to the closest education institutes based on their ages (e.g., daycare, elementary, middle, and high school) whose locations were sourced from US Environmental Protection Agency (EPA) Office of Environmental Information (OEI). Lastly, three types of social networks were created based on being in the same household, working in the same workplace, or attending the same education institute. Household networks are fully connected, while their family members' school and work networks are based on interactions with individuals in these locations. All types of data mentioned above were listed in Table 1. Except certain part of the Step 0 in our method uses GRASS, the rest of our method is coded in Python. The source code, source data, and results data are shared. (Code: `https://github.com/njiang8/Create_Synthetic_Population`; Source Data and Results Data: `https://osf.io/3vsaj/`).

Table 2: Modeling Approaches: Adaptions and Differences

	Barthelemy & Toint [251]	Wise [246]	Our NWMD Model
Create Space	Non-Geographically Explicit	Based on OpenStreetMap	Based on Census Tiger Shapefiles
Create Individuals	Based on Multiple Aggregate Level Data	Based on Age Group from Census Data	Based on Age Group from Census Data
Group Individuals	6 household types	12 household types	12 household types
Assign Daytime Location	No Such Characters	**Work:** Based on Census Bureau's County Business Patterns; **Education:** No Such Characters	**Work:** Based on Census Bureau's County Business Patterns; **Education:** EPA Office of Environmental Information
Create Social Networks	No Such Characters	Used Ego Network and Social Media data	Use Newman-Watts-Strogatz Small-World Network

Figure 5: Giant connected component road network for Ulster and Sullivan counties.

Step 0: Data Preprocessing In this step, a basic environment is created with a transportation layer built from road network data provided by 2010 U.S. Census TIGER/Line Shapefiles [272] which is used to identify the primary and secondary road systems of the whole study area. This information was merged to create a single giant connected component road network file for our study area. Figure 5 displays the road network in Ulster and Sullivan counties of New York. To clean the road data file and create a network topology, we used GRASS (Geographic Resources Analysis Support System) C++ code libraries (also available in QGIS software). The process included simplifying lines, snapping lines to points, breaking lines at each intersection, removing duplicate geometric features, and removing small angles between lines at nodes.

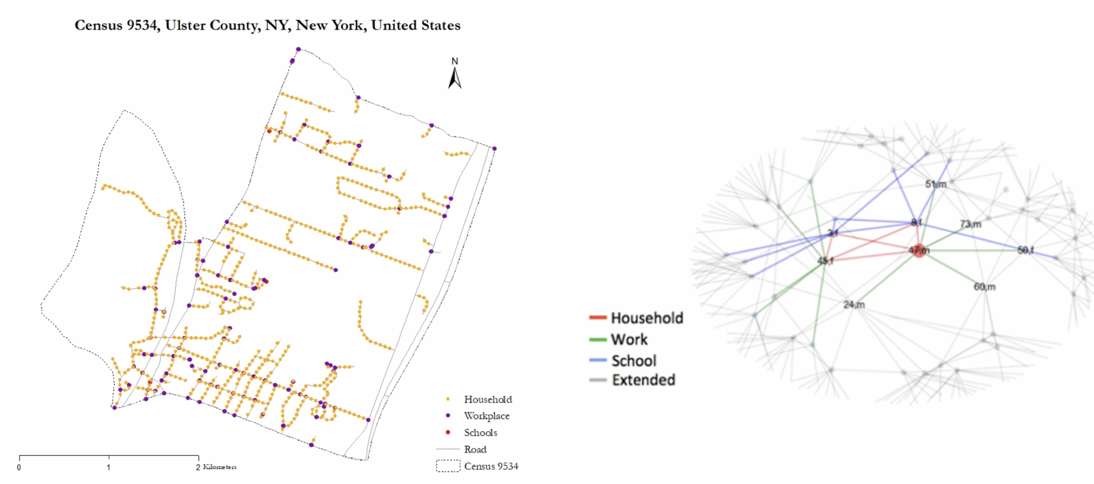
Step 1: Create Individuals and Spaces To create individuals, high-performance computers may allow us to apply combinatorial optimization without intensive computational efforts, but this method requires disaggregate level data to minimize the error during the process, in which the accesses to those data are limited in our study area. Under the constrain of limited data, synthetic reconstruction was selected to create individuals by only using census tract level data. Hence, we created individuals to represent every person within every census tract and assign their sex and age based on information from the U.S. 2010 Census data [273]. Similar to Barthelemy and Toint [251] and Wise [246], our method grouped all into households based on the household types present within a tract and on normal (Gaussian) distributions, but we added constrains related to age differences among the members under the same household to minimize the error between synthetic dataset and Census data, which distinguished our method. In addition, these constraints only apply to household type of family with children under 18. The specifics of the constraints are:

- husband's age – wife's age between (-4, +9)
- father's age - child's age less then or equal to 50
- mother's age – child's age less than or equal to 40

The U.S. Census categorizes households into 11 types: husband-and-wife families, male/female/non-family householders, households with a child less than 18, and male and female single householders over 65. Also, we added one more type for people living in group quarters, which can be institutional (e.g., correctional facilities for adults, juvenile facilities, nursing facilities/skilled-nursing facilities) or non-institutional (e.g., college/university student housing, or military quarters). We assume that for each tract, there's only one group quarters and those who belongs to living in group quarters all live in this location. Hence, there are total of 12 types households in our synthetic population.

Step 2: Assign Daytime Locations In this step, daytime locations (e.g., workplace and school) are assigned to every agent generated in Step 1. However, Barthelemy and Toint's [251] data did not include such character and Wise (2014) only considered workplace. As the results, in our method, home and work locations are generated and placed along the simplified road networks to make the synthetic population geographically explicit. Barthelemy and Toint's [251] synthetic population don't have the geo location. In consideration of general zoning, we restricted businesses to secondary roads with the exception of institutions like religious centers and schools that may be located on residential roads. No businesses are placed on primary roads as these are divided, limited-access highways [272]. In addition, as the lack of information related to the exact home locations or detailed land parcel information, houses are placed on local roads at least 50m apart or on top of each other when population density is high (e.g., representing apartment complexes) based on the number of occupied housing units in each census tract. Workplaces are randomly placed either onto secondary roads at approximate 20m apart or at local road intersections. The number of workplaces in each census tract is disaggregated from County level business establishment counts (and binned-sizes) from the U.S. Census Bureau's County Business Patterns [274]. To determine the size of the population in each workplace, we used a lognormal distribution within census tracts based on findings that job size distributions in U.S. cities are lognormal [275]. The information of commuting patterns related to the work places was derived from the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) dataset [276]. After we aggregated the information to the tract-level, we assigned work-age agents to a random workplace location within a tract based on the origin-destination statistics. Data for school locations are extracted from the Educational Institution dataset retrieved from the US Environmental Protection Agency (EPA), Office of Environmental Information (OEI) [277]. The dataset contains geographic coordinates of educational institutions, enrollments, grade levels, and start age and end age of each institution. We assign school-age agents to the nearest available school location within a tract. School-age agents are sorted into schools based on grade and enrollment levels. In Figure 6(a), a representative result of this step an example of household, workplace and education locations for one census tract within our study area.

Step 3: Create Social Networks To cover those needs mentioned above, we introduce a method that generates a geographically explicit synthetic population along with its social network by extracting information from existing empirical data. Three types of network are created based on living in the same household, working in the same workplace, or attending the same education institute. To add more detail to the synthetic population and social networks, the education network was divided into school and daycare networks. In Barthelemy and Toint's [251] work, social networks were not generated. While Wise's [246] work generated social networks, ego network and social media data were applied to imitate realistic social networks. Our approach is different in the sense, individuals receive a link to each agent located in the same household, work or education place. If the group size of a household, work or education sites is greater than 5 [278], a Newman-Watts-Strogatz [279] small-world network is generated. The resulting ties create individual and household multilayer networks and allow for simulation of the influence family members and group cohorts have on individual behavior. Figure 6(b) shows an example of the multilayer network within a home including an individual's familial ties within their household and proximity ties to people at work and school.



(a) (b)

Figure 6: (a) Mapped locations for Census Tract 9534; (b) Sample social networks of an individual within a household.

3.2.3 Results from the Synthetic Population Generation

While Section 3.2.2 provided details about the our mixed method approach using a series of Census data which generates a data set containing synthetic population living, working and going to education institutions within our study area, it is important to verify and and validate this method, which is what we turn to now. We discuss this here rather than in results Section 4 as this synthetic population was needed in order to build instantiate agent-based model (see Section 3.3. To verify our method, basic statistical analysis and tests were operated and several visualization results are presented in this section. For validation, our results are used to compared with two different benchmarks to test the robustness of our method.

There were some isolated errors within the official Census data for the region of interest. These were internal inconsistencies such as the total population not matching the sum of males and females within a census tract or one of those three figures being blank. A census tract subdivision county population which has approximately 4,000-8,000 people living there. To visualize the significance of these errors, Figure 7 shows the census tracts in the study area. However, the size of the areas does not indicate the amount of the population in question. 8 shows how small these errors are with respect to the overall population, less than 1 percent. The synthesized population represents 99.64% of the 2010 population at the 1:1 level within the model.

Synthetic Population's Social Networks The social network resulting from our approach consists of 22,921,302 nodes representing each person in the total population of the study area, and 29,382,541 ties representing relationships derived from people living in the same household or going to the same workplace or education institute. The majority of edges consist of household ties, and education ties represent the smallest portion of edges as shown in Figure 9 for one specific Census tract. Figure 10 shows the degree distributions for the combined network and each of its edge types. The multi-layer network represents one layer of ties created to represent household relationships, and one to represent relationships present at daytime locations. In the household layer consists of individuals in cliques ranging from 0 to 10 ties with the majority of the population in small groups of 2 to 4 individuals. Households with only one person represent singles living alone, and their social relationships are work related only. We also see that there are a few nodes in the workplaces with degrees ranging from 1 to 3 due to the small size of some workplaces. Because education institutes are occupied by groups of students, as expected we find no isolates.

Verification We verified our synthetic population by comparing selected measures from the official census data to make comparisons with our results. During this process, we found the number of individuals in our synthetic population was not identical to the 2010 Census data that shows just over 23 million (23,004,272) individuals living in the study area, show by Table 1. Looking into the issue, we identified 116 problematic Census tracts out of the approximately 5,500 tracts that had internally inconsistent in the original Census data such as the total males and females not matching the total population, or the no data provided for the number of individuals under 18 years old. Over the whole data set, there were 82,970 individuals living in those problematic tracts. Our method is not able to generate the individuals in those problematic tracts because of the inconsistencies mentioned above. However, the percentage of the individuals

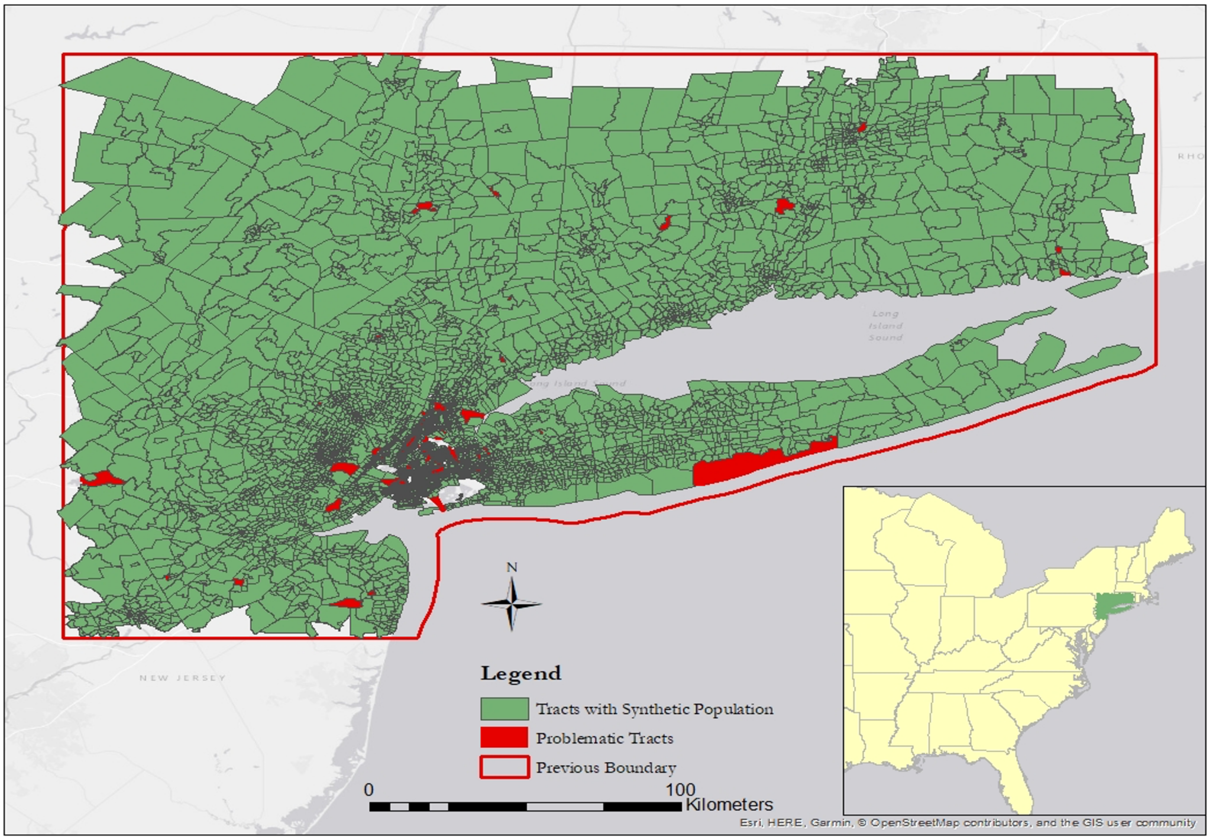


Figure 7: Tracts within Study Area with Internal Discrepancies.

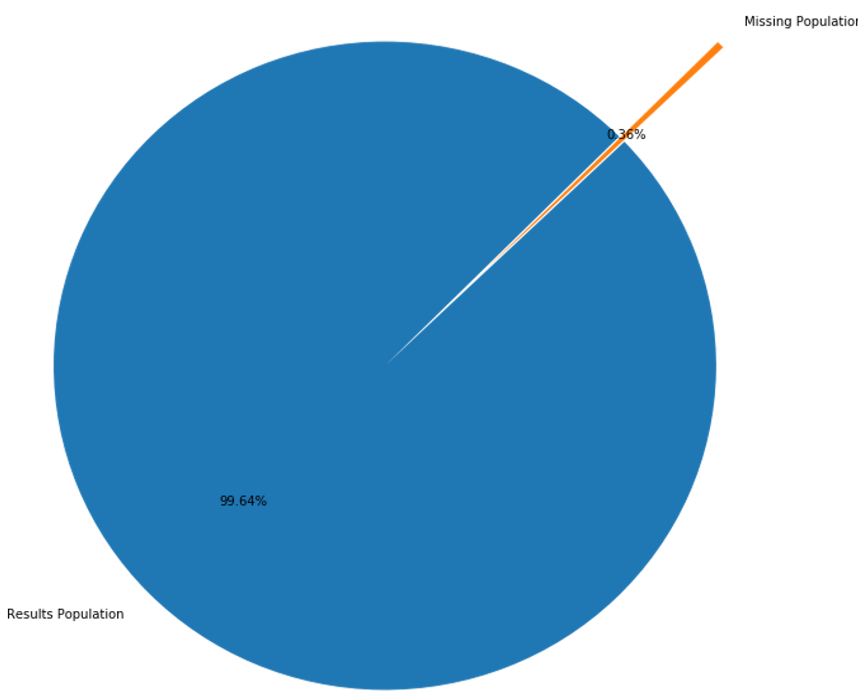


Figure 8: Fraction of Population Not Included due to Discrepancies in the Population Data.

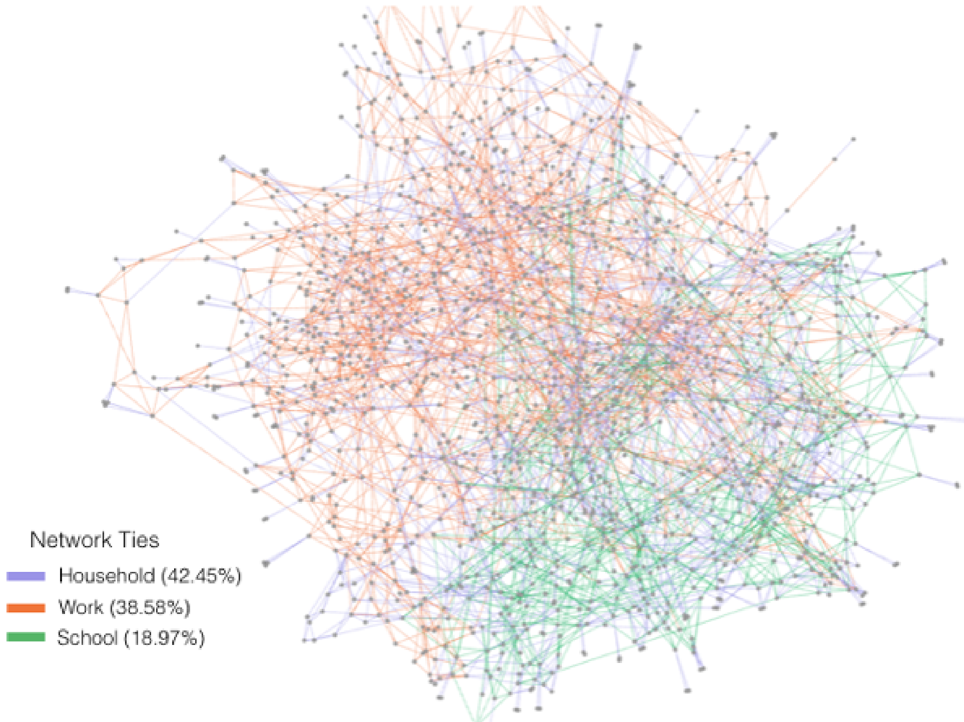


Figure 9: Synthesized Network of One Census Tract.

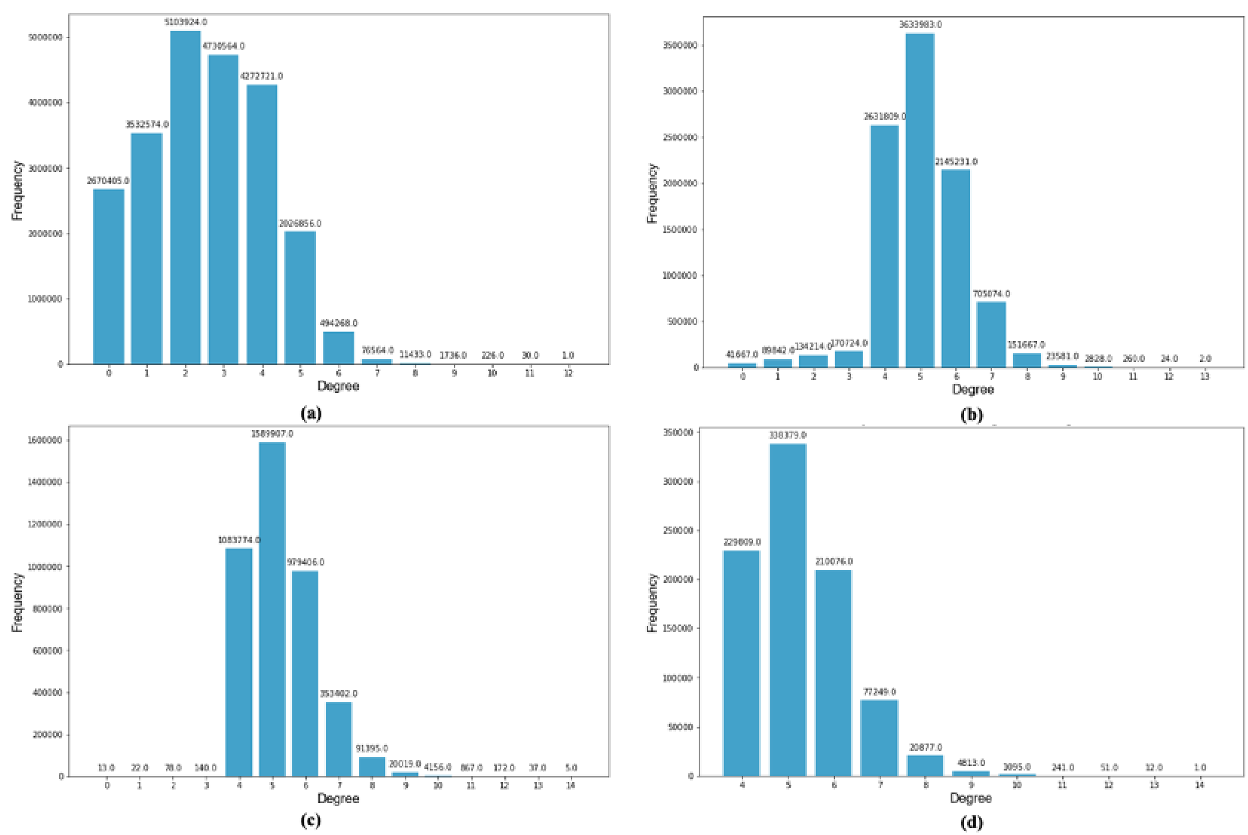


Figure 10: Network Degree Distributions. (a): household network; (b) work networks; (c) school network; (d) daycare network.

Table 3: Population Counts Based on Multiple Sources

	Original Census Tract (5506 Tracts)	Census without Problematic Tracts (5390 Tracts)	Synthetic Population
Individuals	23,004,272	22,921,302	22,921,302
Households	8,468,450	8,453,097	8,457,710
Male	11,108,508	11,053,104	11,053,104
Female	11,895,764	11,868,198	11,868,198

living in problematic tracts is only 0.36% of the total population of the study area. We decided to leave these tracts out of our synthetic population at this point. In the end, a total number of the synthetic population comes to 22,921,302 including 17,697,433 adults (age >=18) and 5,223,869 children (Age < 18), and they were grouped into 8,457,710 households.

Accordingly, as shown in Table 3, the number of total individuals, male and female, are identical to the non-problematic census tracts. However, the total number of households is slightly greater than the number recorded in the census data. This can be explained because we generated households for people living in institutional and non-institutional group homes. When we excluded the population lives in group quarter etc., we got the identical household amount of 8,453,097, which indicated that our method was able to generate the synthetic households without difference between the synthetic household and 2010 census data. RTI synthetic population datasets [280] were generated using IPF [257] with both aggregate and disaggregate data (e.g., PUMAs data), which is also synthetic reconstruction, but our method only uses census tract data. Our rational being is that we wanted to be apple to create synthetic populations when disaggregate data is not available. Therefore, while RTI data [280] has more attributes, such as income, vehicle number of each household etc., there is a 0.08% between RTI's synthetic household and census.

Validation Counts were also collected to ensure the synthetics population's realism, the baseline environment, and representations for the synthetic population is valid compared to other data from the empirical datasets. Two benchmarks were used to validate our population: Census [273], and RTI synthetic population [280] datasets.

First, we compared our results with the Census dataset by using a subset of three measures not used explicitly in generating the synthetic population: 1) the average household size; 2) the number of households with minors (under 18); and 3) the number of households with seniors (over 65). These measures were aggregated to census tract level to operate the comparisons. To capture the differences between our population and Census data, the percentage error is calculated by Equation 1.

$$Error = 100 * (SyntheticResults - CensusData)/CensusData \tag{1}$$

Figure 11 shows outer error percentages of these three measures, the differences on average household size ranged from -0.3 to 0 %, indicating the synthetic population varies only slightly from Census data. The difference for Households with minors ranged from -36 to 391 %, and the difference for households with seniors ranged from -60 to 234 %. Figure 12 shows the overall errors of two measures (households with minor and households with senior). As shown in Figures 12 and 13 displayed, among all tracts the average household size of our synthetic population stays constant with Census, as for another two measures major part of our results with the error less than 100% error. After analyzing the results in detail, only 10 tracts have errors greater than 100% under the measure of households with minors and 8 tracts under households with seniors.

Then, we compared average household size with RTI synthetic population dataset which is also using the 2010 Census to generate the synthetic population with the replication SR method [257]. The purpose is to test the robustness by comparing synthetic population from other organizations with the same generation method. As Figure 13 shows, the difference of household size between our synthetic population and the RTI population is less than 1, which is ranging from -0.22 to 0.78. Also, the V-shape dip in the plot is for the average household size in the Manhattan island of the New York city. Hence, our synthetic population result has minor variances compared with the RTI Synthetic population.

3.3 Agent-Based Model Development

In this section we outline our agent-based model development however, before we do this we want to give readers an understanding of agent-based modeling therefore we first give an brief overview of such a style of modeling (Section 3.3.1). After this brief introduction we discuss the overall design and resulting agent architecture (Sections 3.3.2 and 3.3.3). We then move onto the parts of the model, starting with the physical environment (Section 3.3.4). The next

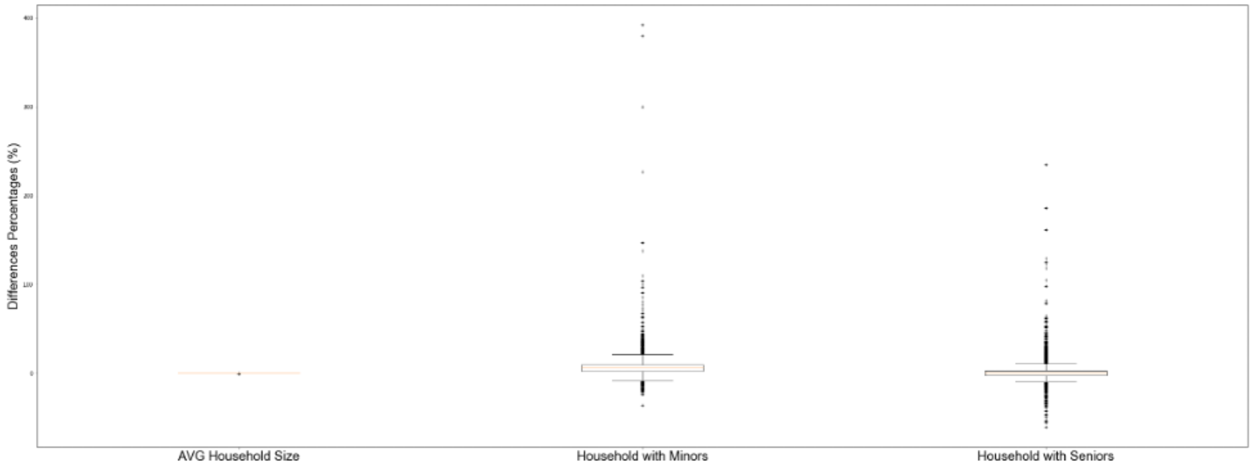


Figure 11: Box plots to show three measures: average household size, households with minors and households with seniors.

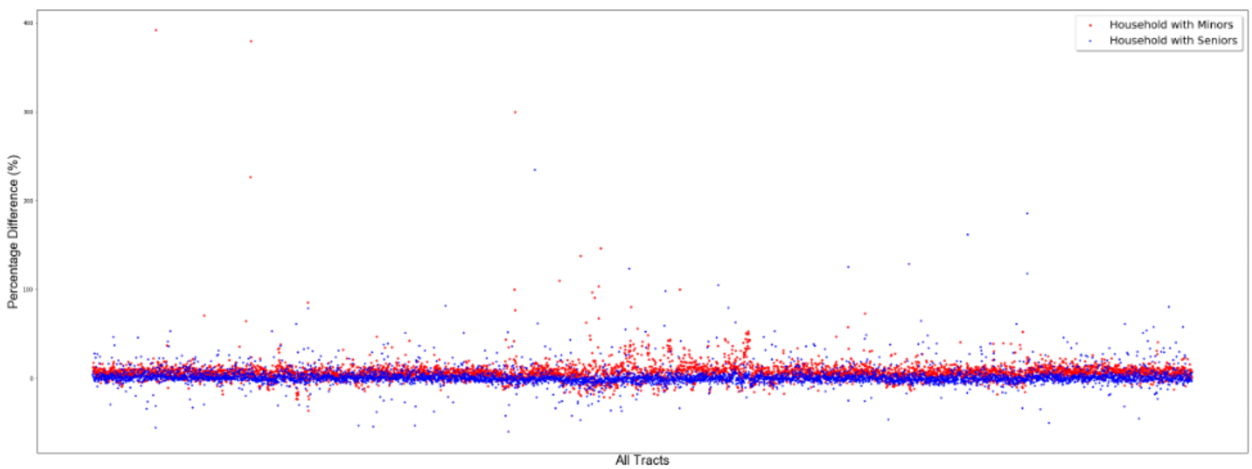


Figure 12: Percentage difference of synthesized population for each Census tract.

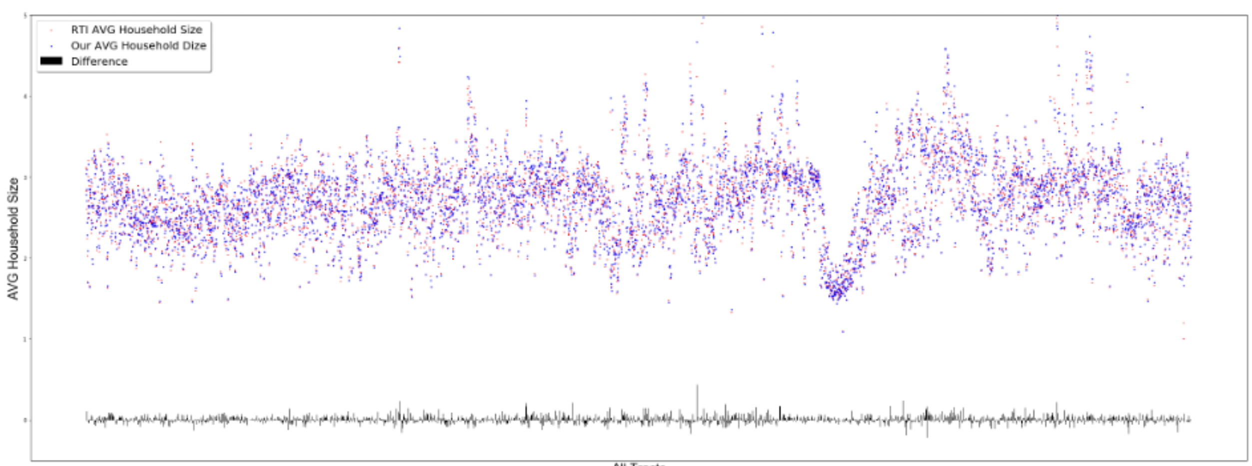


Figure 13: Validation: our Results VS. RTI synthetic population.

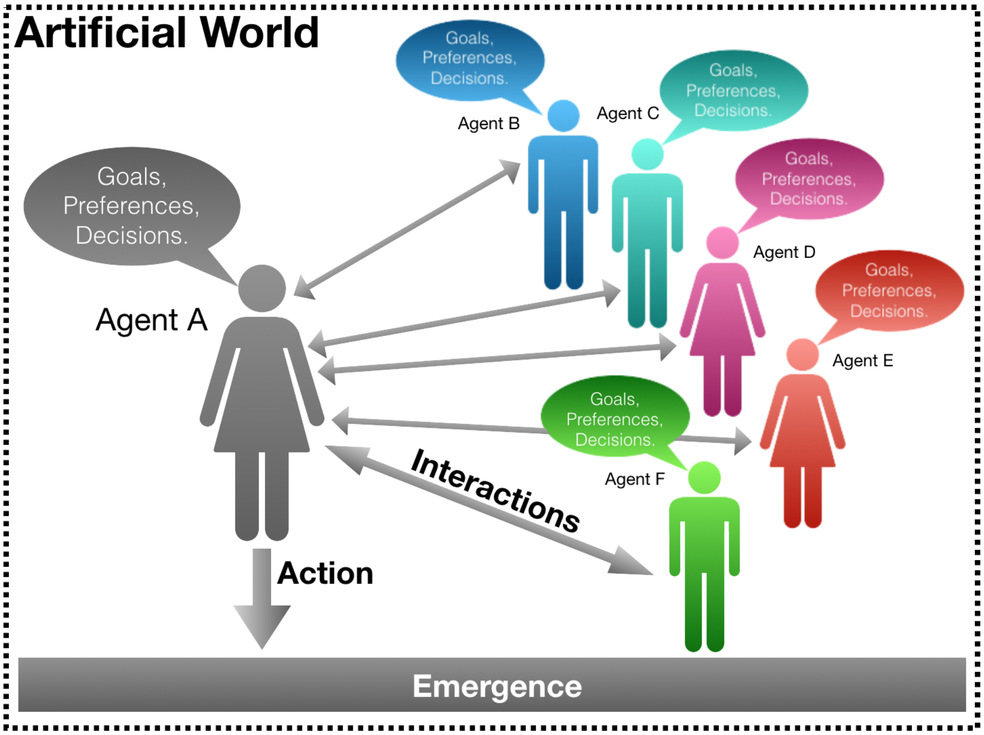


Figure 14: A schematic of an agent-based model, showing how interactions between agents lead to emergent phenomena within an artificial world (source: [283]).

sections discuss parts of the design, the design of agents (Section 3.3.5), their individual behaviors (Section 3.3.6) and their group behaviors (Section 3.3.7). The final section presents the NWMD scenario used to study the populations' reaction to a nuclear WMD event (Section 3.3.8)

3.3.1 A Brief Introduction to Agent-based Modeling

Before introducing our agent-based model of how society might react to a WMD, we would like to provide the reader with a brief overview of agent-based modeling and why we chose to use it in this project (similar to how we introduced Synthetic Population Generation in Section 3.2.1

Agent-based models were first developed in the 1970s and popularized with Shelling's [35] segregation model being one of the best know examples. In this model it showed how individual decisions in a heterogeneous population can lead to segregation. Within the virtual space of an ABM, individuals can form emergent groups, interact within communities and adapt with their environment. Since then, with the growth of computational power and data, agent-based models have evolved into one of the main modeling paradigms for modeling complex adaptive systems [281]. Agent-based modeling, as with other modeling techniques (e.g., spatial interaction models, micro-simulation) is a way to take the complexities of the real-world and, through abstraction, reductionism, and simplification, to focus on the important task at hand [282]. The main difference between agent-based modeling and other styles of modeling is that the focus is on interactions of individual entities and their behaviors, and how through such interactions more aggregate patterns emerge (e.g., how individual cars can lead to the emergence of traffic jams). Broadly defined, an agent-based model can be considered as an artificial world inhabited by autonomous and heterogeneous agents, each with their set of goals and preferences (which links back to notions of complex adaptive systems introduced in Section 2). It is through interactions with other agents that the agent makes decisions and decides what actions are to be carried out based on specific goals. These interactions lead to more aggregate patterns emerging as shown in Figure 14.

Apart from the individual entities within agent-based models interacting with each other, these entities are also interacting and are affected by the artificial world (or environment) which they inhabit, similar to how the physical world around us affects our lives. For example, take land use change. Developers may buy agricultural land, convert the land to residential use and then sell it to residents who then move into it (e.g., [284]). Agents can also perceive their environment and respond to it (e.g., changing climatic conditions may alter farming practices [285]). Initially,

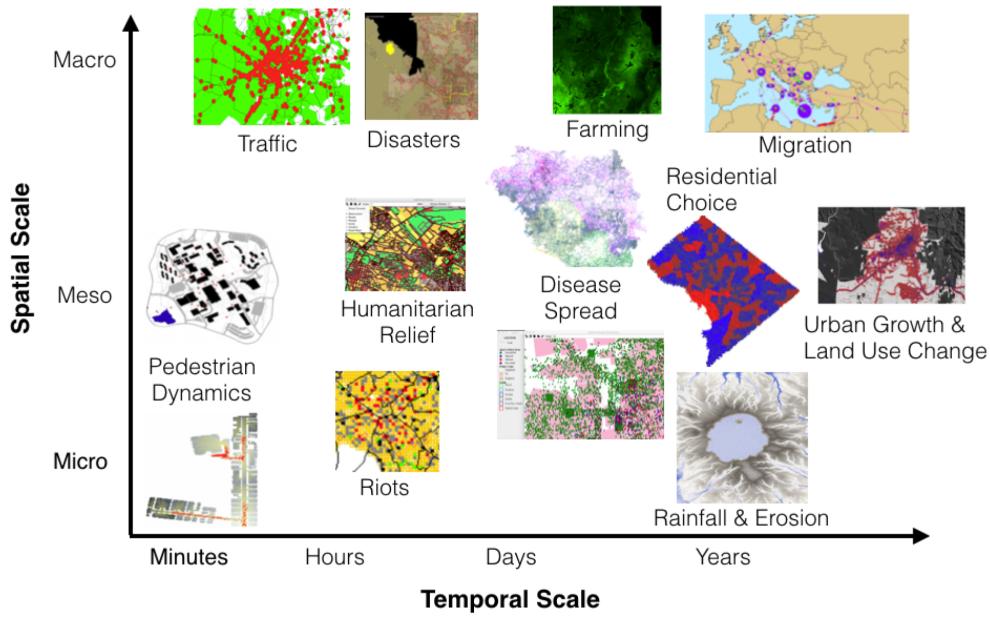


Figure 15: A selection of MASON Agent-based models models across various spatial and temporal scales (source: [283])

.

many agent-based models represented space rather abstractly such as the Schelling [35] model. However, perhaps with the demonstration of the Sugarscape model by Epstein and Axtell [286], which showed how the environment can affect agents' wealth and survival, modelers started to realize that the artificial world that the agents inhabited could be stylized on geographical data. From earlier works such as Gimblett [287], Benenson and Torrens [288] to current day work (e.g., [281]), researchers have utilized data not only to represent the physical aspects of the artificial world (e.g., land cover, road networks) but also to help inform the social aspects (e.g., census data to help with knowing how many agents live in an area) as we will show in Section 3.3.2. Such data takes the abstract representations of space and makes it more grounded in real-world locations as we show in Figure 15. As a result of this, agent-based models have been developed to explore the micro movement of pedestrians over second to macro migration over years and many things in-between (readers wishing to know more about the applications of agent-based modeling are referred to [281]).This notion of how we can ground a model to a real world location will be shown in will be further highlighted in Section 3.3.2.

It might therefore not come to a surprise to readers that agent-based models have been used to explore a variety of disaster scenarios. For example, Yang et al. [289] used home-work relationships for evacuation decisions and fell back on traffic patterns for group movement. However, one think many of these models lack is network dynamics integrated with their agent decision-making and spatial movement. For example, Haer et al. [205], and Widener and Gunter [217] only models flood and hurricane evacuations; Grinberger and Felsenstein's [290] disaster recovery model does not enable agents to travel through space; Wise [246] models fire evacuation; and Dawson et al. [291] models flood evacuation without social networks. This is something we overcome in this project.

3.3.2 NWMD Model Design

The conceptualization of disaster as a complex adaptive system (CAS) as shown in Figure 1, (Section 2.2), is adapted here for modeling a NWMD event in a mega-city as shown in Figure 3. This figure shows information and interactions of the three systems which make up this CAS (i.e., the physical (Section 2.1.1), social (Section 2.1.2), and individual (Section 2.1.3)), that is represented in our NWMD agent-based model. The framework is used as a foundation to organize the empirical data relevant to a detonation of a NWMD in CAS and the data collected in simulation experiments. By using high-quality data in a geospatially explicit model, the agent-based model can be characterized as an intermediate or mid-level agent-based model, a model that has theoretical basis with some applicability to real world problems. Models at this level can be used to test theoretical concepts in a simulation of the real world.

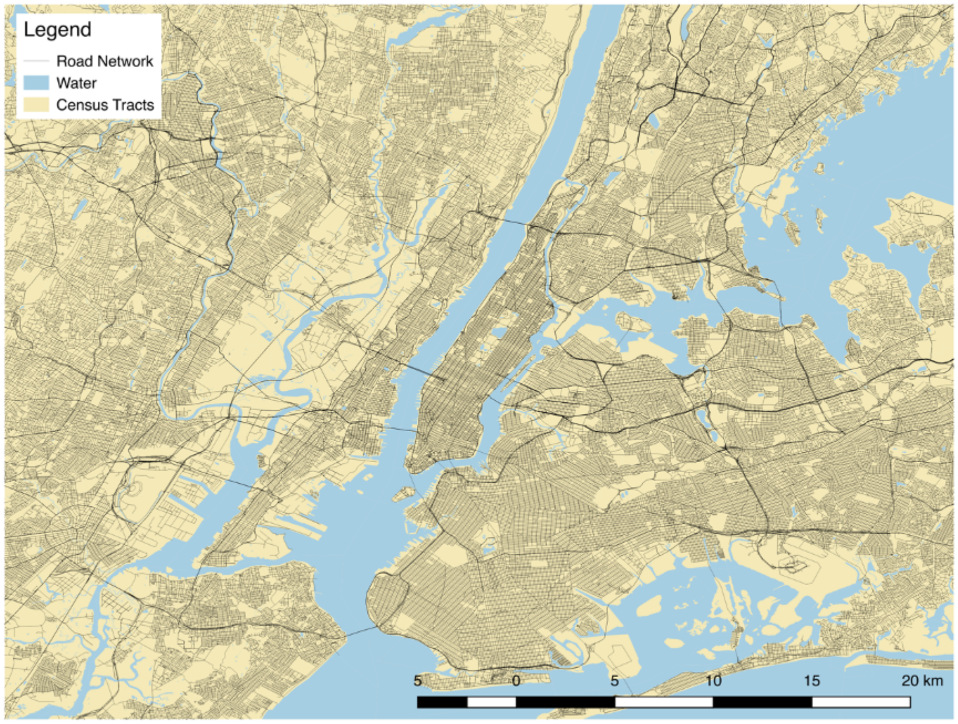


Figure 16: Small NYC Area Map.

So, with the intention was to develop an accurate representation of geographical space, the interactions of agents over time, and export social networks for subsequent analysis, we chose to use the Java-based, agent-based modeling framework MASON [292, 293]. It provided an industrial-grade, agent-based modeling development environment which separates the modeling from the visualization. We could run the model with visualization for development and without for large production runs.

The MASON Java-based agent-based modeling framework was chosen for its ability to import and create layers of geospatial data, so that agents can interact in a spatially explicit environment as well as interacting with each other. The modeled area includes the New York City commuter region to allow not only everyday patterns of commuting but also city-wide aid and rescue from unaffected areas to the area of destruction. The total population for the area at the time of the 2010 U.S. Census was approximately 23 million. We also used a smaller study area, approximately 11 million with 27,106 km of roads in an area of 47.6 x 63.1 km, as shown in Figure 16 for some experiments.

To simulate the population's reaction to a NWMD detonation, the agent-based schedules agent movements (representing individual people) and interactions on a 1-minute time step (which will be further discussed in Section 3.3.6). This timescale allows agents to simulate routine commuting, initial reaction to the NWMD detonation, and the emergence of victim groups after the event. Agents interact with each other and an ABM modeled environment (Section 2.1.1 consisting of a road network consisting of nodes and edges and Geographic Information Systems (GIS) standard shapefiles [294] that delineate census tracts boundaries and water (rivers, lakes, ocean), as well as the emergent groups of individuals that form based on the individual agents' shared tasks and activities as will be discussed in Section 3.3.7.

In the following sections, we provide and discuss the data sources used to create the model (Section 3.3.4), routine agent behavior (i.e., before detonation, Section 3.3.5), the modeling of the weapon's effects, the reactive behavior after weapon detonation, and finally the NWMD simulation scenario. However, if readers want to know more they are refereed to doctoral dissertations by our team members [15, 20].

3.3.3 NWMD Model Architecture

The NWMD model's architecture is diagrammed in Figure 17. At the heart of the model is the representation of the world were the agents and their virtual environment is implemented and the controls to the agent schedules. The world builder code builds the virtual environment (i.e., the artificial world) from input data files for the road networks, water

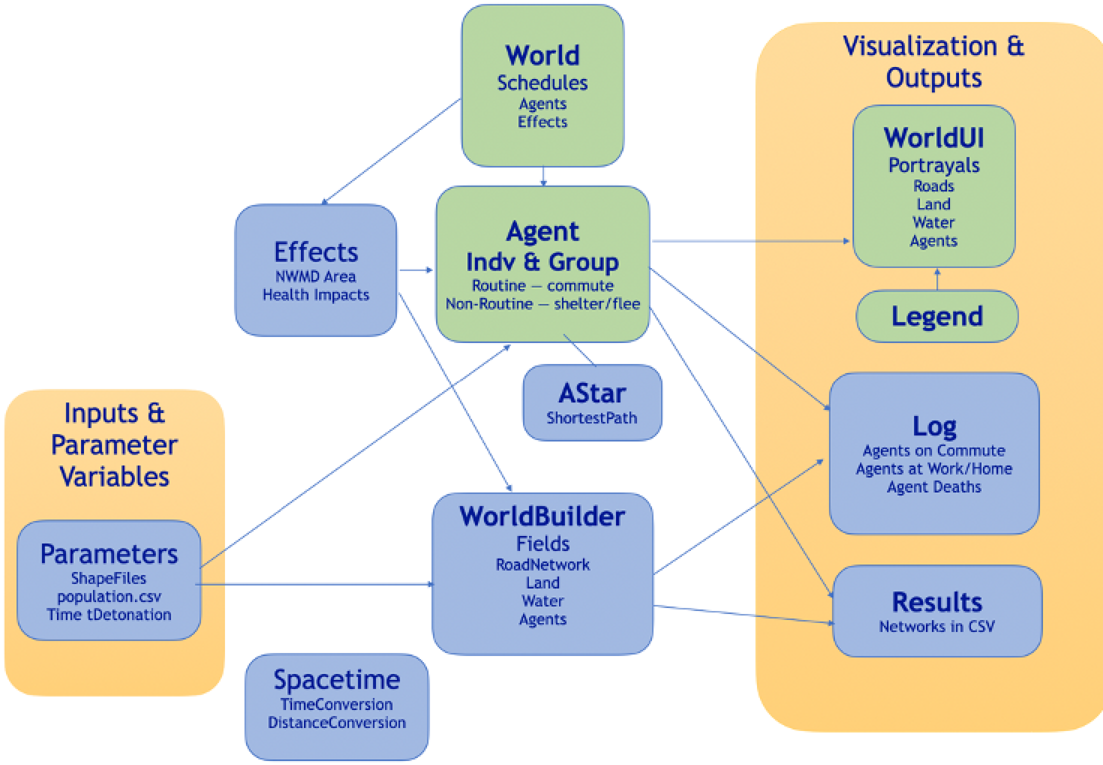


Figure 17: ABM Model Architecture.

areas, and urban landscapes as well as the agent population (from Section 3.2). The agents have various properties relating to basic health, location, and movement attributes. Movement along road networks is calculated using an popular A-Star algorithm [295], a simple method for route planning along a shortest path. There are two types of agents, individuals and groups. Individuals expand on the agent data with demographic attributes, household information, daily location goals (home, work, or shelter), and code for routine and reactive behavior (see Section 3.3.6). The second type of agent is the group. Group agents that simulate carpool groups for the routine commuting behaviors and emergent groups that form when agents flee from the NWMD detonation (see Section 3.3.7). An effects model simulates all the NWMD effects (see Section 3.3.4), mainly indicating the area of impact, degradation of agent health, and the destruction of nodes and edges on the road network.

All of the model data inputs and parameters are controlled in Parameters. This code allows experimenters central control of the model data and effects, specifically the map and population files, the NWMD effects, and turning carpool and emergent grouping behavior on and off. Outputs of the model include a visualization of the agents and effects, a log of agent-environment interactions as well as input and output data, and results data includes code for exporting model data. There is also code (Spacetime) to clearly describe and calculate model spatial and temporal dimensions, such as calculations for time steps and degree/time conversions.

3.3.4 Modeling the Physical Environment

The model includes representations of the land, water, road network, and census tracts as well as the effects of the nuclear WMD (which relates the the physical system of CAS as described in Section 2.1.1). As part of the generation of the synthetic population, we synthesized homes located on the road network as described in Section 3.2.2. In addition to homes for every household, we also generated work places, school locations, and day care centers. These physical locations were nodes on the road network.

These physical locations were also on a continuous representation of the land layer. Agent movements were continuous and between these locations traveling via the road network prior to the weapon's explosion. After the weapon detonated, surviving victims of the explosion were modeled as only able to move "on foot" until they got out of the damage area and could travel on the road network (which is further discussed in Section 3.3.6).

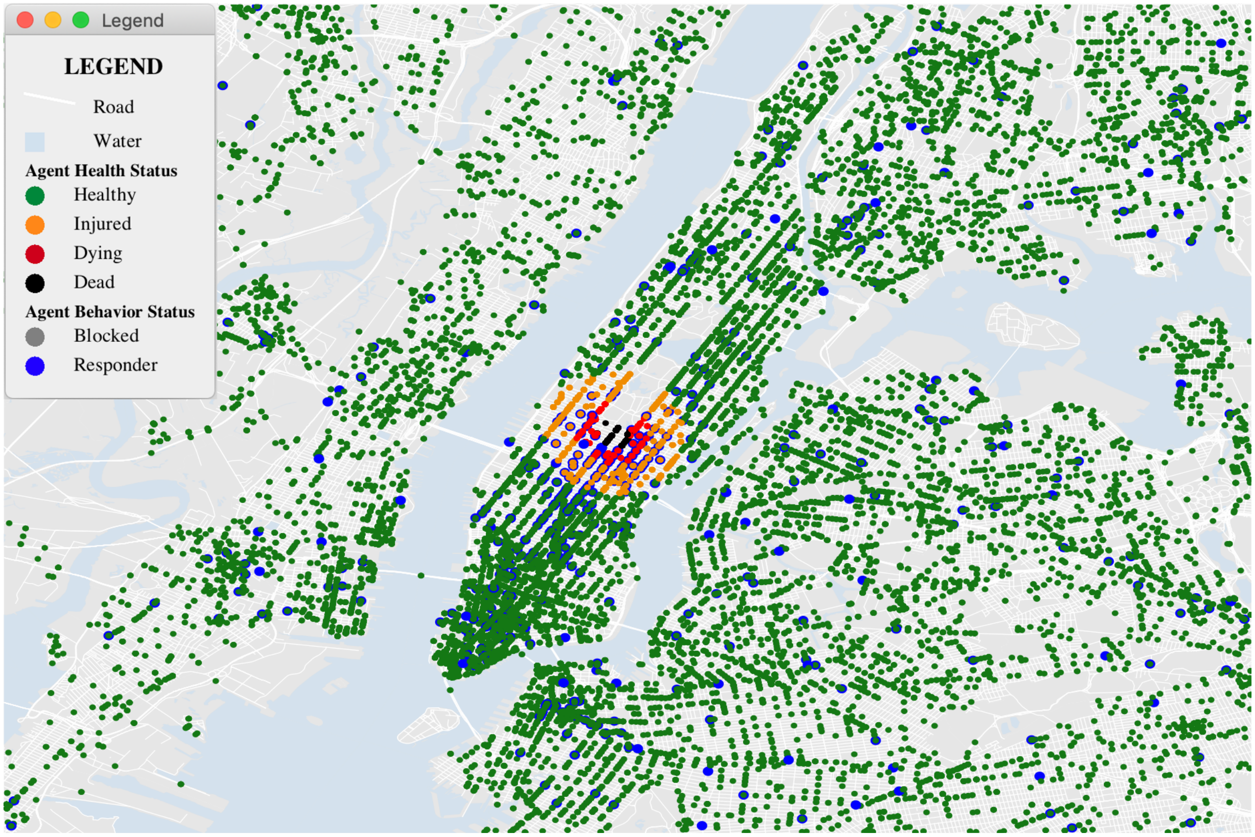


Figure 18: Modeled Effects of a 10Kt Nuclear WMD Just South of Central Park Showing Health Status of Population.

As part of the modeling of the physical environment, we included in a simple model of the weapon's effects. We used an approach with three defined rings at different distances from the location of the weapon. However, we did not use the physics-based descriptions but focused on the effects for the population. The inner ring is defined by the weapon's fire ball. Within this distance all the people are killed instantly. From the first radius to a second radius, the people are modeled as mortally injured, meaning they will die within a short time after the explosion. The zone between the fire ball and the second radius includes significant infrastructure damage, fires, and radiation effects. Vehicular traffic was not considered possible in this zone. The only movement of agents was simulated as on foot. We used a third radius to define a third zone between those mortally injured and those who will be injured, but are mobile and will survive. Outside the third radius, we modeled no infrastructure damage. These three radii and the location, time, and yield (size) of the nuclear WMD explosion were parameters for the model. We routinely used 430, 1,200, and 2,500 meters from the center for the three radii to model the effects of a 10 kiloton weapon. These infrastructure and biological effects on people were based on the publicly available reference, [296]. Figure 18 shows the three zones of a 10 kiloton nuclear WMD's and the weapon's effects on the affected population.

3.3.5 Modeling the Population's Individuals, Locations, and Social Networks

Synthesis of the model population was derived from the 2010 U.S. Census data, as described in Section 3.2, to generate a one-to-one representation of the entire population that covered not only New York City, but also near-by states whose state and local governments may be contacted for emergency support. The population's area of 262 x 234 km and 225,977 km of roads includes all of Connecticut and parts of Massachusetts, Rhode Island, New York, New Jersey and Pennsylvania and covers with over 22.3 million people. (See Figure 7.) Not only does this create the heterogeneity and complexity of the real-world, but place, space, and location can be used to integrate data across information systems [297], and thus, the qualitative and quantitative methodologies used in this report. This explicit model lays the foundation for future versions of the ABM to be used by researchers and policymakers.

The population was synthesized for the purpose of representing a population responding to a "no-warning" or "no-notice" disaster, such as a NWMD detonation, during the immediate impact and response phase. An Iterative Proportional Fitting (IPF) synthesis process [253] was modified and used to create a generic population of the area, and the

Table 4: Population Data Sets.

Population	% Full Population
22,921,302	100 %
229,544	1.0 %
22,960	0.1 %
2,298	0.01 %
406	0.002 %

synthesized population is intended for use in future modeling. The key modification the process was the addition of social ties (i.e., social networks) that capture the relationships relevant in a disaster scenario.

While nuclear weapons of mass destruction exist, thankfully they have only been used in anger twice. Therefore, there is little know about how people will react to them. Current social theory and empirical data emphasizes the importance of social networks for information and decision-making during disasters (Section 2). The control population is generated with connected household networks and small-world network ties to represent relationships between co-workers and school. Family ties are known to have significant influence on human behavior and decision-making in disasters [132, 182, 298], and at the time of an emergency humans form *ad hoc* groups to improve their chances for survival [191, 299, 300, 21, 182].

As such the synthesized population for disaster needed to include both individuals and the social networks most relevant to a disaster response, specifically family/household ties, schoolmates, and work colleagues. During the population synthesis individuals and their households are derived based on census demographics. In a later step of the synthesis process these individuals are assigned work and school locations based on county business and nationwide school and daycare data (See Section 3.1 for a description of these data sources). Immediate family and group cohorts from work and school are expected to be relevant to agent decision-making at the time of the disaster event and are represented in the model with social network ties. In the last stage of the synthesis process, social groups (i.e., social networks) were created based on those living in the same household, working in the same workplace, or attending the same school. Individuals receive a link to each agent located in the same household, work, or school place. If the group size of a household, work or school is greater than five [278], a Newman-Watts-Strogatz [279] small-world network is generated. Without this limitation household and coworker/schoolmate networks become too large to reasonably represent an agent's close social interactions. The resulting ties create individual and household multilayer networks and allow for simulation of the influence family members and group cohorts have on individual behavior.

For the exploratory purposes of the NWMD model, we sampled the synthesized population for a smaller commuter region as discussed in Section 3.2.3. This smaller region allows simulation of the commuting behavior and an area unaffected by the NWMD detonation from which first responders can be drawn. By decreasing the size of the population from 22.3 million, we also reduced the computational requirements for model simulation in experimentation. This smaller area includes approximately half the population size of the original population size and only covers parts of New York and New Jersey for a population size of 11.35 million. We used a separate random sampling algorithm to create four population data files to input agent data into the agent-based model. The respective population sizes and percent of the 1:1 synthetic population are shown in Figure 4. The algorithm sampled the synthesized population by household until an indicated population size is reached. Complete households were necessary to provide an environment with representative social networks at reduced population scales. The methodology for population synthesis discussed in this section builds on previously presented work [14].

Agents' Descriptions The agent-based model includes three types of agents; the NWMD, the individuals, and the groups. With respect to the NWMD agent, this agent is created in the effects code. When the NWMD detonates, it impacts individual agents and the road network edges and nodes in three ring zones. The detonation impacts ground zero at -73.977290 west longitude and 40.764290 north latitude, and the three ring zones ripple out at 430, 1,200, and 2,500 meters from the center. The individual agents and the road network are deferentially impacted according to these rings. These doses are calculated based on the publicly available reference of Glasstone and Dolan [296]. All agents in the closest ring, Zone 1, are killed. Every step after the detonation individual agents in Zone 2 are killed at random and the health of individual agents in Zone 3 is either degraded or randomly reaches the point of death. The stochasticity of deaths in these zones are used in lieu of physical models of both the specific health characteristics of individual agents and the detonation effects. The third type of agents, group agents, are made up of individual agents following an individual leader. Individual and group agent actions are driven by fast-and frugal heuristics [7, 301] through which agents determine when and where to move and whether to join emergent groups.

Table 5: Individual Agent Attributes.

Attribute	Type	Units
	Basic	
agentID	String	
healthStatus	int	1-10
dead	boolean	true/false
dose	double	
age	int	0-100
sex	String	m/f
tract	String	
county	String	
	Goals	
goal	String	commute, findshelter, shelter, flee
goalPoint	Coordinate	
goalNode	Node	Object
homeID	String	
workID	String	
hmRdID	String	
wrkRdID	String	
homeNode	Node	Object
workNode	Node	Object
	Schedule & Movement States	
StayAtHome	boolean	true/false
toWork	boolean	true/false
atWork	boolean	true/false
atHome	boolean	true/false
onCommute	boolean	true/false
fleeing	boolean	true/false
isHomeless	boolean	true/false
commutedist	double	kilometers
tcommuteStart	int	timestep
tcommuteEnd	int	timestep
tcommuteTime	int	minutes
commutePath	ArrayList <GeomPlanarGraphDirectedEdge>	List of Edges
	Grouping	
inGroup	boolean	true/false
isLeader	boolean	true/false
indvGrpID	String	
carGrpID	String	
idsToHouseMembers	HashMap <String, Indv>	agentIDs, Object
hholdnet	ArrayList<String>	List of agentIDs
emergnet	ArrayList<String> List of agentIDs	

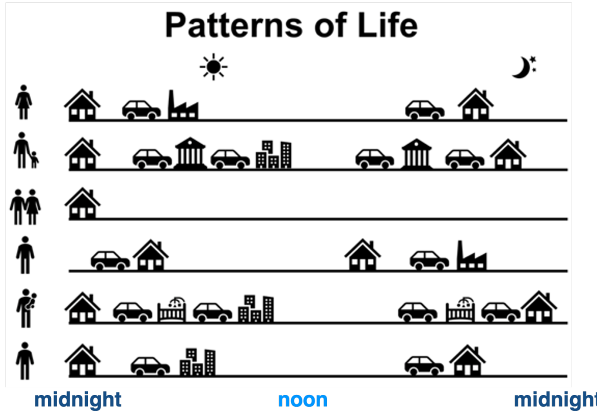


Figure 19: Sample Household Types, Makeup, and Commuting Patterns.

3.3.6 Modeling Individual Behaviors

Agent behaviors were basically of two types: routine and non-routine. The routine behavior were commuting to and from work places, schools, and day care centers. The non-routine or reactive behaviors were after the weapon exploded. People modeled as within the immediate effects of the weapons are modeled as victims. Those unaffected directly by the weapon were treated as trying to follow their "normal" routine behaviors. The following discussions of these behaviors provides more details.

Routine Behavior We now have the necessary representation of population to discuss normal commuting behaviors based on household makeup in detail. Households are situated on the road network, work places for working individuals within a household, and school and daycare locations are also set (See Section 3.2). The synthesized population has 12 different household types and their normal commuting behavior is established based on how many adults work outside the home and whether children need to go to school or day care. A sample of household types and make up along with their normal commuting patterns are shown in Figure 19. Some agents simply commute to and from their place of work during daytime hours. A single parent with a young child may need to drop off the child at school before going to work and picking up the child at the end of the day. Some households may be retired or working from home and not commuting at all. Still others may work a different shift starting each day at work and commuting home to sleep. These possible patterns are based on the household type, make up in ages of the members and their daily commuting schedules.

Reactive Behavior No recent actual nuclear weapon events can be used as empirical evidence for deriving agent behavior in the model. Therefore evidence from similar events have been used to guide the individual and group actions (one of the reasons we went into detail in Section 2 about disasters at large). An NWMD detonation is envisioned as a no-warning, man-made disaster, so we analyzed evidence for no-warning, man-made disaster events that were either nuclear in nature or occurred in NYC. Specifically, the atomic bombings of Nagasaki and Hiroshima and the 9/11 World Towers attack in New York City were used as approximate empirical evidence for the agent population's reactive responses. After the bombings of Nagasaki and Hiroshima, victims who were ambulatory moved away from the impact areas and displayed the symptoms of "shock" and "awe" [302, 303] that are found in many disasters, as discussed in Sections 2.1.1 and 2.1.3. In more recent disasters, such as the evacuations around the nuclear facilities at Three-Mile Island in Pennsylvania and Fukushima, Japan [304, 305], people left these areas as household groups. Perhaps, the best evidence of individual and group behavior for the no-warning, man-made event in NYC is the 9/11 World Towers Attack. New Yorkers responded without panic and banded into emergent groups, relying on family and other social ties for support with evacuation and shelter decisions [239, 191, 306].

The model of individual agents represents how individuals in the agent-based model routinely behave and later in the simulation respond to the NWMD detonation. These agents provide the basis for bottom-up behavior and dynamics in the model using fast and frugal decision trees [301]. Individual agent attributes are provided in Table 5. Simulated behavior is characterized by their daily routines consisting of staying at home or commuting to work and school (Section 3.3.6) and their reactive response to the NWMD detonation [303, 302, 155, 191, 239, 306]. Each step, each agent decides between routine and non-routine behavior. If the agent is healthy, it engages in routine behavior either commuting in a carpool or on its own. During their commute, agents travel via car along a commute path on the ABM's road network, represented as nodes and edges. If the nuclear weapon has exploded, in the non-routing/reactive behavior

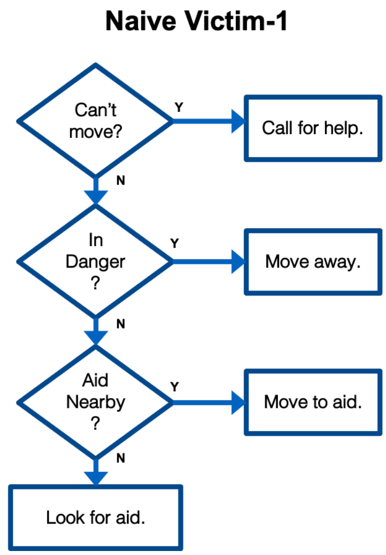


Figure 20: Decision Tree for Naive Victim (Simple Fleeing Behavior).

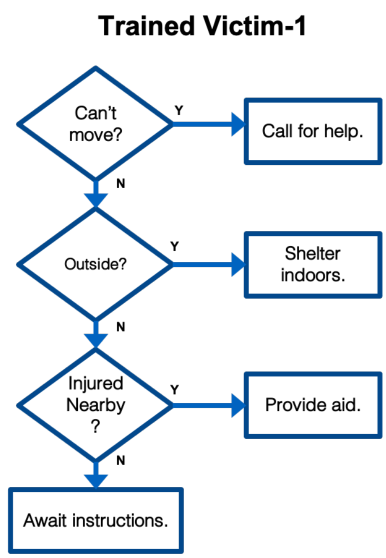


Figure 21: Decision Tree for Trained Victim (Following Guidance to Seek Cover and Await Instructions.

code, an agent decides whether to move toward ground zero as a first responder, join a group, shelter at home or work, or get away from the NWMD detonation. Health is used as a proxy to signal an individual's observation of the detonation which would realistically include not only any injuries, but also a flash of light, destruction of their environment, or information from other agents [303, 302, 306]. Decision trees are also used for agents to determine whether to join an existing group or make a new group. To join a group, an individual must have the same goal, whether to find join a group, shelter, shelter, or flee. If an agent changes its goals or dies, the code removes it from the group. Figure 20 is the decision tree for naive victims. It basically has survivors fleeing the area. Figure 21 is the decision planned for modeling victims following the Federal Emergency Management Agency's (FEMA's) guidance to seek shelter from fallout and await guidance when it is safe to evacuate the area. Figure 22 shows the decision tree for first responders.

3.3.7 Modeling Groups

During emergencies and disasters people turn to family and friends for material and emotional support, and the structure and composition of these networks have real effect on how well a community responds and recovers from these events [237] as discussed in [15]. In the response phase of an emergency or disaster, social networks are used for information

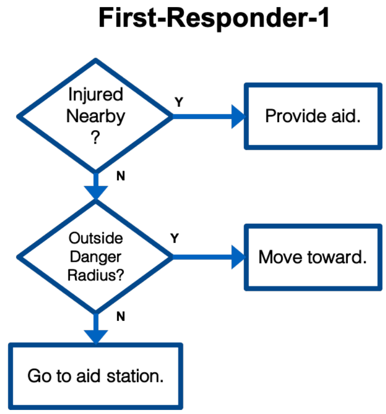


Figure 22: Decision Tree for Naive Victim.

and physical support as individuals, groups, and families decide to evacuate, shelter, or find and give aid [237, 238]. Social networks in this phase of an emergency are very dynamic as people scramble to find safety for themselves and loved ones. Often *ad hoc* emergent groups form temporarily with short-term goals to find shelter or to provide aid and rescue [133, 239]. Research in this area is rarely available because the collection of social network data during the response phase is prohibitive and after-the-fact accounts can be unreliable due to trauma [240]. Agent-based modeling can provide some insight into the dynamics of social networks in the response phase of a disaster when paired with empirical demographic and geographic data. In Section 2, we discussed how more than a century of disaster research has evolved into understandings of CASs framed as three CAS, the physical, social, and individual, interact in spaces that produce disasters. The social interaction of individuals in groups in disaster provide observation data of human behavior in disasters. These data are collected in qualitative research and measured and experimented with in computational social science techniques such as social network analysis and agent-based models. In this section, we first discuss empirical research in group behavior during disasters in Section 3.3.7, followed by some of the disaster work conducted using social network analysis in Section 3.3.7. The remaining Section 3.3.7 will introduce how these CASs can be used in an agent-based modeling experiment.

Group Behavior in Disaster Research Group behavior in disaster can be understood best from the perspective of the individual relationships that draw group members together in specific contexts. Drawing from decades of research Barton [21] defines the social units (individual and group) that make up social structures (i.e., networks) in disaster. These networks are the result of social relationships such as family ties that effect group behavior throughout the phases of disaster (preparation and mitigation, response, and recovery) decisions [131] or emergent citizen groups in response and recovery [133]. A significant body of research, expounded in Section 3.3.7, has been completed identifying and explaining group behavior arising from familial and kin relationships as discussed (e.g., [185]). Much of group behavior in disaster can be explained by these familial relationships and network relationships developed in normal routines such as work and daily personal encounters [239]. Individual and group decision-making can be considered a branch of group behavior in disasters in which the relationships and interactions between the individual and group influence decisions such as whether to evacuate or join search and rescue groups. Familial relations have significant effect on evacuation decisions [307, 298, 308, 121, 121, 309]. Information from non-familial relationships also impacts evacuation behavior. Neighborhood groups can also be a significant factor in evacuation decision-making as shown in studies of Three Mile Island Nuclear Event [304] and Hurricane Rita in Houston Texas [310]. Family and community groups can have both positive and negative effects on evacuation rates [311]. Technology has also become a factor in group behavior [312] as seen specifically in the effects of Twitter on evacuation from wildfires [313]. An interesting branch off this evacuation decision-making work is a study on the effect of emotional intelligence and group decision-making in emergencies [314]. The researchers integrated a conceptual model of group emotional with statistical approaches to enhance decision making in emergency environments. Less traditionally studied, are the group behaviors that arise from emergent or *ad hoc* groups form as a result of a disaster. These groups form, evolve, and disband based individual and group goals, and [315] theorized that emergent groups arise to address needs in a crisis, when they cannot be addressed in an existing organization. For example, research on social networks forming for the purpose of seeking help [238] or alternatively for the purpose of giving aid [142] and providing emergency response [110, 316] some of the roles that emergent groups play in disaster response and recovery. [317] found that emergent groups can be structured based analytic qualities; specifically, group domains, tasks, human and material resources, and activities. “Borantia” or volunteer groups in Japan are an example of an emergent group that has become the equivalent of the formalized Australian State

Emergency Service (SES) system [318]. In this case the emergent group filled a gap in emergency services that were formalized by the Australian government with the same tasks and activities. As an example of an event similar to the NWMD detonation in NYC being modeled in my ABM, emergent groups were a significant factor in the emergent response to the 9/11 Trade Towers attack in New York City [319]; 99% of people in these towers survived due to the help from emergent social groups [239]. Although the threat of nuclear weapons continues to be a concern, research and models regarding human behavior in NWMD events continues to be limited because there are few of these events from which to gather data. Empirical research on individual and group behavior specific to a NWMD event is limited historical information from the World War II (WWII) nuclear bombings of Hiroshima and Nagasaki, Japan in 1945. Group behavior in the WWII nuclear bombings reflected the unprecedented use of nuclear technology, and citizens of Hiroshima and Nagasaki had no previous experience or conceptualizations of the effects of a nuclear bomb. Therefore, the behavior response and evacuations of the towns reflected both their shock and their health status [320]. More recent nuclear events include the evacuation of the Pennsylvania Three Mile Island nuclear facility in 1979. Group behavior in evacuations resulting from nuclear facility failures depended on both the credibility of authority notifications and use of force as well as the influence of family and friends [304]. As found here and discussed in Section 2.1.2, the social networks of family and friends are significant factors in the survivability of people in emergency events such as NWMD detonations.

Group Social Networks Social network analysis has historically focused on the analysis of network ties with temporal snapshots of the social network, but today the study of spatial and temporal effects on networks is quickly growing both in applications such as vaccination debates [321], wildfires [313], and migration [322] and theoretical understandings [323, 324, 325]. As discussed in Section 2.1.2, social networks also perform important functions in emergency and disaster events [237, 326].

Unfortunately, research on social networks in NWMD events does not exist. Without empirical data on social networks, analysis in this area is not possible. Instead evidence on social networks using SNA must rely on a proxy event, such as the 9/11 World Trade Towers attack and the Boston Marathon bombing evacuation and emergency response. Researchers used data on organizational networks to reveal the coordinated and emergent responses of existing inter-organizational networks in the case of the 9/11 World Trade Towers attack [241] and the effectiveness of organizational emergency planning in the Boston Marathon bombing [242].

Groups Within Agent-based Models Agent-based models provide an experimental platform in which agents representing heterogeneous populations can act in a spatially explicit virtual world and have great potential for modeling CAS such as cities and disasters [327, 308]. Within the simulated space of an agent-based model, individuals can form groups (i.e., create networks), interact within communities, and adapt with their environment [267]; for examples see [265, 328, 13]. They can also use place-based variables in their decision-making that have significant impact on emergency and disaster recovery such as physical exposure, local government, local planning, citizen participation, and social networks [329]. By adding the dimensions of space and time in social simulations, ABMs become powerful tools for experimenting with dynamic social networks.

A common topic of group behavior in emergencies that accounts for space and time in agent-based models is the dynamics of evacuation behavior. One recent example of this body of work applied to disasters is a study that explores group dynamics and evacuation flow in building evacuations using a mechanical statistics paradigm [330]. [331] used an agent-based model to show the importance of social networks in emergency planning and response for evacuation before hurricanes, and [332] used an agent-based model to explore factors in an individual's decision to evacuate an area can minimize the outcome of a volcanic eruption. The only example of an agent-based model exploring an evacuation response to a nuclear event focuses on wireless communication [333]. The strength of this model is the use of cell phone data for validation of the simulation results. However, the ABM does not explore the influence of social networks on individual decision-making or the rise emergent groups. A growing body research uses computational social science methodology that integrate social network analysis and agent-based models [16, 281].

Previous ABMs of emergencies and disasters that include social networks are limited [290, 331, 246, 271]. For example, Grinberger and Felsenstein's disaster recovery model does not enable agents to travel through space; while Widener, Horner, and Metcalf only modeled hurricane evacuations with a static social network. Wise models fire evacuation and how social connections may impact the decision to leave, and Barrett et al. models social networks that do not explicitly capture interactions in space. None of these models include network dynamics in their agent decision-making. A challenge in this area involves the high computational costs of modeling social networks, especially those tied to geographical information. Research into improving modeling capabilities through high performance or distributed computing is ongoing [334, 335, 336].

Unfortunately, in addition to the challenge of a dearth of data that can characterize a disaster caused by a NWMD detonation, the collection, experimentation, and analysis of dynamic social networks that are not geocoded in social media remain a logistical challenge. Our solution is to develop an agent-based model that create virtual spaces in which agents (individual systems in Section 2.1.3) are modeled interacting with their social and physical systems (Sections 2.1.2 and 2.1.1). The design of the model can be conceptualized in the complex system of systems discussed in Section 3.1, and Figure 3 organizes the empirical data, system, and agent interactions in an ABM framework. The ABM developed in this project is designed to include location and social networks in the agent's decisionmaking process in a NWMD detonation and is described in Section 3.3.5. We demonstrate how an agent-based model that integrates social networks with a spatially explicit environment improves the fidelity of an emergency response simulation. The results of the model are presented in Section 4 and show how data of emergent networks changing over time are gathered in the model as well as the data that characterizes the response of agents to a NWMD detonation in NYC. The simulated behavior before the event is characterized by their daily routines consisting of staying at home or commuting to work and school. Their reactive response is triggered by the weapon's detonation. We turn to modeling of the weapons effects next

The weapon's effects code is run with each step and models the NWMD detonation's three blast effect ring zones (Section 3.3.4). At the time of the detonation, the code updates the health of all the individual agents and removes edges and nodes of the road network, depending on the zone. Agents in Zone 1, closest to ground zero, are designated as dead with their health status set to 10, and all the edges and nodes in the zone are removed from the model. The health of agents in Zone 2 is set in the range of 4 thru 6 depending on their distance from ground zero, and 50 percent of the edges and nodes are removed from the road network. These agents are modeled as receiving radiation that will kill them within 24-48 hours. In Zone 3, agents' health status are set in the range of 1-3, and the road network is left undamaged. These agents receive a dose of radiation that will possibly kill them in 1-14 days. All of the agents in Zones 2 and 3 have their goals set to flee. After detonation a random number of the individual agents in Zone 2 die. In Zone 3, a random number of individual agents die based on their dose of radiation.

3.3.8 NWMD Model Scenarios

In the NWMD model, each step represents one minute of real time, and at each step every agent is activated to make and implement their decisions, i.e., move towards a goal or stay in place. The one-minute time step was chosen as the timescale for the model because it provides a level of detail that sufficiently encompasses agent decision-making and movement over a 3-day period and in the NYC commuter area. A 3-day period provides a timeline that includes the detonation, initial responses, and the 72-hour period Urban Search and Rescue (US&R) are self-sufficient [337]. In the baseline agent behavior, i.e., prior to the weapon's detonation, agents follow daily schedules in a 24-hour day. Agents either stay at home or travel to work or school during their commutes. For simplification agents only travel by car and do not utilize any alternative forms of transportation such as walking or taking the subway, train, or bus until the NWMD detonation. Individual differences in the agent population with varying work-day time schedules and destinations create realistic patterns of commuter behavior (see 3.3.6). Figure 19 provides a sample of the variations in agent daily patterns based on household compositions with agents starting their daily routine at home at midnight. In experiments with our model agents all begin their commutes at 0730 (7:30am in the morning) and 1830 (6:30pm in the evening). Within the code, an experimenter can set the agents in the model to work 3 shifts, 0800-1700, 1700-2400, and 2400-0800, however these are not implemented in all results as the focus was on the aftermath of the NWMD.

In the agent-based model, the effects of the NWMD detonation are modeled as a ground burst and based on Model parameters which specify the location, yield, and time of the detonation, and the model then implements a ground burst at the specified time. The simulation scenario is based on a 10Kt weapon in Manhattan detonated during the workday. Damage to the area and agents are caused by the force of the blast, fire, and radiation. As discussed in this section above, the road network nodes and edges are destroyed in the zone closest to the detonation. In the second zone only 50% of the nodes and edges are destroyed, while they remain intact in the third zone. Agents' health deteriorates depending on the weapon size and its distance from ground zero resulting in the amount of radiation they receive. Blast effects are represented in three zones based on agent health status: immediately killed, mortally injured, and injured, but likely to survive. With the visual results of the model (see the ABM architecture in Section 3.3.3) agent health is displayed as a change of color: black equates to dead, red equates to dying, orange equates to injured, and green equates to healthy. Upon injury, agents begin to move away on foot from the impact area at varying speeds. Mortally injured agents walk slower than those likely to survive (as we will show in Section 4). A notional timeline of agent behavior and goals is provided in Figure 23.

As one would expect to see in a real-life situation, the detonation of a NWMD in NYC causes agents to alter their normal behavior and daily routines (as disscsed in Section 3.3.6. Agents whose health is directly affected change their goals from the daily commuting schedule to finding safety for themselves in a shelter. Depending on the availability of a shelter and agent knowledge, they will either choose to form *ad hoc* groups (i.e., emergent networks), find a temporary shelter, shelter, or flee from the blast, fire, and radiation areas [239, 191, 306]. If individual agents in a location have the same goals, they either join an existing group or form a new emergent group. These agents, who have not already

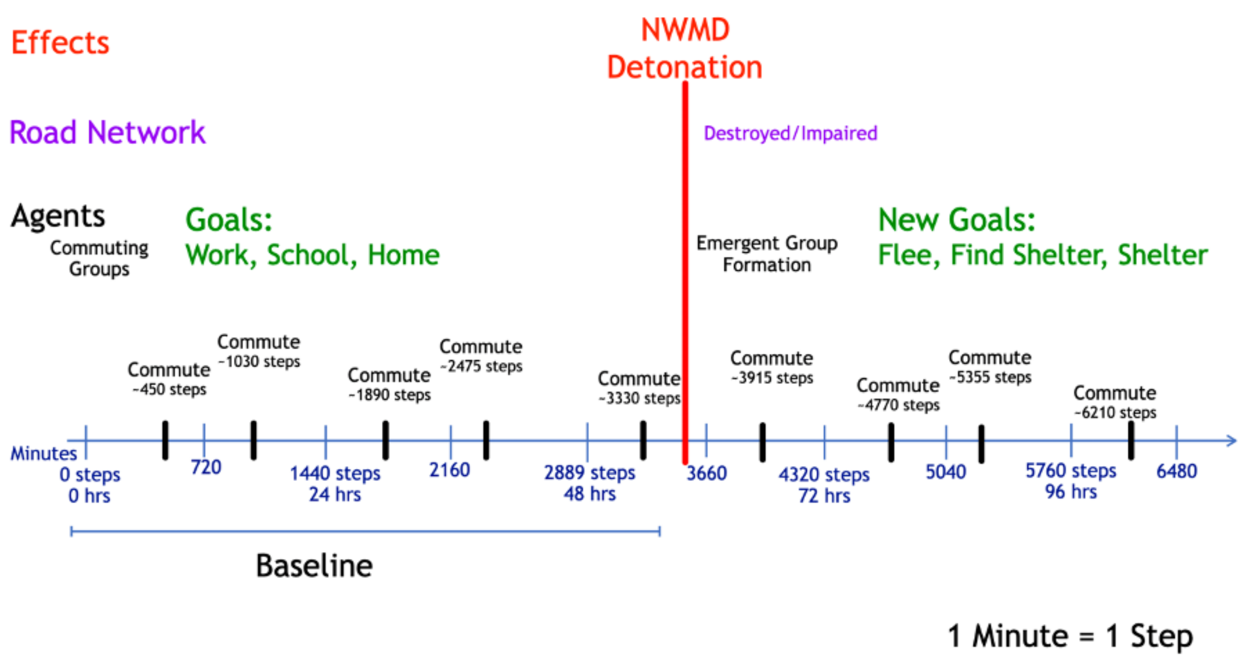


Figure 23: Timeline of a NWMD Detonation in NYC.

sheltered, walk to outside the impact zones and then return to the road network as they find and travel to a temporary shelter. Unaffected agents (i.e., those with a health status of 1) continue with their daily routines until they learn of the event, in which case they locate household members, adjust their commutes to avoid any evacuated areas, and return home. Agent behavior is currently modeled to represent only the first 15 minutes after a NWMD explosion. The timeline could be extended to simulate time period to days following the event, but 15 minutes are sufficient in this simulation because our intent is to show that the social networks of emergent groups in disaster can be studied using quantitative methods such as ABMs and social network analysis.

With this description of the model's design, we now turn to presenting and describing the results of several model runs after confirming the model runs as planned (verification) and that those results are representative of what would actually happen (validation).

4 Results

In this section, we discuss results from running the model described in Section 3. As a basic research project, the first topic presented as results is the verification and validation of the model. These processes were conducted throughout the research effort to maintain and confirm the quality of the model during its development and are discussed throughout this section. After the overview, we discuss the performance of the model's code as we scale up the number of agents instantiated in the model. Next, we demonstrate routine traffic (more confirmation that the model behaves as expected), and then behavior of the model after the nuclear WMD event. After the NWMD event, we characterize the reaction of the population overall as well as look closely at the formation of *ad hoc* groups a result of damage and emergency evacuation. Each of these topics is discussed in the below sections.

4.1 Verification and Validation

Our verification and validation of this project is discussed in two parts. The confirmation of the quality of the population synthesis process was presented in Sections 3.2.3 and 3.2.3. Here, the basic verification of the Agent-Based Model is presented and this process was done at several points during the project. The process included conducting code walk-throughs with other modelers, profiling the code's performance to ensure it was running as expected, and parameter testing (e.g., varying agent population size and seeing how this impacted commuting times, agent locations, and agent health status) to ensure the model was working as intended (i.e., verification). For the purpose of validation, we started with representative runs of the simulation to show the baseline of how agents travel from their home locations to their work or school locations based on a work-day activity schedule, i.e., routine behavior and these matched the real world. These tests are discussed in Section 4.2 below. These tests were done for a range of implemented population sizes. The effect of increasing populations can be seen in the figure. The agents traveling to and from work and school provide a baseline of normal behavior in the model through commuting individually and in carpools.

The validation of the agents' reaction to the NWMD event would normally be the comparison of the modeled reaction to available data from the real world. However, as discussed in Section 3.3.5 there are no directly equivalent NWMD events to use as model validation. Data from the Hiroshima and Nagasaki bombings, the only nuclear bombing events to date, was focused on the physical effects of the weapons [303], and the attack of the World Towers during 9/11 did not involve a nuclear weapon. Therefore, to validate the model's behavior in response to the scenario discussed in Section 3.3.8, we drew from evidence of the WWII, attack on 9-11, the reactions to nuclear-related disasters (Three Mile Island [338], Chernobyl [339], and Fukushima [305]), and the behavior described in the disaster research discussed in Sections 2.1.3 and 2.1.2. The model can be considered validated at a Level One level of performance and analysis is in qualitative agreement with empirical macro-structures [340]. The importance of verification and validation within the project will be demonstrated by their part in the discussion in the rest of this section on results.

4.2 Traffic Dynamics and Scaling

Agent-based models have been developed and used extensively in urban transportation studies [341]. So this project started with a traffic model. The synthetic population was an input to the model. It specified how to initialize the agents and their basic attributes (e.g., age, sex, home location, work location, etc.). The main purpose of this version of the Agent-Based Model was to demonstrate the traffic dynamics during the morning rush hour by simulating agents commuting from home to work. By doing so, traffic delays can be captured along the road network, which reflects the traffic dynamic. Before showing the overall traffic dynamics, a series of experiments were done to verify the model behaved as expected. The small full scale experiments proved the model did capture the traffic delays.

Additional experiments were run to validate agent movement on the road network. The experiments also demonstrated the model could represent a number of individuals with a single agent, instead of using a single agent to represent one individual. In these experiments, one selected road segment is used to simulate the traffic delay with various sizes of population (e.g., size of 1, 10, 100, 1000). The more people traveling on the selected road segment, the worse traffic delay can be captured by the model. In Figure 24, we show an example of a 2.1 km road segment in Manhattan and the speeds of various number of vehicles. As can be seem, increasing the number of vehicles decreases the average speed and increases the time of the trip. In the these scaling experiments, only one agent is loaded to the model to simulate 100 and 1000 agents commuting on the same road segment. The model generates the same commute speed as the 100 and 1000 agents' speed from the non-scaling experiment as shown by Figure 24. The experimental results also indicated that the model was capable to scale up properly. As a result, a percentage of population could be loaded to the model and simulate the traffic behavior of whole population. This allowed us to be comfortable that model runs with less than 1:1 population could produce behavior characteristics representative of the full population.

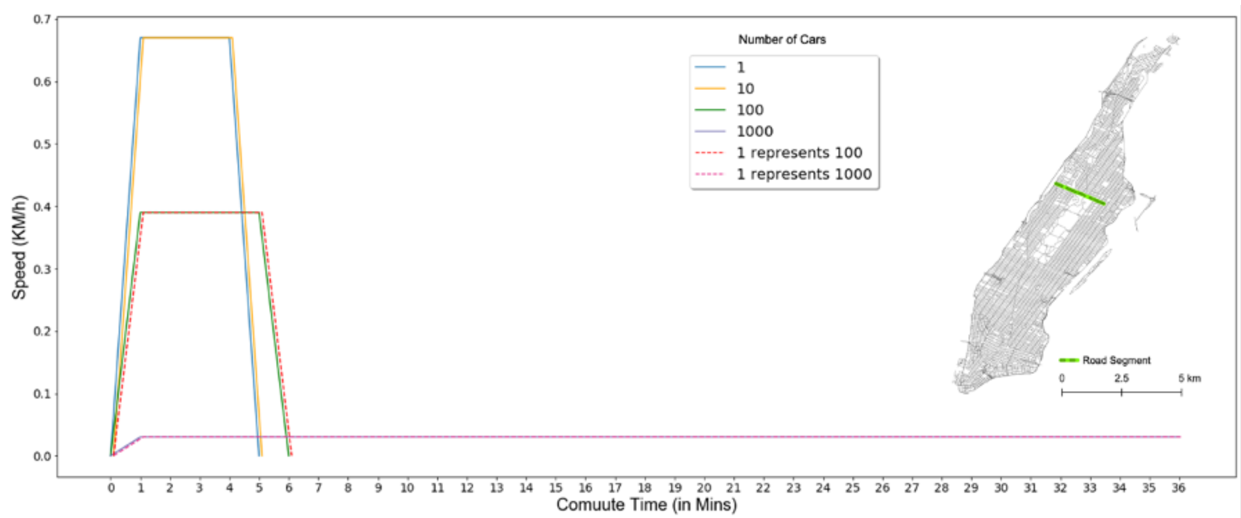


Figure 24: Selected Road Segment Traffic Delay Showing Traffic Delays Depending on the Number of Agents.

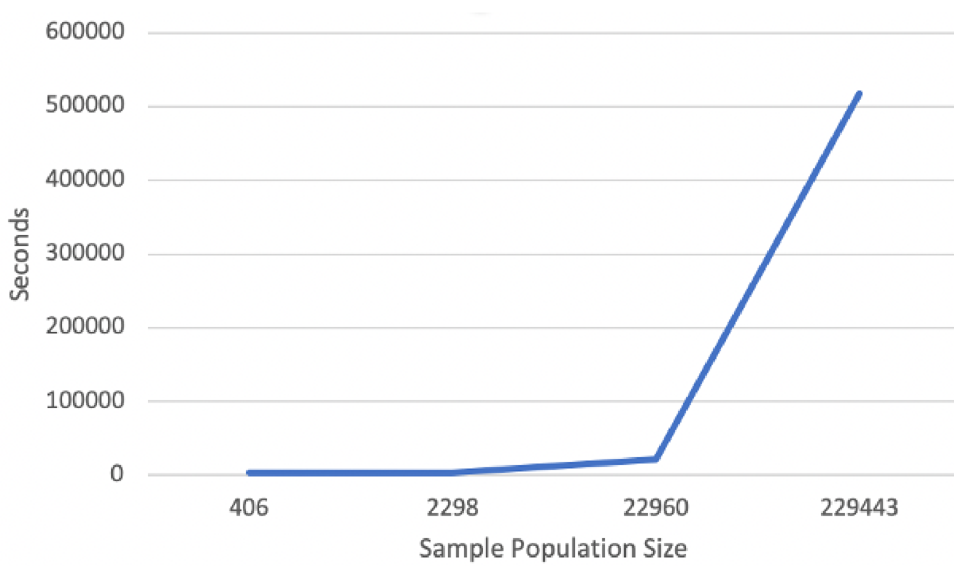


Figure 25: Model Performance as Computation Time to Step 615.

Agent-Based models can be expensive to run at full scale. The project set out to model the population of the commuting area of New York on a 1:1 scale. That means millions of agents. Computational time for runs of our model with different numbers of agents simulating 10 hours of normal behavior and 15 minutes after a nuclear MWD event on a current laptop took over 135 hours with only one tenth of the population represented. (See Figure 25.) (The laptop was a Macbook Pro with a 3.1 GHz Quad-Core, Intel Core i7 processor.) This is an obvious problem and will be addressed in Section 5.2 as an area for further work.

An approach to overcome the computational challenge of a running the model with a large number of agents was to explore behavior with a lower number of agents with adjustments to the movement rates so as to be at the same movement rate as they would be with the full population. We developed these factors experimentally by simulating agent movements and then calibrating them based on the fraction of the population instantiated in the model. Although originally done to verify the agent movement rates, the derived information can be used to have a computationally useful smaller population of agents moving as they would with the full scale population instantiated. See Section 4.2 and Table 4. Commuter traffic with different population sizes confirmation that smaller populations provide reasonable insights into the behavior of the entire population.

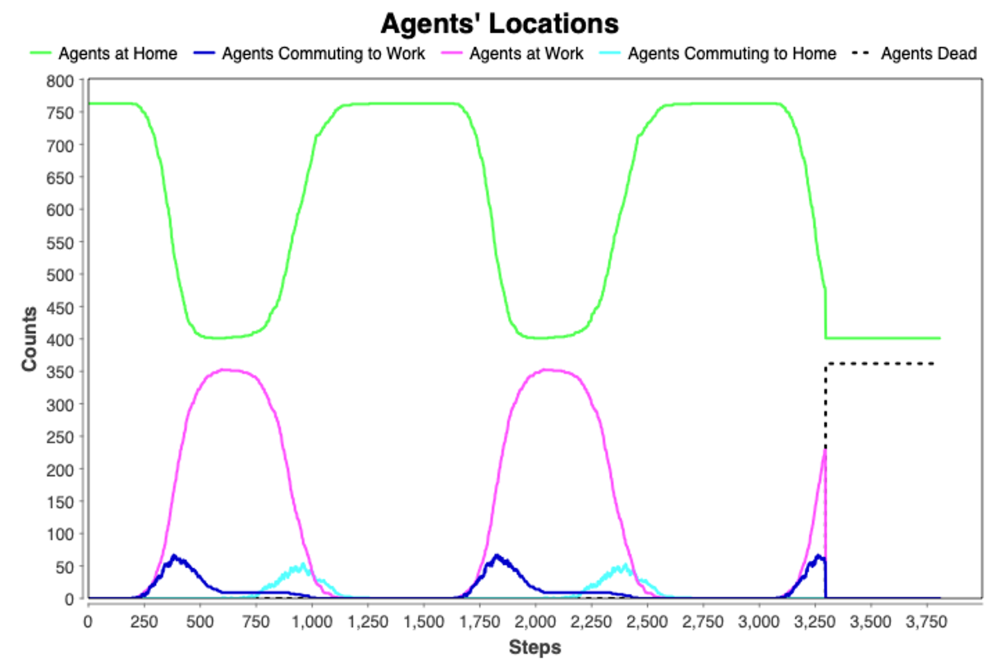


Figure 26: Commuting Behavior, i.e., Counts of Routine Behavior each Step (minute) Two Days Prior to NWMD Event.

4.3 Routine Traffic Patterns

As part of validating the model's representation of the population's behavior, we explored what routine travel patterns developed before the nuclear WMD event. Agents that commuted to their place of work were counted each minute of the simulation run and the counts plotted. Figure 26 shows counts of agent behavior activities two days prior to the NWMD event. Shown are most agents staying at home and just less than half spending time work during each day. Also shown are the number recognized as commuting in the morning and afternoon. The steps shown are counting the minutes since the simulation's start at the midnight of the first day. Each day consists of 1440 steps, i.e., minutes. This is a single test run with approximately 780 agents. This plot is from a simulation run set up to model a single work-shift per day. We also modeled and successfully tested three-shift schedules with overlapping commuting times. These results confirmed our we could model represented the behavior of the population of NYC reasonably well.

The verification experiments above also confirmed the model captures the traffic delay and scale up properly. To demonstrate the traffic dynamics of a high-density urban area with less computational efforts, a sample population was extracted for this experiment, which contains 2,306 individuals who worked in Manhattan and either lived in Manhattan or nearby area. This sample population was used in the model to capture each individual's geo-location at each time step. Figure 27 shows a heat map of traffic density during the morning commute, the darkest red color indicates the area is suffering the worst traffic delay. Again, our synthetic population dataset demonstrated its validity in simulating traffic dynamics. With confidence in the model's demonstrated behavior during routine operations, we added the nuclear WMD event and its effects for the population to react to.

4.4 Reaction to NWMD Event

Our previous work on agent commuting under normal conditions provided a basis for experiments with the NWMD event. We chose to have the NWMD event to occur at 10am during a work day. This timing as a parameter easily adjustable for further experiments. Agents are affected in the event based on the modeled weapon effects, as described in Section 3.3.4. Agents representing injured people attempt to flee the weapon effects zone. Simple, "fast and frugal" decision trees were used to model behaviors based on their situation [342].

Typically, our experiments had the population commute to work normally and the weapon modeled as exploding in the morning. The within the first minute, the agents working within the first radius are simulated as killed. Between the first radius and the second, agents are mortally injured and, as naive victims, they start to flee the area. Their their movement rate was modeled based on their health status simulating travel on foot through rubble. Victims between the

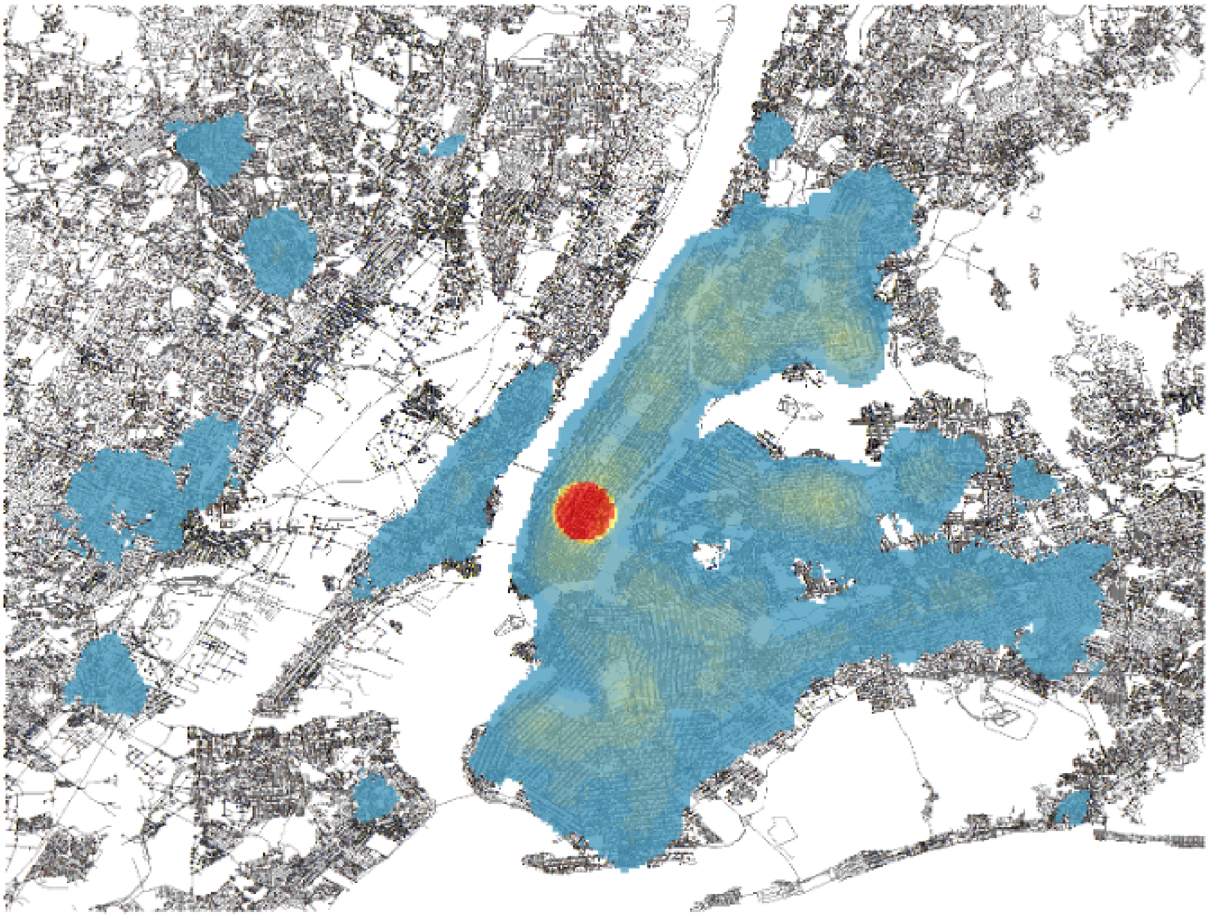


Figure 27: Worst Traffic Delays during Morning Commute.

second and third radii exit their buildings and flee on foot toward transportation nodes. In addition, first responders, typically set to a random 10% of the population, start moving toward the source of the explosion and provide assistance to those victims they find. We expect these behavior are representative of the population's immediate reaction to the weapon. We also believe these behaviors would continue for hours after the event before organized responses would come into play. A different behavior we did include in our model was the formation of groups of victims after the event. This behavior is discussed in the next section.

In what follows we present and discuss a set of simulations and results of the ABM to test the extent the agent-based model captures the behavior of individuals and groups in response to the detonation of a 10 kt NWMD in New York City (Section 3.3.8). The model was run with different population sample sizes, and data was collected on the number and range of emergent groups as well as their social networks' nodes and edges. We explored these groups to see how they scale with population size and demonstrate how such a model can characterize their dynamics with basic statistics of the model population, emergent groups, and social network analysis measures. The experiments were restricted to emergent networks because agents in the current ABM only make decisions post-detonation based on their health and location. Information regarding household, school, and work social networks had not been included in their decision trees. Parameters in the model were set so that agents form carpools under routine conditions and emergent groups after the NWMD event. During each simulation run, agents traveled from home to work with some commuting in carpool groups. At 1000 a.m. or Step 600, the NWMD detonates south of Central Park. The simulation is then run for another 15 minutes (to Step 615) because emergency management planning by the US Government indicate that the initial minutes of an emergency are critical for their response. As discussed in Section 3.3.6, agents whose health was impacted by the bomb respond by moving away from the point of impact. As individual agents encounter other agents in the same location and with the same goals (to join a group, find shelter, shelter, or flee), they will form emergent groups.

As discussed in Section 3.3.6, results from verification of the model demonstrated that routine patterns of behavior provide a baseline before agent behavior respond to the NWMD detonation. We next show runs where agents responded to a 10kt NWMD detonation just south of Central Park. Figure 28 shows the health status of agents in and around the impact zone. In this figure, affected agents were initially clustered around the detonation site at Step 600 (10am) and color-coded based on their health status, ranging from healthy (green), sick (orange), lethal (red), and dead (black). The greater the distance agents are from ground zero the greater their chance of survival.

Over the duration of the simulation, agents that are ambulatory, move away from the impacted area at varying speeds depending on their health status. The speed differentiation creates a visual pattern of dispersion of the injured agents in 28 due to the less injured victims fleeing faster than more the severely injured victims. (Not identified in Figure 28 are agents designated as first responders who initially move toward the affected area.) The impact area is relatively small compared with the entire commuter region. As discussed in Section 3.3.6 the larger area is used simulate healthy first responder that can move towards ground zero. As noted in Section 3.3.6, a NWMD event causes agents to alter their normal behavior and daily routines (e.g., commuting behavior and locations) by moving away from the impact area in varying states of shock and levels of injury [4, 303]. One of the key patterns in a disaster is the movement of people inbound and outbound from the impact area. Besides the movements of agents, another behavior we modeled was the formation of social groups of co-located survivors. This behavior is discussed in the next section.

4.5 Groups Formation After NWMD Event

New social groups (emergent groups) are one element of the social capital leveraged as part of the community responses to disaster [239]. These groups can provide a critical function in disaster by delivering people, resources, and actions to address unexpected needs in the response (See Sections 2.1.3 and 2.1.2). However, this form of community resilience is unpredictable because emergent groups are not planned. The individuals in that make up these groups, their resources, their goals, and their actions are unknown before the event. As a result, empirical data on their dynamics is difficult to collect as they form, and in some cases impossible to collect until well after the fact. An agent-based model of the CAS in disaster can reveal emergent patterns in these complex systems [236] (See Section 2.1.2). Combined with social network analysis, they can be used to test and reveal patterns of emergent groups that arise from human interactions in complex adaptive systems.

In the final agent-based model, when the model was run past the NWMD detonation, groups began forming. Table 6 shows the number of emergent groups and the range of the group sizes as of Step 615 (15 minutes after the event). As the population size in the model increases, so does the number of emergent groups. The range of emergent group sizes increases with the population sizes, but the ranges do not vary across simulation runs. Emergent groups never form for the population size of 406. Just as the scaling analysis in Table 4, emergent groups are not significant for analysis.

Simulation data in Table 6 shows consistency across 10 runs with only two variations, Runs 8 and 10. These variations are a result of the stochasticity introduced in the ABM through variations of the number and locations of road network nodes and edges that are destroyed by the nuclear bomb, as discussed in Section 3.3.4. In Zone 2 (the middle zone) of the impact area 50% of the nodes and edges of the road network are destroyed. There are no functioning road network nodes and edges in the center Zone 1, and the road network nodes and edges remain in place. When the destroyed nodes and edges in Zone 2 differ across simulation runs, the range of emergent group sizes can be seen to decrease as shown in Run 8 of population size 22,960, and the number of group sizes in Run 10 of population size 229,443 decreases from 172 to 165. Due to the variations of the number of emergent groups between Run 9 and Run 10 of population size 229,443, a detailed exploration of the groups sizes of each run is graphed in Figure 12. Although the range of sizes remains 2-100, the number of groups with the same group size differs between runs. These variations may be due to both the stochasticity introduced in the road network damage along with the order of activation of the agents. The MASON toolkit uses random number generator for agent step schedules placement [292, 293].

However, across the four populations, increasing numbers and variety of groups formed. For a population of 406, no groups formed and for a population of 2K, 2 groups were formed consistently. There was more variety for the next to populations, 22K and 229K, as shown in Figures 30 and 31. The distribution of group sizes suggests a relationship between the number of groups and group sizes for emergent groups. The distribution could be described by an exponential curve with a heavy tail, but due to the small sample size of number of populations, the scaling relationship cannot be confirmed.

Not only can the number and sizes of emergent groups be explored in our ABM, the social network data for these groups can be exported at any time step. Similar to the scale of emergent group numbers and sizes, the variety of emergent group networks scale over the three populations sizes on the last time step. The network diagram of emergent groups in simulation with a population size of 22,960 are shown in Figure 32. The largest emergent group can be seen in the bottom left corner of the diagram. Unfortunately, when we diagram the networks of emergent groups from a run

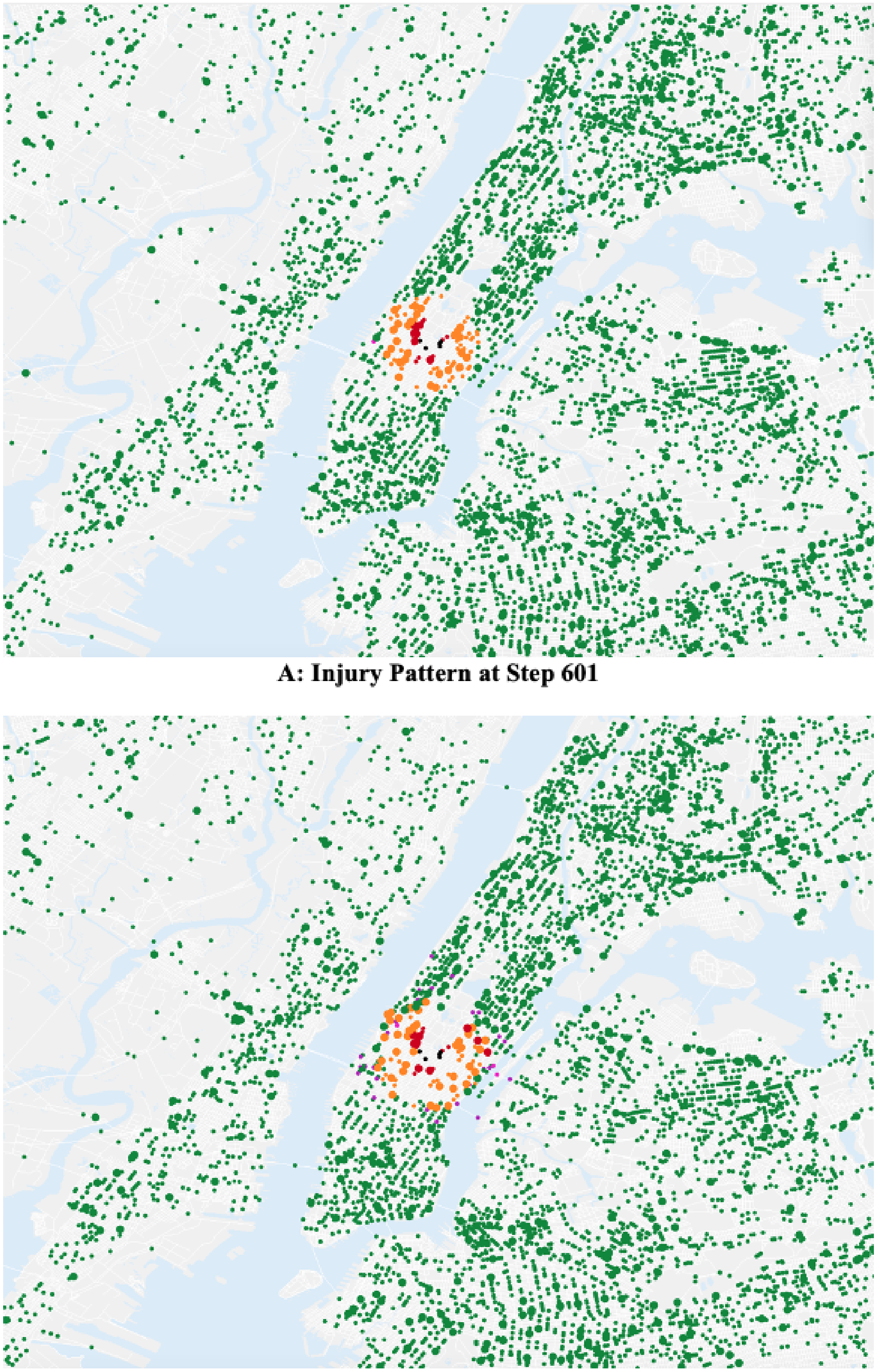


Figure 28: Injury Pattern with Population Size 22,980 at 1 Minute and an Hour after the Explosion.

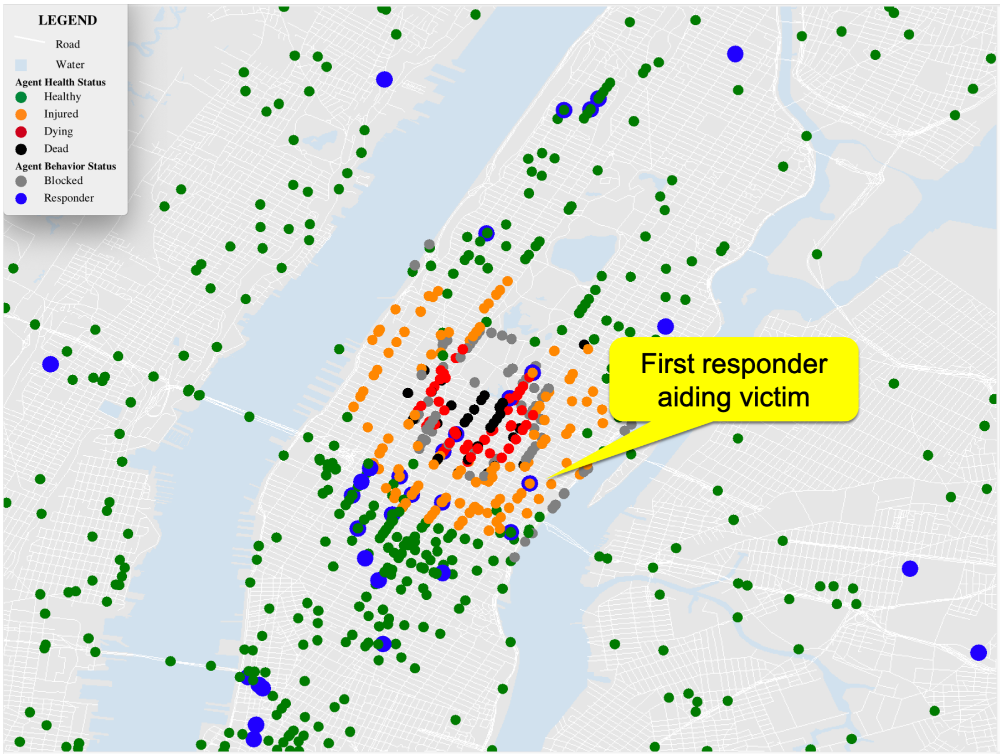


Figure 29: Model After 15 Minutes.

Table 6: Emergent Groups Post NWMD Detonation.

	Population Size:	**406**	**2,298**	**22,960**	**229,443**
Run 1	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 2	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 3	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 4	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 5	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 6	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 7	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 8	No. of Groups Size Range	0 -	2 2-2	43 2-36	172 2-100
Run 9	No. of Groups Size Range	0 -	2 2-2	43 2-37	172 2-100
Run 10	No. of Groups Size Range	0 -	2 2-2	43 2-37	165 2-100

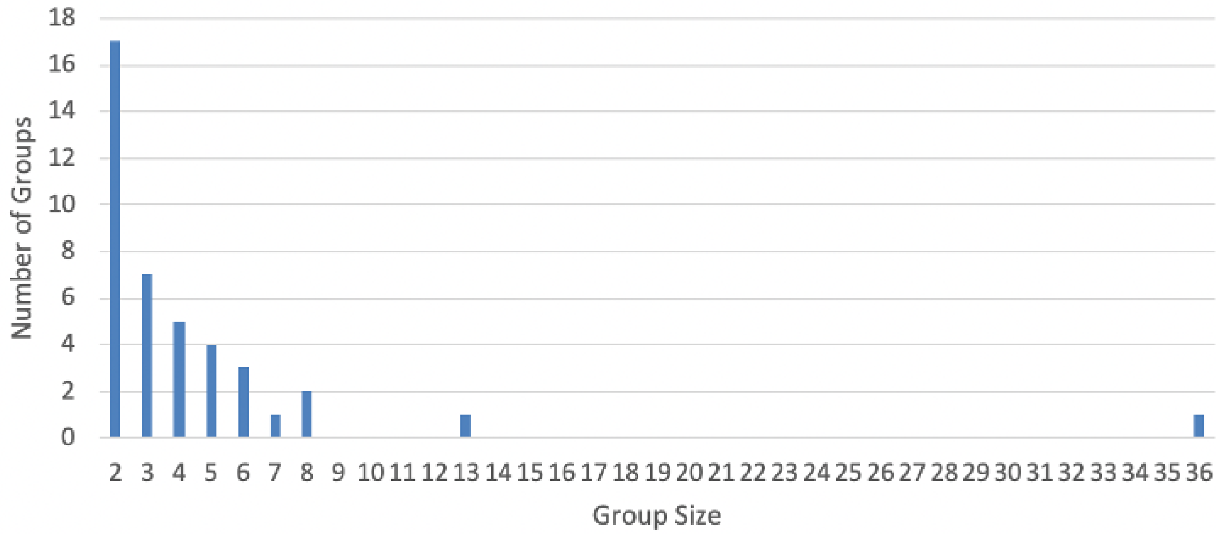


Figure 30: Population 22,960: Distribution of Number of Groups at Each Size at Step 602.

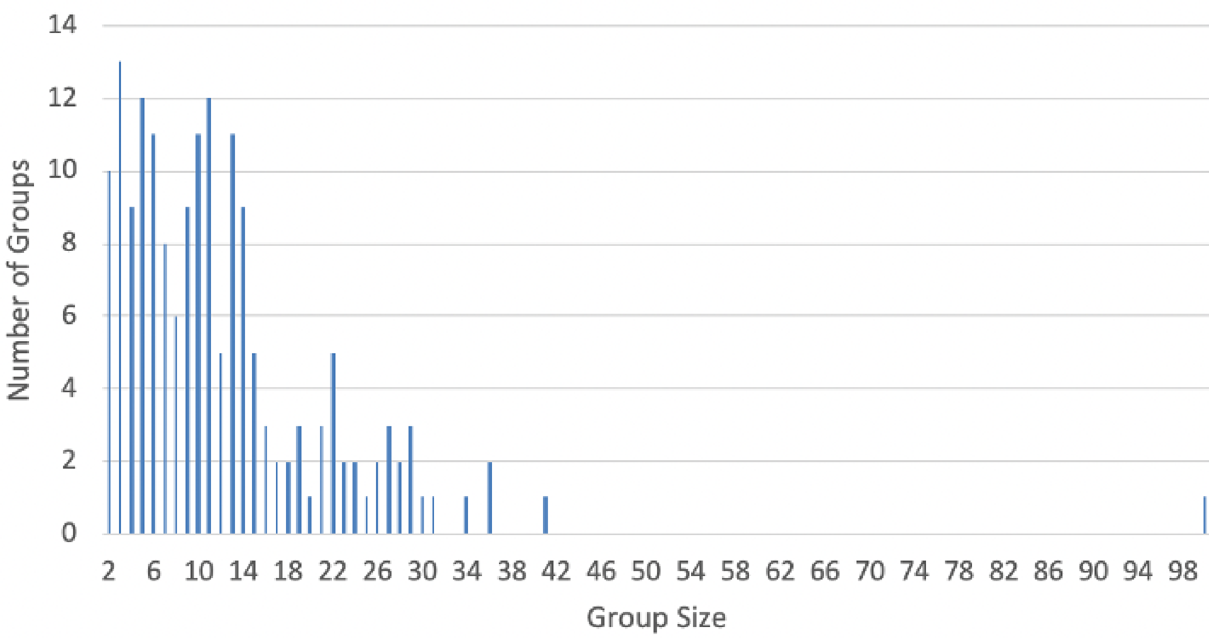


Figure 31: Population 229,443: Distribution of Number of Groups at Each Size at Step 602.

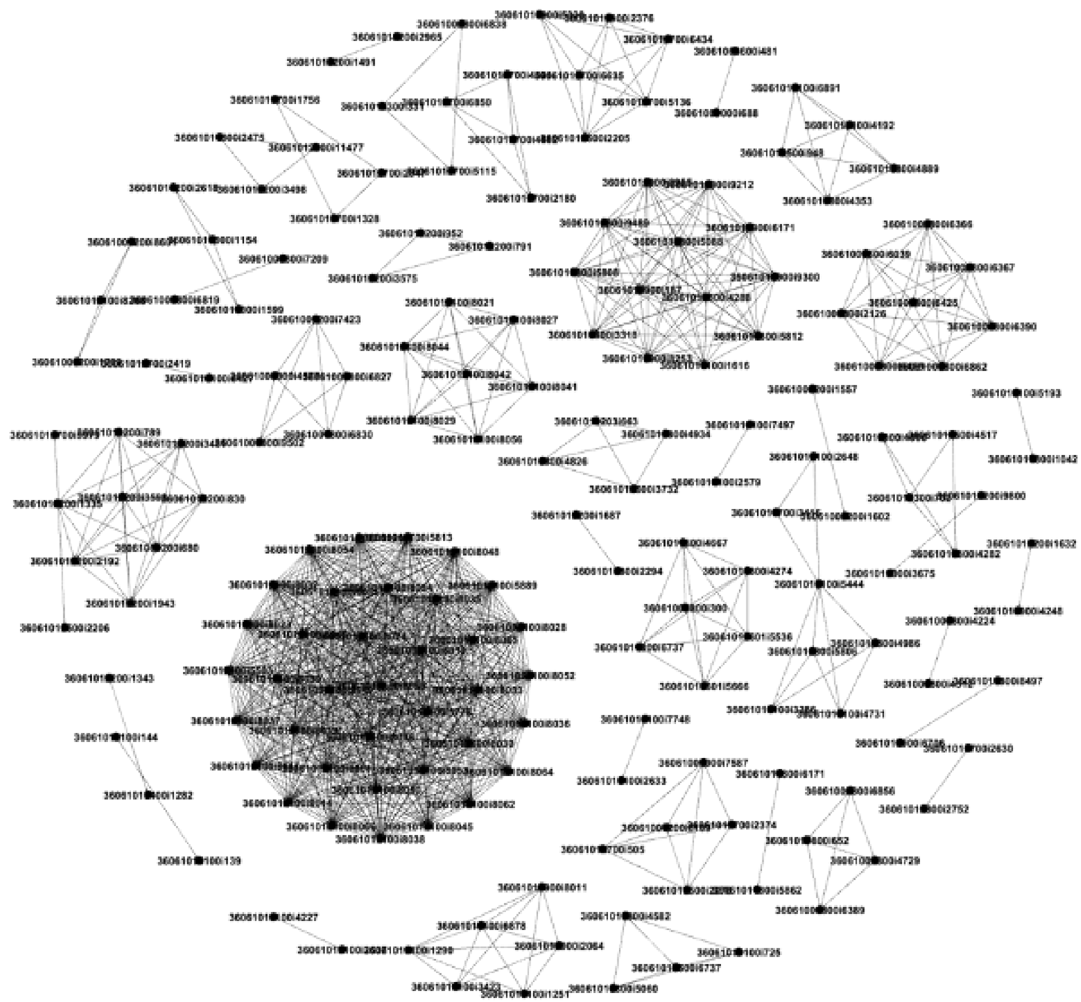


Figure 32: Emergent Groups for Population Size 22,960.

with a population of 229,443, the network diagram of 172 emergent groups produces a diagram too dense to easily visualize each group. Obviously, there is interesting dynamics in the formation of groups after the nuclear WMD event.

4.6 Characterizing Reaction to NWMD Event

We characterized the reaction of the population of NYC to a nuclear WMD by describing the status and activities of our modeled population. As our model ran, we collected data on how many agents were different conditions. 33 shows the results. With a simulated population of approximately 23,000 agents, approximately a third, about 7,500 work outside of their homes. The plot shows them from midnight until an hour after the nuclear event. They commute to work and at 10am, i.e., 600 minutes after midnight, the nuclear WMD event goes off and immediately some are killed, some start fleeing, and some are first responders headed toward the blast. In this plot, those directly affected by the weapon are no longer considered "at work". Others, outside the affected area, start for home and are counted as commuting. Some of the commuters become blocked by rubble because their commute route would take them through the damaged area. Also note, in this scenario, there is no delay included to model the people's shock, disbelief, and delay due to being in damaged buildings.

Figure 33 is the result of a single run of the simulation. As an Agent-Based Model, this is one of several possible results. To explore the variety of results, we made 25 runs with the same parameters. The plot of those 25 runs is shown in 34. There is little variation among the 25 runs prior to the explosion because the agents live and work at fixed locations and

Figure 33: Characterizing Reaction to Nuclear WMD Event (a single run).

Figure 34: Characterizing Reaction to Nuclear WMD Event (25 runs).

follow the same patterns of movement. After the event, there is stocastisicty due to a different set of first responders being selected randomly and random assignment of health status to the victims affected by the nuclear WMD.

5 Summary and Areas of Further Work

While nuclear weapons of mass destruction exist, thankfully they have only been used in anger twice. Therefore, there is little known about how people react to them. To overcome this dearth of understanding, we treated a nuclear WMD event as a complex adaptive system (Section 2). We focused on the main properties of complex adaptive systems, specifically heterogeneity, webs of connections (i.e., social networks), relationships and interactions, and adaptations arising from individual actions and decisions (which relate back to our four core ideas first introduced in Section 1). By building on psychological theory and previous disaster studies, we developed an agent-based model that captures how the characteristics of the environment, the agents, and the event affect their behavior, i.e., their reaction to the simulated NWMD. The model represents the road network and weapon effects as part of the physical environment. It also includes the individual and social environments through agents' social networks and carpools prior to the NWMD event and emergent group dynamics after the event.

The work presented in this report not only demonstrates the extent an agent-based model can characterize individuals and emergent groups respond to a NWMD detonation in NYC, but also furthers disaster research with a new conceptualization and implementation of disaster as a complex adaptive system of systems 1 and extends the use of computational social science methodologies in disaster research. It expands the spaces of scientific inquiry in disaster research from traditionally qualitative methodologies to computational social science's quantitative methodologies. The conceptualization can be used as a foundation for ensuring agent-based models adequately account for the interacting systems and variables in disasters. As discussed in Sections 2.1.2 and 2.1.3 these are required to improve understanding of the nonlinear dynamics, the crossing of scales, and the feedback of forces already observed and studied in disaster research. The Complex Adaptive System framework enables modelers to adequately simulate the dynamics of the complex adaptive systems that evolve into disasters such as features of heterogeneity, webs of connections and relationships, and adaptations with the system, and dynamics such as emergence, flows of information and resources, and shifts of the system between stability and instability.

The combination of agent-based modeling technology with social network analysis and GIS for disaster study is a novel methodology in computational social science that enables experimentation with spatial characteristics in the physical system as well as the agent interactions that make up the social and individual systems. Experimentation with social networks in an agent-based model adds to the knowledge of how social networks evolve in bounded space and time, and the agent-based model demonstrates how social networks can be operationalized and analyzed within a spatially explicit model. This work integrates disaster research from empirical and theoretical literature within an agent-based model. By developing a conceptualization of complex adaptive system of systems in disaster (Section 2.2, agent-based models of disaster can be designed to better characterize the effects of physical (Section 2.1.1 (Section 2.1.2, and Section 3.3.5) systems and better represent the balance of forces within these systems. Integration of agent-based modeling with social network analysis and GIS demonstrates how the combination of these computational techniques can illuminate agent interactions with network diagrams and create emergent networks that can be studied for network characteristics and attributes.

In this report, we demonstrate a mixed method that generates a synthetic population from available census data and generates their synthetic social networks in the area of Connecticut and New Jersey, New York and part of Pennsylvania (Section 3.2. In addition, a series of verification and validation of the population synthesis confirmed its robustness (Section 3.2.3). Meanwhile, the synthesized network of the population represents multiple network relationships with individual agents connected both to household members and individuals located at education sites and workplaces. To the end, two use cases indicate the utility and re-usability of the synthetic population dataset and its synthetic social network.

The mixed method addresses some of the current challenges in population synthesis for agent-based models. By including synthesized social ties for agents, the model can better represent human relationships and the behavior that arises from those relationships, such as movement of people to visit/migrate, daily commuting patterns, and information dissemination and decision-making in the context of disasters or purchasing decisions. These synthesized networks no longer restrict agent-based models to simulations of interactions based only on physical proximity connections (i.e., adjacent cells), rather they allow distant and multi-layer network connections and interactions to better represent and impact agent behavior. Our method also created a heterogeneous population that replicates aggregate statistical descriptions from empirical data, and yet maintains the anonymity of the actual population's personal information. Techniques for anonymizing data are critical to the utilization of big datasets in simulation as the agent-based modeling community builds models at higher resolution and closer to real-world conditions.

5.1 What We Learned

This section discusses what we learned from this project. We address, modeling infrastructure, modeling agents based on the census data, and agent behaviors.

About Modeling Infrastructure Modeling the infrastructure needs to be done within the vision for the whole model and with an eye toward what minimum functionality is required for the purpose of the model and with an understanding of potential refinements and extensions may be supported. In this model, we represented homes, schools, and workplaces as points on the road network and agents could move along the that network to get move between particular locations. That was enough to place agents at the time of the event (e.g., generating basic patterns of life). The creation of these locations was part of the population synthesis portion of the project (Section 3.2. Movement between locations by individual agents involved route planning for which we used the A* algorithm [295]. Moving individual agents was relatively easy but required calibration and validation (see Section 4.2). However, when it came to moving agents in groups, this was a significant part of a doctoral dissertation [15]. This required addressing the formation of groups, selecting a leader, i.e., an exemplar for modeling purposes, managing members joining and leaving the group, and dealing with contingencies that a single agent did not have to deal with. One contingency particular to this model was the potential death of members or the leader of the group. See Section 5.2.

About Synthesizing A Population An important part of the design of this model was to have our agents represent that actual population of New York mega-city and surrounding commuting region. We learned we could use the 2010 census and related sources (see 3.1) to establish where each person was at any given time during the day. This started with our use of heuristics for forming households based on the demographics at the census tract level served our purposes very well. We successfully created agents for basically for each person in the census (99.64%) matching the available demographics data, such as sex and age of household members based on the counts and types of households in a given census tract. With households, other heuristics could match adults with work places based on the home work commuting data. We use actual school locations to place the schools and heuristics to match children of households with their appropriate local school or daycare center based on their ages. We also created departure times for those going to work, school, or daycare centers. This supported the model having agents in appropriate places when the nuclear WMD event occurred. We also created each agent's social network based on appropriate parameters linking them to other agents in the work, school, and daycare places. The social networks for agents is a particular strength of our synthesized population because few large-scale agent-based models incorporate social networks influencing agent movement [2].

About Agent-Based Modeling We learned several things about agent-based modeling, among which, dealing with the very large numbers of agents is first. Although running the model with the full population would take days to run, we found there did not seem to be a material difference in the behavior we observed as we ran the model with larger and larger fractions of the synthesized population. Running the model with less than the full population, however is not simply setting a specific fraction of the population to be instantiated. It did require careful attention to which agents were included, particularly entire households, rather than fractional households seemed important to to select, but fractional work places, schools, and daycare centers did not cause problems, although social networks were impacted.

To meet the goal of characterizing the reaction of a population of a mega-city to a nuclear WMD event, we ran our model with several different fractions of the entire population to explore resulting behaviors (Section 4). With the number of agents approximately 20,000, the systems response was still interactive, meaning results could be visualized in near real time, while for larger populations this was not the case.

We also developed an approach to have the behavior reasonably replicate the dynamics of the whole population. This approach was to calibrate the movement speed of the agents based on the fraction of the population implemented. That way a smaller number of agents would move more slowly from point to point simulating traffic (as discussed in Section 4.2). Possible future approaches are discussed in Section 5.2.We confirmed the importance of verification and validation. It certainly took a lot of effort but the verification of the quality of the synthetic population was worth the effort. We learned the magnitude and specifics of where our approach did not match the available data by verifying our approach at the census tract level. We synthesized our population for 99.64% of the population and know which tracts have problems with the source census data. We also know how well our synthesized population compares in the area of demographics to the actual census and another effort to synthesize a population (see 3.2.3. On the subject of validating a model of a population's reaction to a nuclear WMD, we learned other types of disasters can provide useful data for modeling purposes and the only uses of nuclear weapons can provide a useful validation source (Section 4.6).

[2]Note: census data is for a specific day, as of 1 April of the census year. As an interesting point, the 2010 census reported a small number of people living in Central Park.

About Disasters The scientific study of disasters is only 100 years old this year, given the first scientific work in the area was Prince's [71] PhD thesis at Yale on the Halifax Disaster of 1917 discussed in Section 2.1.2. With this project, we learned three things about disasters. First, the research on disaster is becoming more organized. So, we wrote a review and proposed an organization [16] and that lead to a dissertation on the topic [15]. Second, there are two major categories of disasters based on their notification, developing conditions or no-notice events. The focus here was intentionally on no-notice events, where people become alerted with the NWMD event or afterwards. That way we could model their reactive behavior (Section 3.3.6), expected to be more constrained than their their anticipatory behavior. Third, as discussed in Section 2, disasters can be thought of as complex adaptive systems with physical, individual, and social major components. This successfully guided the methodology used in this research (See Section 3.) Fourth, we learned from first hand accounts that, in general, will go through a series of phases starting with shock and denial, but that people were well behaved during the early part of a disaster, they helped each other, they followed instructions, and they were generally calm. We also learned that non-notice events are very different from slowly developing disasters when fear and anticipation may be the primary drivers of behavior. As such, we would expect that modeling such a disaster would be much more difficult that the task we set for ourselves in this project. Finally, there is obviously much more than can be done to study a populations reaction to a disaster. The next section identifies some of these.

5.2 Areas for Further Work

Disasters are instances of community stress uniquely located at the intersection of human, natural and technological systems that continue to change and adapt to new circumstances. The complexity inherent in the number of interacting variables in space and time can be prohibitive of research and understanding. Data for verification and validation of this agent-based model does not exist, and therefore the quantitative data gathered in the model cannot be used to prove specific behavior in the case of a NWMD detonation in NYC. Full validation for this work would require a NWMD event for observation and the means with which to measure the behavior of individuals and emergent groups throughout the event. However, one of the utilities of agent-based modeling is to test ideas and hypotheses that cannot be done in the real world for theoretical, practical, and ethical reasons. A model that is constructed on well know principles and theory should be able to generate patterns and behaviors that one would observe in the real world [343, 344]. For example, our model behavior builds on years of disaster research and theories (Section 2).

The challenges of research in the area of complex adaptive systems and disasters are numerous. Research on complex subject matter requires the management of heterogeneity, multiple dimensions (time, space, scale, etc.), many interdependencies, interacting scales (local to global), unpredictable outcomes, and the need for interdisciplinary expertise. More critical to progress in this area is the unobservability of events and indicators. The features of the physical, social, and individual systems often do not survive a disaster event and reported events and experiences are reshaped by memory and circumstance. Additionally, the environment in disasters is often too dangerous for observation and measurement or too unpredictable for researchers to plan any field work. However, one potential area for this is synthesis data, models and theories through the lens of computational social science as shown in figure 35 .

Addressing Scaling A challenge for agent-based models that have a large number of agents is the performance of the model [281, 345]. Specifically, the challenge is not the number of agents themselves but the nature of the model. In the sense that our agents are geographically explicit, and have social networks which are constantly interacting. While we could have minimized the geographical complexity along with the social networks, this would of detracted from the core model, where agents are and who they are connected to has an impact on the resulting behavior. With our time step at 1 minute and, at full scale (i.e., 23 million agents), we could have waited for months for a full parameter sweep. To reasonably deal with this challenge, we employed and explored a number of approaches. First and most frequently used was to test the model using a small fraction of the population. We developed synthetic populations of different sizes for research purposes. These smaller populations allowed verification of the modeling and exploration of the behavior of the population. We also explored scaling the agent movement rates based on the number of people simulated. See Section 4.2 for a discussion of the scaling of traffic dynamics. One potential way to overcome this is to use some sort of distributed computing infrastructure (e.g., [346]) or high performance computing (e.g., [347]), however even with these approaches, there will still be a major challenge when it comes to network communication.

Other Forms of Transportation Our model included the population's commuting, but abstractly based on the road network. By using the road network, our agents were able to travel from home to their routine daytime locations (Section 4.2). A significant and unexpected method of exiting Manhattan after the September 11, 2001 attacks was by an unorganized, volunteered boat lift that was reported to have moved more than 500,000 people off the island during that day [348]. Future models should include other forms of public transportation and *ad hoc* means of travel besides "on foot", i.e., including trains, subways, busses, taxis, and others. This would require more data which is

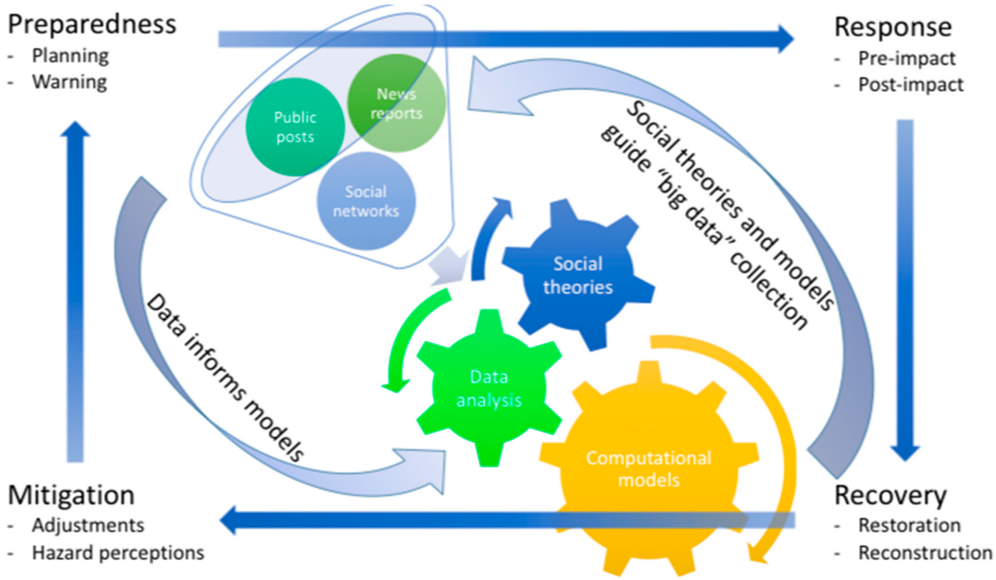


Figure 35: Interactions of Data Analysis, Computational Models, and Social Theory in Computational Social Science of Disasters (source: [14]).

becoming more available through open data initiatives (such as `https://opendata.cityofnewyork.us/`). It would also require more detailed models of modes of transportation which could be done through coupling various types of models together extending more traditional transport interaction/integration models (LUTI) models [341, 281, 349].

Modeling Types of Work This model is focused on identifying where people are at the time the NWMD detonates. To achieve that we only needed to model where they spend most of their time, i.e., at home, work, school, or daycare. Our model represented people either at home, at their daytime location, or commuting between. That level of granularity could be improved for different purpose by modeling the nature of social interactions for different types of primary daytime activities more specifically and including secondary activities such as social dining, recreational activities, and mass gatherings. This would be useful in modeling how different activities increase or decrease the likelihood of spreading a COVID-19-like disease. Some occupations, activities or recreational activities are more vulnerable to the spread than others.

The characteristics of activities that could be useful to include include amount of social contacts, jobs involved with frequent moving, e.g., transportation operators, jobs involved with interacting with disease carriers such as hospitals as places of employment and in-patients population. There are many features of our agents "lives" that would be useful to include for other research goals.

Adding more Agent Diversity Our models represented agent the population as having uniform initial health conditions, movement capabilities, level of civil defense and first aide training, and use of communication networks. These could be made more diverse. For example, future models could be expand the healthy and fit assumption to explicitly include the unhealthy portion of the population and those with mobility challenges for common reasons such as age, temporary illnesses, or permanent disabilities. In addition, some people are not mobile for other reasons such as incarceration or roles in vital public service positions. We had developed decision trees for both naive victims as well as trained victims (see Section 3.3.5). A future model could explore the effects of different levels of civil defense training and percentage of first responders within the populace on the casualties following a NWMD event. Furthermore, while possible, the model did not get to capture people's communications with their social network (e.g., cell phones, social media), which is important post-event need everyone has: to check on the condition of their loved one. The lack of ability to communicate or the learning of family members' distress would certainly affect agents' behavior after a NWMD, such as motivating them to neglect their own health in an attempt to assist family members. This was a particularly interesting avenue of further interest we would like to pursue.

Additional Applications: A Disease Modeling Example As hinted at above, the model code base and resulting data products (Appendix B) and the synthetic population be used to explore how people might react to a WMD but it can also be applied to other problems where space matters. For example, it can be applied in the epidemic research.

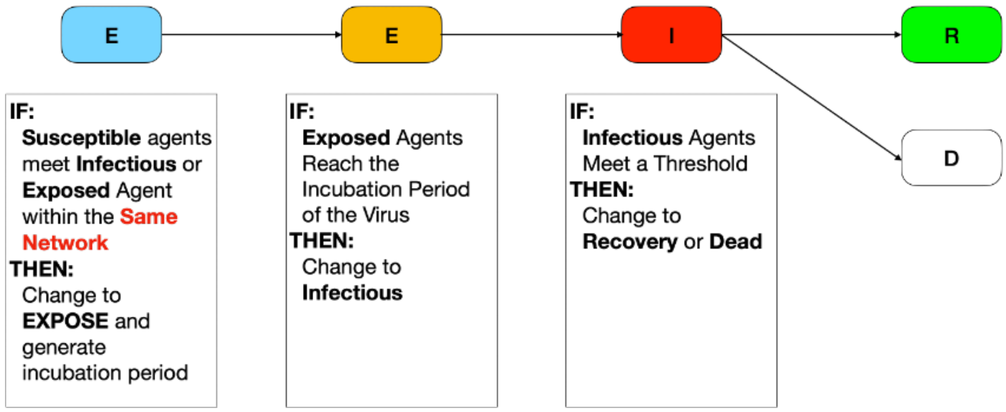


Figure 36: Disease Model Structure.

For example, here we preset a simple disease model built by using Python with the synthesized population dataset and its social network (Section 3.2) to model how a disease may spread through their social networks. This done by capturing the health status of each agent at each time step while the agents change the status. The purposed model demonstrates the usage of the synthetic population dataset and social networks, which indicates our synthetic population dataset has the reusability in other programming languages and other research purposes. As for the model, a standard compartmental model, the SEIR model (Susceptible - Exposed - Infectious - Recovered - Susceptible) is the foundation of this ABM. The SEIR model is shown by Figure 36. To reduce the intensive computational efforts, we extracted a sample population from one county in the state of New York, Ulster county, to run this model. Within this county, based on the 2010 census, there were a total 153,253 people living in 58,094 households. The spatial arrangement of homes, work places, schools, day care centers in the county is shown by Figure 37. In this model, one day is divided into 3 time periods, 8 hour each. In one time period, we make the agents have interactions through their work or education social networks, based on the agents being at work or education sites. These interactions are possibilities to spread the disease. During the other two periods, the model has all agents at home and interacting only through household networks. Figure 38 shows the SEIR curve, which indicates that the SEIR model can be replicated by using our synthetic population and their social networks. In addition, the relative locations where agents get infected can be inferred by the application of social networks. In Figure 39, the cumulative number of infectious agents are divided into home infected and work infected. As a result, the utility and reusability of our synthetic population in another programming language is proved through this use case.

This is an SEIR model of only one county, for demonstration purposes. It could for the basis of a model of just downtown NYC, the suburbs, or the entire region and it could support more sophisticated models of disease spread to address specific research questions beyond the original purpose for building the model of the road network and synthesized population.

Acknowledgements This if the final report of a 4-year long, basic research project to characterize the reaction of a mega-city to a nuclear weapon of mass destruction. This work was sponsored by the Defense Technology Research Agency (DTRA) under Grant number HDTRA1-16-0043. Dr. Paul Tandy was the Program Manager throughout the project and provided valuable guidance and advice thought-out the project which we are truly indebted for. The opinions, findings,and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the sponsors. The authors would also like to acknowledge other graduate research assistants who worked on this project including Diana "Pat" Guillen-Piazza, Mary Sproul, and volunteers: Samantha Vawter (GMU undergrad), 3 USAFA cadets, and 2 highschool students.

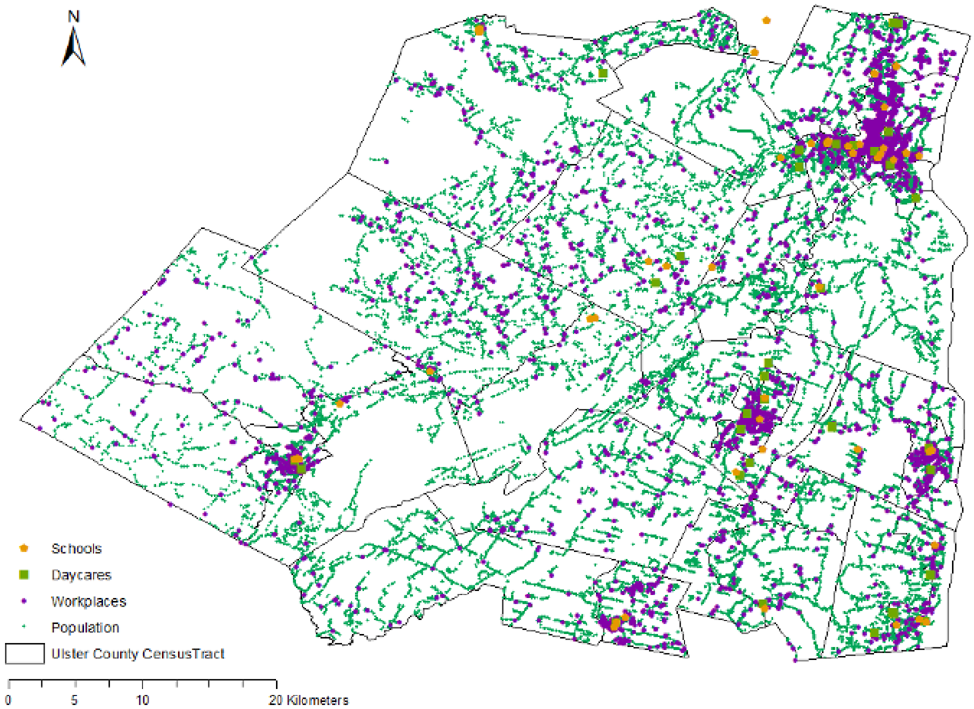


Figure 37: Locations in Ulster County.

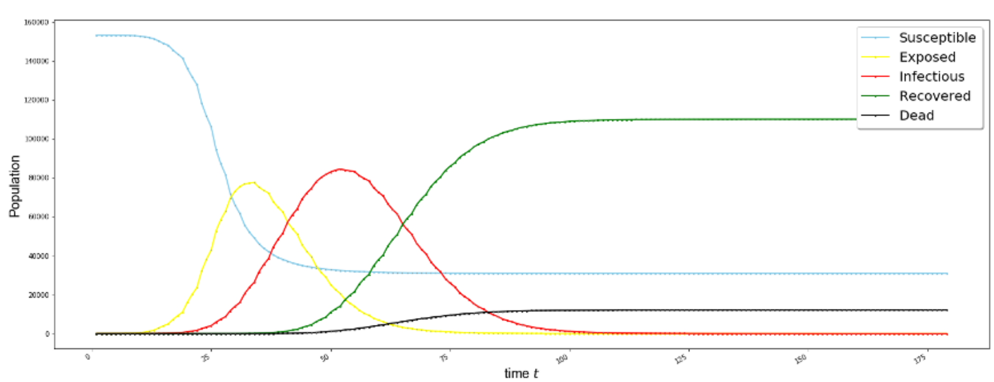


Figure 38: Disease Model Results.

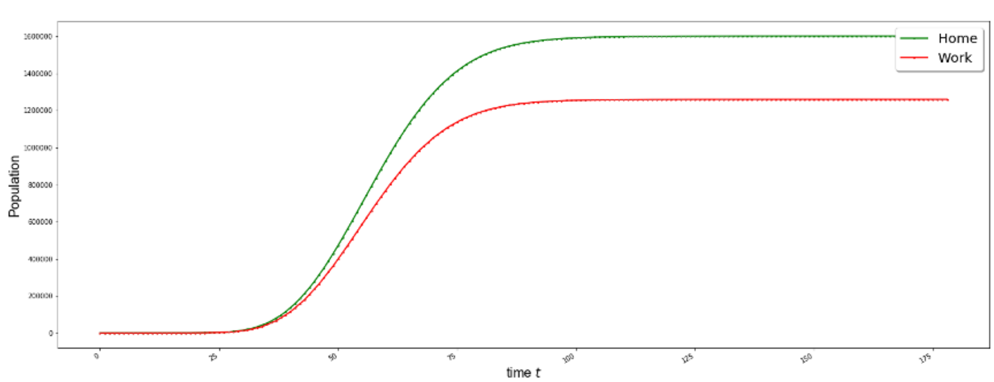


Figure 39: Locations of Where People Become Infected.

References

[1] George EP Box and Norman R Draper. *Empirical model-building and response surfaces.* John Wiley & Sons, 1987.

[2] Scott E Page. Computational models from a to z. *Complexity*, 5(1):35–41, 1999.

[3] US Department of Energy National Nuclear Security Administration. US Department of Energy National Nuclear Security Administration. https://www.energy.gov/nnsa/missions/maintaining-stockpile.

[4] Samuel Glasstone and Philip J. Dolan. The effects of nuclear weapons. Technical Report 972902, 1977. http://www.osti.gov/servlets/purl/972902-YMX5Q5/.

[5] Nidhi Parikh, Samarth Swarup, Paula E. Stretz, Caitlin M. Rivers, Bryan L. Lewis, Madhav V. Marathe, Stephen G. Eubank, Christopher L. Barrett, Kristian Lum, and Youngyun Chungbaek. Modeling human behavior in the aftermath of a hypothetical improvised nuclear detonation. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 949–956. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[6] Christopher Barrett, Keith Bisset, Shridhar Chandan, Jiangzhuo Chen, Youngyun Chungbaek, Stephen Eubank, Yaman Evrenosoğlu, Bryan Lewis, Kristian Lum, Achla Marathe, and others. Planning and response in the aftermath of a large crisis: An agent-based informatics framework. In *Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World*, pages 1515–1526. IEEE Press, 2013.

[7] William G Kennedy. Modelling human behaviour in agent-based models. In Alison J Heppenstall, Andrew T Crooks, Linda M See, and Michael Batty, editors, *Agent-Based Models of Geographical Systems*, pages 167–180. Springer, New York, NY, 2012.

[8] Herbert A. Simon. *The Sciences of the Artificial*. MIT Press, Cambridge, MA, 3rd edition, 1996.

[9] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. An integrated theory of the mind. *Psychological review*, 111(4):1036, 2004.

[10] Joshua M Epstein. *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science*. Princeton University Press, Princeton, NJ, 2014.

[11] Susan T Fiske. *Social Beings: Core Motives in Social Psychology*. John Wiley & Sons, 2018.

[12] Abraham Harold Maslow. A theory of human motivation. *Psychological Review*, 50:370–396, 1943.

[13] Andrew T Crooks and Sarah Wise. Gis and agent-based models for humanitarian assistance. *Computers, Environment and Urban Systems*, 41:100–111, 2013.

[14] Annetta Burger, Talha Oz, Andrew Crooks, and William G. Kennedy. Generation of realistic mega-city populations and social networks for agent-based modeling. In *Proceedings of the 2017 International Conference of The Computational Social Science Society of the Americas*, pages 1–7, New York, NY, USA, 2017. Association for Computing Machinery.

[15] Annetta Burger. Disaster through the lens of complex adaptive systems: Exploring emergent groups utilizing agent based modeling and social networks. PhD Dissertation, George Mason University, Fairfax, VA, 2020.

[16] Annetta Burger, Talha Oz, William G Kennedy, and Andrew T Crooks. Computational social science of disasters: Opportunities and challenges. *Future Internet*, 11(5):103, 2019.

[17] Na Jiang, Annetta Burger, Andrew T Crooks, and William G Kennedy. Integrating social networks into large-scale urban simulations for disaster responses. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation*, pages 52–55, 2020.

[18] Annetta G Burger, William G Kennedy, Andrew T Crooks, Na Jiang, and Dianna Guillen-Piazza. Modeling social networks in an agent-based model of a nuclear weapon of mass destruction event. In *The 2019 Computational Social Science Society of Americas Conference*, Santa Fe, NM, 2019.

[19] Annetta G Burger, William G Kennedy, Andrew T Crooks, Na Jiang, and Dianna Guillen-Piazza. Modeling society reacting to a nuclear weapon of mass destruction event. In *2019 International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, Washington DC, 2019.

[20] Talha Oz. Collective stress in the digital age. PhD Dissertation, George Mason University, Fairfax, VA, 2020.

[21] Allen H. Barton. *Communities in Disaster: A Sociological Analysis of Collective Stress Situations*. Doubleday Anchor, New York, NY, 1970.

[22] Charles E. Fritz. Disasters. In R.K. Merton and R.A. Nisbet, editors, *Contemporary Social Problems*, pages 651–694. Harcourt, New York, NY, 1961.

[23] Anthony Oliver-Smith. "What is a disaster?": Anthropological perspectives on a persistent question. In Anthony Oliver-Smith and Susanna Hoffman, editors, *The Angry Earth: Disaster in Anthropological Perspective*, pages 18–34. Routledge, New York, NY, 1999.

[24] Ronald W. Perry and E. L. Quarantelli, editors. *What is a Disaster? New Answers to Old Questions*. Xlibris, Philadelphia, PA, 2005.

[25] Enrico Louis Quarantelli, editor. *What is a Disaster? Perspectives on the Question*. Routledge, New York, NY, 1998.

[26] Havidán Rodríguez, Joseph E Trainor, and William Donner, editors. *Handbook of Disaster Research*. Handbooks of Sociology and Social Research. Springer, New York, NY, 2007.

[27] Dennis Mileti. *Disasters by Design: A Reassessment of Natural Hazards in the United States*. Joseph Henry Press, Washington, DC, 1999.

[28] Margareta Wahlstrom and Debarati Guha-Sapir. The human cost of weather-related disasters 1995–2015. Report, United Nations Office for Disaster Risk Reduction (UNDRR), Geneva, Switzerland, 2015.

[29] Leigh Tesfatsion. Agent-based computational economics: Modeling economies as complex adaptive systems. *Information Sciences*, 149(4):262–268, 2003.

[30] Simon Levin, Tasos Xepapadeas, Anne-Sophie Crépin, Jon Norberg, Aart De Zeeuw, Carl Folke, Terry Hughes, Kenneth Arrow, Scott Barrett, Gretchen Daily, Paul Ehrlich, Nils Kautsky, Karl-Göran Mäler, Steve Polasky, Max Troell, Jeffrey R. Vincent, and Brian Walker. Social-ecological systems as complex adaptive systems: Modeling and policy implications. *Environment and Development Economics*, 18(2):111–132, 2013.

[31] Timon McPhearson, Dagmar Haase, Nadja Kabisch, and Åsa Gren. Advancing understanding of the complex nature of urban systems. *Ecological Indicators*, 70:566–573, 2016.

[32] Lilian Alessa and Andrew Kliskey. The role of agent types in detecting and responding to environmental change. *Human Organization*, 71(1):1–10, 2012.

[33] Dagmar Haase. Participatory modelling of vulnerability and adaptive capacity in flood risk management. *Natural Hazards*, 67(1):77–97, 2013.

[34] Jonas Joerin, Rajib Shaw, Yukiko Takeuchi, and Ramasamy Krishnamurthy. Assessing community resilience to climate-related disasters in Chennai, India. *International Journal of Disaster Risk Reduction*, 1:44–54, 2012.

[35] Thomas C. Schelling. Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2):143–186, 1971.

[36] Dewald Van Niekerk, Emmanuel Raju, and Christo Coetzee. Disaster resilience and complex adaptive systems theory: Finding common grounds for risk reduction. *Disaster Prevention and Management: An International Journal*, 25(2):196–211, 2016.

[37] David Etkin. *Disaster Theory: An Interdisciplinary Approach to Concepts and Causes*. Butterworth-Heinemann, Boston, MA, 2016.

[38] John Henry Holland. *Hidden Order: How Adaptation Builds Complexity*. Helix Books, Cambridge, MA., 2003.

[39] J. Stephen Lansing. Complex adaptive systems. *Annual Review of Anthropology*, 32(1):183–204, 2003.

[40] Simon A. Levin. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems*, 1(5):431–436, 1998.

[41] Michael Batty. *The New Science of Cities*. MIT Press, Cambridge, MA, 2013.

[42] Brian D. Fath, Carly A. Dean, and Harald Katzmair. Navigating the adaptive cycle: An approach to managing the resilience of social systems. *Ecology and Society*, 20, 2015.

[43] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, NY, 1993.

[44] Philip Buckle. Disaster: Mandated definitions, local knowledge and complexity. In Havidán Rodríguez, Enrico Louis Quarantelli, and Russell Rowe Dynes, editors, *What is Disaster? New Answers to Old Questions*, pages 173–200. Xlibris, Philadelphia, PA, 2005.

[45] Louise K. Comfort. Self-organization in complex systems. *Journal of Public Administration Research and Theory: J-PART*, 4(3):393–410, 1994.

[46] Enrico L. Quarantelli and Russell R. Dynes. Response to social crisis and disaster. *Annual Review of Sociology*, pages 23–49, 1977.

[47] Peter Timmerman. *Vulnerability, Resilience and the Collapse of Society: A Review of Models and Possible Climatic Applications*. Institute for Environmental Studies, University of Toronto, Toronto, Canada, 1981.

[48] Antonella Cavallo and Vernon Ireland. Preparing for complex interdependent risks: A system of systems approach to building disaster resilience. *International Journal of Disaster Risk Reduction*, 9:181–193, 2014.

[49] Claudio Cioffi-Revilla. Universal laws of disaster. In *2016 IEEE Global Humanitarian Technology Conference (GHTC)*, pages 272–279, 2016.

[50] L. K. Comfort, Yesim Sungu, David Johnson, and Mark Dunn. Complex systems in crisis: Anticipation and resilience in dynamic environments. *Journal of Contingencies and Crisis Management*, 9(3):144–158, 2001.

[51] Lowell Juilliard Carr. Disaster and the sequence-pattern concept of social change. *American Journal of Sociology*, 38(2):207–218, 1932.

[52] J. Eugene Haas, Robert W. Kates, and Martyn J. Bowden, editors. *Reconstruction Following Disaster*. MIT Press, Cambridge, MA, 1977.

[53] Richard Gist and Bernard Lubin, editors. *Response to Disaster: Psychosocial, Community, and Ecological Approaches*. Brunner/Mazel, Philadelphia, PA, 1999.

[54] Steward TA Pickett, Mary L Cadenasso, J Morgan Grove, Christopher G Boone, Peter M Groffman, Elena Irwin, Sujay S Kaushal, Victoria Marshall, Brian P McGrath, Charles H Nilon, et al. Urban ecological systems: Scientific foundations and a decade of progress. *Journal of Environmental Management*, 92(3):331–362, 2011.

[55] Richard G Little. Controlling cascading failure: Understanding the vulnerabilities of interconnected infrastructures. *Journal of Urban Technology*, 9(1):109–123, 2002.

[56] David M. Abramson, Lynn M. Grattan, Brian Mayer, Craig E. Colten, Farah A. Arosemena, Ariane Bedimo-Rung, and Maureen Lichtveld. The resilience activation framework: A conceptual model of how access to social resources promotes adaptation and rapid recovery in post-disaster settings. *The Journal of Behavioral Health Services & Research*, 42(1):42–57, 2015.

[57] Mohamed F. M Firdhous and P. M. Karuratane. A model for enhancing the role of information and communication technologies for improving the resilience of rural communities to disasters. *Procedia Engineering*, 212:707–714, 2018.

[58] David Mendonça and William A Wallace. Factors underlying organizational resilience: The case of electric power restoration in New York City after 11 September 2001. *Reliability Engineering & System Safety*, 141:83–91, 2015.

[59] Sidney Dekker. *Drift into Failure: From Hunting Broken Components to Understanding Complex Systems*. CRC Press, Boca Raton, FL, 2016.

[60] Sidney Dekker and Shawn Pruchnicki. Drifting into failure: Theorising the dynamics of disaster incubation. *Theoretical Issues in Ergonomics Science*, 15:534–544, 2014.

[61] David Mendonça, William A. Wallace, Barbara Cutler, and James Brooks. Synthetic environments for investigating collaborative information seeking: An application in emergency restoration of critical infrastructures. *Journal of Homeland Security and Emergency Management*, 12(3):763–784, 2015.

[62] Charles Perrow. The meltdown was not an accident. In Michael Lounsbury and Paul M Hirsch, editors, *Markets on Trial: The Economic Sociology of the U.S. Financial Crisis*, Research in the Sociology of Organizations, pages 309–330. Emerald Group Publishing Limited, Bingley, UK, 2010.

[63] Nick F. Pidgeon and M. O'Leary. Man-made disasters: why technology and organizations (sometimes) fail. *Safety Science*, 34(1):15–30, 2000.

[64] Joseph H. Saleh, Karen B. Marais, Efstathios Bakolas, and Raghvendra V. Cowlagi. Highlights from the Literature on Accident Causation and System Safety: Review of major ideas, recent contributions, and challenges. *Reliability Engineering & System Safety*, 95(11):1105–1116, 2010.

[65] Barry A. Turner and Nick F. Pidgeon. *Man-made Disasters*. Butterworth-Heinemann, Boston, MA, 1997.

[66] David A. McEntire. *Disaster Response and Recovery: Strategies and Tactics for Resilience*. Wiley, Hoboken, NJ, second edition edition, 2015.

[67] Robert E. Overstreet, Dianne Hall, Joe B. Hanna, and R. Kelly Rainer. Research in humanitarian logistics. *Journal of Humanitarian Logistics and Supply Chain Management*, 1(2):114–131, 2011.

[68] Erin Smith, Jason Wasiak, Ayan Sen, Frank Archer, and Frederick M. Burkle. Three decades of disasters: A review of disaster-specific literature from 1977-2009. *Prehospital and Disaster Medicine*, 24(4):306–311, 2009.

[69] Kathleen J. Tierney, Michael K. Lindell, and Ronald W. Perry, editors. *Facing Hazards and Disasters: Understanding Human Dimensions*. National Academies Press, Washington, DC, 2006.

[70] William H. Form, Charles P. Loomis, Roy A. Clifford, Harry E. Moore, Sigmund Nosow, Gregory P. Stone, and Charles M. Westie. The persistence and emergence of social and cultural systems in disasters. *American Sociological Review*, 21(2):180–185, 1956.

[71] Samuel Henry Prince. *Catastrophe and Social Change: Based Upon a Sociological Study of the Halifax Disaster*. PhD thesis, Columbia University, New York, NY, 1920.

[72] Anthony F. C. Wallace. Human Behavior in Extreme Situations: A Survey of the Literature and Suggestions for Further Research. Technical Report 390, National Academy of Sciences – National Research Council, 1956.

[73] Charles E. Fritz and Eli S. Marks. The NORC studies of human behavior in disaster. *Journal of Social Issues*, 10(3):26–41, 1954.

[74] Pitirim Aleksandrovich Sorokin. *Man and Society in Calamity*. Transaction Publishers, New Brunswick, NJ, 2010(1942).

[75] Gilbert F. White. Human Adjustments to Floods. PhD Dissertation, Department of Geography, University of Chicago, Chicago, IL, 1945.

[76] Ian Burton, Robert William Kates, and Gilbert F. White. The human ecology of extreme geophysical events. Working Paper 1, Department Geology, University of Toronto, Toronto, Canada, 1968.

[77] Kenneth Hewitt and Ian Burton. *The Hazardousness of a Place: A Regional Ecology of Damaging Events*. University of Toronto Press, Toronto, Canada, 1971.

[78] Robert W. Kates. Natural hazard in human ecological perspective: Hypotheses and models. *Economic Geography*, 47(3):438–451, 1971.

[79] Ian Burton, Robert W. Kates, and Gilbert F. White. *The Environment as Hazard*. Oxford University Press, New York, NY, 1978.

[80] Dennis S. Mileti, Thomas E. Drabek, and J. Eugene Haas. Human systems in extreme environments: A sociological perspective. Monograph 21, Institute of Behavioral Science, The University of Colorado, Boulder, CO, 1975.

[81] Michael Barkun. Disaster in history. *Mass Emergencies*, 2:219–231, 1977.

[82] Thomas Lum and Rhoda Margesson. Typhoon Haiyan (Yolanda): US and international response to Philippines disaster. *Current Politics and Economics of South, Southeastern, and Central Asia*, 23(2):209–246, 2014.

[83] Ljubica Mamula-Seadon and Ian McLean. Response and early recovery following 4 September 2010 and 22 February 2011 Canterbury earthquakes: Societal resilience and the role of governance. *International Journal of Disaster Risk Reduction*, 14:82–95, 2015.

[84] Kenneth Hewitt, editor. *Interpretations of Calamity From the Viewpoint of Human Ecology*. Allen & Unwin, Boston, MA, 1983.

[85] Phil O'Keefe, Ken Westgate, and Ben Wisner. Taking the naturalness out of natural disasters. *Nature*, 260:566–567, 1976.

[86] Piers Blaikie, Terry Cannon, Ian Davis, and Ben Wisner. *At Risk: Natural Hazards, People's Vulnerability, and Disasters*. Routledge, New York, NY, 1994.

[87] Frederick L. Bates and Carlo Pelanda. An ecological approach to disasters. In Russell R. Dynes and Kathleen J. Tierney, editors, *Disasters, Collective Behavior, and Social Organization*, pages 145–159. University of Delaware Press, Newark, NJ, 1994.

[88] Terry Cannon. Vulnerability analysis and the explanation of 'natural' disasters. In Anne Varley, editor, *Disasters, Development and Environment*, pages 13–30. John Wiley & Sons Ltd, New York, NY, 1994.

[89] Fikret Berkes, Johan Colding, and Carl Folke. *Navigating Social-ecological Systems: Building Resilience for Complexity and Change*. Cambridge University Press, Cambridge, UK, 2003.

[90] Lance H. Gunderson and C. S. Holling, editors. *Panarchy: Understanding Transformations in Human and Natural Systems*. Island Press, Washington, DC, 2002.

[91] Per Olsson, Carl Folke, and Fikret Berkes. Adaptive comanagement for building resilience in social-ecological systems. *Environmental Management*, 34(1), 2004.

[92] Brian Walker, Lance Gunderson, Ann Kinzig, Carl Folke, Steve Carpenter, and Lisen Schultz. A handful of heuristics and some propositions for understanding resilience in social-ecological systems. *Ecology and Society*, 11(1):13, 2006.

[93] Mark Pelling. Learning from others: The scope and challenges for participatory disaster risk assessment. *Disasters*, 31(4):373–385, 2007.

[94] W. Neil Adger, Nigel W. Arnell, and Emma L. Tompkins. Adapting to climate change: Perspectives across scales. *Global Environmental Change*, 15(2):75–76, 2005.

[95] W. Neil Adger. Social and ecological resilience: Are they related? *Progress in Human Geography*, 24(3):347–364, 2000.

[96] Louise K. Comfort, Namkyung Oh, and Gunes Ertan. The dynamics of disaster recovery: Resilience and entropy in hurricane response systems 2005-2008. *Public Organization Review*, 9(4):309–323, 2009.

[97] Naim Kapucu. Interorganizational coordination in complex environments of disasters: The evolution of intergovernmental disaster response systems. *Journal of Homeland Security and Emergency Management*, 6(1):47, 2009.

[98] O Pérez-Maqueo, A Intralawan, and ML Martínez. Coastal disasters from the perspective of ecological economics. *Ecological Economics*, 63(2-3):273–284, 2007.

[99] Adam Rose. Economic resilience to natural and man-made disasters: Multidisciplinary origins and contextual dimensions. *Environmental Hazards*, 7(4):383–398, 2007.

[100] Hongjian Zhou, Jinhong Wan, Huicong Jia, et al. Resilience to natural hazards: A geographic perspective. *Natural Hazards*, 53(1):21–41, 2010.

[101] Susan L. Cutter, Lindsey Barnes, Melissa Berry, Christopher Burton, Elijah Evans, Eric Tate, and Jennifer Webb. A place-based model for understanding community resilience to natural disasters. *Global Environmental Change*, 18(4):598–606, 2008.

[102] Olivia Patterson, Frederick Weil, and Kavita Patel. The role of community in disaster response: Conceptual models. *Population Research and Policy Review*, 29(2):127–141, 2010.

[103] Wendy S.A. Saunders and Julia S. Becker. A discussion of resilience and sustainability: Land use planning recovery from the Canterbury earthquake sequence, New Zealand. *International Journal of Disaster Risk Reduction*, 14:73–81, 2015.

[104] Benigno E. Aguirre. On the concept of resilience. Preliminary Paper 356, University of Deleware Disaster Research Center, Newark, DE, 2006.

[105] Susan L. Cutter. The landscape of disaster resilience indicators in the USA. *Natural Hazards*, 80(2):741–758, 2016.

[106] Fran H. Norris, Susan P. Stevens, Betty Pfefferbaum, Karen F. Wyche, and Rose L. Pfefferbaum. Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *American Journal of Community Psychology*, 41(1-2):127–150, 2008.

[107] Adam Rose. Defining and measuring economic resilience to disasters. *Disaster Prevention and Management: An International Journal*, 13(4):307–314, 2004.

[108] Oluwadunsin Ajulo, Jason Von-Meding, and Patrick Tang. Upending the status quo through transformative adaptation: A systematic literature review. *Progress in Disaster Science*, page 100103, 2020.

[109] Daniel J. Alesch and William Siembieda. The role of the built environment in the recovery of cities and communities from extreme events. *International Journal of Mass Emergencies & Disasters*, 30(2):197–211, 2012.

[110] Stephen J. Guastello. Self-organization and leadership emergence in emergency response teams. *Nonlinear Dynamics, Psychology, and Life Sciences*, 14(2):179–204, 2010.

[111] Lance Gunderson. Ecological and human community resilience in response to natural disasters. *Ecology and Society*, 15(2):18, 2010.

[112] Laurie A. Johnson and Haruo Hayashi. Synthesis efforts in disaster recovery research. *International Journal of Mass Emergencies & Disasters*, 30(2):212–238, 2012.

[113] Judith Steele and Niraj Verma. A systems approach to long-term urban disaster recovery. *Journal of Security Education*, 1(4):145–157, 2006.

[114] Gianluca Pescaroli and David Alexander. Critical infrastructure, panarchies and the vulnerability paths of cascading disasters. *Natural Hazards*, 82(1):175–192, 2016.

[115] Crawford S Holling. Understanding the complexity of economic, ecological, and social systems. *Ecosystems*, 4(5):390–405, 2001.

[116] Russell R. Dynes and Enrico Louis Quarantelli. The place of the 1917 explosion in Halifax harbor in the history of disaster research: The work of Samuel H. Prince. Preliminary Paper 182, University of Deleware Disaster Research Center, Newark, DE, 1993.

[117] T. Joseph Scanlon. Disaster's little known pioneer: Canada's Samuel Henry Prince. *International Journal of Mass Emergencies and Disasters*, 6(3):213–232, 1988.

[118] Enrico L Quarantelli. The nature and conditions of panic. *American Journal of Sociology*, 60(3):267–275, 1954.

[119] Ruth Leeds. Altruism and the norm of giving. *Merrill-Palmer Quarterly of Behavior and Development*, 9(3):229–240, 1963.

[120] William A. Anderson. Some observations on a disaster subculture: The organizational response of Cincinnati, Ohio, to the 1964 flood. Research Note 1964-0309, Disaster Research Center, Columbus, OH, 1965.

[121] Thomas E. Drabek and Keith S. Boggs. Families in disaster: Reactions and relatives. *Journal of Marriage and the Family*, 30(3):443, 1968.

[122] Ralph H. Turner. Types of solidarity in the reconstituting of groups. *The Pacific Sociological Review*, 10(2):60–68, 1967.

[123] Russell R. Dynes, Dwight Billings, Sally Maggard, and Kai T. Erikson. Two views of an award winning book. *Social Forces*, 57(2):721–723, 1978.

[124] Lynda Ann Ewen and Julia A. Lewis. Buffalo Creek revisited: Deconstructing Kai Erikson's stereotypes. *Appalachian Journal*, 27(1):22–45, 1999.

[125] Kai Erikson. *Everything in its Path: Destruction of Community in the Buffalo Creek Flood*. Simon and Schuster, New York, NY, 1976.

[126] Enrico L Quarantelli and Russell R Dynes. Community conflict: Its absence and its presence in natural disasters. *Mass Emergencies*, 1:139–152, 1976.

[127] Anthony Oliver-Smith. Post disaster consensus and conflict in a traditional society: The 1970 avalanche of Yungay, Peru. *Mass Emergencies*, 4:39–52, 1979.

[128] Stephen R. Couch and J. Stephen Kroll-Smith. The chronic technical disaster: Toward a social scientific perspective. *Social Science Quarterly*, 66(3):564–575, 1985.

[129] Beverly H. Cuthbertson and Joanne M. Nigg. Technological disaster and the nontherapeutic community: A question of true victimization. *Environment and Behavior*, 19(4):462–483, 1987.

[130] William R. Freudenburg. Contamination, corrosion and the social order: An overview. *Current Sociology*, 45(3):19–39, 1997.

[131] Thomas E. Drabek, William H. Key, Patricia E. Erickson, and Juanita L. Crowe. The impact of disaster on kin relationships. *Journal of Marriage and the Family*, 37(3):481–494, 1975.

[132] Orjan Hultaker. Family and disaster. *International Journal of Mass Emergencies and Disasters*, 1(2):7–18, 1983.

[133] Enrico Louis Quarantelli. Emergent citizen groups in disaster preparedness and recovery activities. Final Project Report 33, Disaster Research Center, University of Deleware, Newark, DE, 1984.

[134] Robert Bolin. Family recovery from natural disaster: A preliminary model. *Mass Emergencies*, 1:267–277, 1976.

[135] Robert Bolin and Patricia Bolton. Race, religion and ethnicity in disaster recovery. Program on Environment and Behavior Monograph 42, Institute of Behavioral Science, University of Colorado, Boulder, CO, 1986.

[136] Frederick L. Bates. The Social Network. Unpublished Manuscript, Department of Sociology, University of Georgia, Athens, GA, 1989.

[137] Walter Gillis Peacock. In search of social structure. *Sociological Inquiry*, 61(3):281–298, 1991.

[138] Benigno E. Aguirre, Dennis Wenger, and Gabriela Vigo. A test of the emergent norm theory of collective behavior. *Sociological Forum*, 13:301–320, 1998.

[139] Philip Berke and Timothy Beatley. *After the Hurricane: Linking Recovery to Sustainable Development in the Caribbean*. Johns Hopkins University Press, Baltimore, MD, 1997.

[140] John J. Beggs, Valerie A. Haines, and Jeanne S. Hurlbert. Situational contingencies surrounding the receipt of informal support. *Social Forces*, 75(1):201–222, 1996.

[141] Ben Wisner and Henry R. Luce. Disaster vulnerability: Scale, power and daily life. *GeoJournal*, 30(2):127–140, 1993.

[142] Krzysztof Kaniasty and Fran H. Norris. In search of altruistic community: Patterns of social support mobilization following Hurricane Hugo. *American Journal of Community Psychology*, 23(4):447–477, 1995.

[143] Robert C Bolin and Lois Stanford. *The Northridge Earthquake: Vulnerability and Disaster*. Routledge, New York, NY, 1998.

[144] Betty Hearn Morrow and Elaine Enarson. Hurricane Andrew through women's eyes: Issues and recommendations. *International Journal of Mass Emergencies and Disasters*, 14(1):5–22, 1996.

[145] Raymond E. Wiest, Jane SP Mocellin, and D. Thandiwe Motsisi. *The needs of women in disasters and emergencies*. Disaster Research Institute, University of Manitoba, Winnipeg, Canada, 1994.

[146] Valerie A. Haines, Jeanne S. Hurlbert, and John J. Beggs. Exploring the determinants of support provision: Provider characteristics, personal networks, community contexts, and support following life events. *Journal of Health and Social Behavior*, 37(3):252–264, 1996.

[147] Betty Hearn Morrow. Identifying and mapping community vulnerability. *Disasters*, 23(1):1–18, 1999.

[148] Walter Gillis Peacock, Betty Hearn Morrow, and Hugh Gladwin, editors. *Hurricane Andrew: Ethnicity, Gender, and the Sociology of Disasters*. Routledge, New York, NY, 1997.

[149] Greg Bankoff. Cultures of coping: Adaptation to hazard and living with disaster in the Philippines. *Philippine Sociological Review*, 51:1–16, 2003.

[150] Joshua D. Lichterman. A "community as resource" strategy for disaster response. *Public Health Reports*, 115(2-3):262–265, 2000.

[151] Eve Coles and Philip Buckle. Developing community resilience as a foundation for effective disaster recovery. *The Australian Journal of Emergency Management*, 19(4):6–15, 2004.

[152] Susan L. Cutter, Bryan J. Boruff, and W. Lynn Shirley. Social vulnerability to environmental hazards. *Social Science Quarterly*, 84(2):242–261, 2003.

[153] Ben Wisner. Disaster risk reduction in megacities: Making the most of human and social capital. In Alcira Kreimer, Margaret Arnold, and Anne Carlin, editors, *Building Safer Cities: the Future of Disaster Risk*, pages 181–196. World Bank, Washington, DC, 2003.

[154] Russell R. Dynes. The importance of social capital in disaster response. Preliminary Paper 327, Disaster Research Center, University of Delaware, Newark, DE, 2002.

[155] Russell R. Dynes. Community Social Capital as the Primary Basis for Resilience. Preliminary Paper 344, Disaster Research Center, University of Delaware, Newark, DE, 2005.

[156] Yuko Nakagawa and Rajib Shaw. Social capital: A missing link to disaster recovery. *International Journal of Mass Emergencies and Disasters*, 22(1):5–34, 2004.

[157] Rajib Shaw and Katsuihciro Goda. From disaster to sustainable civil society: The Kobe experience. *Disasters*, 28(1):16–40, 2004.

[158] Emily Chamlee-Wright. *The Cultural and Political Economy of Recovery: Social Learning in a Post-disaster Environment*. Routledge, New York, NY, 2010.

[159] Daniel P. Aldrich and Michelle A. Meyer. Social capital and community resilience. *American Behavioral Scientist*, 59(2):254–269, 2015.

[160] Graham A. Tobin and Linda M. Whiteford. Community resilience and volcano hazard: The eruption of Tungurahua and evacuation of the Faldas in Ecuador. *Disasters*, 26(1):28–48, 2002.

[161] Brigit Maguire and Patrick Hagan. Disasters and communities: Understanding social resilience. *The Australian Journal of Emergency Management*, 22(2):16–20, 2007.

[162] Ljubica Mamula-Seadon, Karen Selway, and Douglas Paton. Exploring resilience: Learning from Christchurch communities. *TEPHRA*, 23:5–7, 2012.

[163] Carl Folke, Thomas Hahn, Per Olsson, and Jon Norberg. Adaptive governance of social-ecological systems. *Annual Review of Environment and Resources*, 30(1):441–473, 2005.

[164] Riyanti Djalante, Cameron Holley, and Frank Thomalla. Adaptive governance and managing resilience to natural hazards. *International Journal of Disaster Risk Science*, 2(4):1–14, 2011.

[165] Margot Hurlbert and Joyeeta Gupta. Adaptive governance, Uncertainty, and risk: Policy framing and responses to climate change, drought, and flood. *Risk Analysis*, 36(2):339–356, 2016.

[166] Benjamin Wisner, J. C. Gaillard, and Ilan Kelman, editors. *The Routledge Handbook of Hazards and Disaster Risk Reduction*. Routledge, New York, NY, 2012.

[167] Lewis M. Killian. The significance of multiple-group membership in disaster. *American Journal of Sociology*, 57(4):309–314, 1952.

[168] James S Tyhurst. Psychological and social aspects of civilian disaster. *Canadian Medical Association Journal*, 76(5):385–393, 1957.

[169] Albert J. Glass. Psychological aspects of disaster. *Journal of the American Medical Association*, 171(2):222–225, 1959.

[170] Helene Rank Veltfort and George E. Lee. The Cocoanut Grove fire: A study in scapegoating. *The Journal of Abnormal and Social Psychology*, 38(2S):138—-154, 1943.

[171] Ralph Crawshaw. Reactions to a disaster. *Archives of General Psychiatry*, 9(2):157–162, 1963.

[172] Richard S. Lazarus. *Psychological Stress and the Coping Process*. McGraw-Hill, New York, NY, 1966.

[173] Susan Folkman and Richard S. Lazarus. If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology*, 48(1):150, 1985.

[174] Susan Folkman, Richard S. Lazarus, Christine Dunkel-Schetter, Anita DeLongis, and Rand J. Gruen. Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology*, 50(5):992, 1986.

[175] Ronald W. Perry and Michael K. Lindell. The psychological consequences of natural disaster: A review of research on American communities. *Mass Emergencies*, 3:105–115, 1978.

[176] Anthony R. Mawson. Panic behavior: A review and a new hypothesis. In *9th World Congress of Sociology*, Uppsala, Sweden, 1978.

[177] Ralph H. Turner and Lewis M. Killian. *Collective Behavior*. Prentice-Hall, Englewood Cliffs, NJ, 3rd edition, 1987.

[178] Drue M. Johnston and Norris R. Johnson. Role extension in disaster: Employee behavior at the Beverly Hills supper club fire. *Sociological Focus*, 22(1):39–51, 1989.

[179] John Drury, Chris Cocking, and Steve Reicher. Everyone for themselves? A comparative study of crowd solidarity among emergency survivors. *British Journal of Social Psychology*, 48(3):487–506, 2009.

[180] Mark Levine and Kirstien Thompson. Identity, place, and bystander intervention: Social categories and helping after natural disasters. *The Journal of Social Psychology*, 144(3):229–245, 2004.

[181] Anthony V. Rubonis and Leonard Bickman. Psychological impairment in the wake of disaster: The disaster-psychopathology relationship. *Psychological Bulletin*, 109(3):384–399, 1991.

[182] James D. Goltz, Lisa A. Russell, and Linda B. Bourque. Initial behavioral response to a rapid onset disaster: A case study of the October 1, 1987 Whittier Narrows earthquake. *International Journal of Mass Emergencies and Disasters*, 10(1):43–69, 1992.

[183] Stevan E. Hobfoll. Conservation of resources: A new attempt at conceptualizing stress. *American psychologist*, 44(3):513, 1989.

[184] Stevan E. Hobfoll. The influence of culture, community, and the nested-self in the stress process: Advancing conservation of resources theory. *Applied Psychology: An International Review*, 50(3):337–421, 2001.

[185] Fran H. Norris, Matthew J. Friedman, Patricia J. Watson, Christopher M. Byrne, Eolia Diaz, and Krzysztof Kaniasty. 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry*, 65(3):207–239, 2002.

[186] Anthony R. Mawson. Understanding mass panic and other collective responses to threat and disaster. *Psychiatry: Interpersonal and Biological Processes*, 68(2):95–113, 2005.

[187] Anthony R. Mawson. *Mass Panic and Social Attachment: The Dynamics of Human Behavior*. Ashgate, Aldershot, UK, 2007.

[188] Douglas Paton. Disaster preparedness: A social-cognitive perspective. *Disaster Prevention and Management: An International Journal*, 12(3):210–216, 2003.

[189] Donald L. Rosenstein. Decision-making capacity and disaster research. *Journal of Traumatic Stress*, 17(5):373–381, 2004.

[190] Paul C. Van Fenema. Collaborative elasticity and breakdowns in high reliability organizations: Contributions from distributed cognition and collective mind theory. *Cognition, Technology & Work*, 7(2):134–140, 2005.

[191] Amanda Ripley. *The Unthinkable: Who Survives When Disaster Strikes and Why*. Crown Publishers, New York, NY, 2008.

[192] J. Richard Eiser, Ann Bostrom, Ian Burton, David M. Johnston, John McClure, Douglas Paton, Joop van der Pligt, and Mathew P. White. Risk interpretation and action: A conceptual framework for responses to natural hazards. *International Journal of Disaster Risk Reduction*, 1:5–16, 2012.

[193] Ervina Espina and Mendiola Teng-Calleja. A social cognitive approach to disaster preparedness. *Philippine Journal of Psychology*, 48(2):161–174, 2015.

[194] Albert Bandura. Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1):1–26, 2001.

[195] Charles C. Benight and Albert Bandura. Social cognitive theory of posttraumatic recovery: The role of perceived self-efficacy. *Behaviour Research and Therapy*, 42(10):1129–1148, 2004.

[196] Daniel Coq-Huelva, Angie Higuchi, Rafaela Alfalla-Luque, Ricardo Burgos-Morán, and Ruth Arias-Gutiérrez. Co-evolution and bio-social construction: The Kichwa agroforestry systems (Chakras) in the Ecuadorian Amazonia. *Sustainability*, 9(10):1920, 2017.

[197] Richard B Norgaard. Sociosystem and ecosystem coevolution in the Amazon. *Journal of Environmental Economics and Management*, 8(3):238–254, 1981.

[198] Jules Pretty. Interdisciplinary progress in approaches to address social-ecological and ecocultural systems. *Environmental Conservation*, 38(2):127–139, 2011.

[199] P. Ormerod. *Butterfly Economics: A New General Theory of Economic and Social Behaviour*. Faber and Faber, London, UK, 1998.

[200] Christopher Small and Terry Naumann. The global distribution of human population and recent volcanism. *Global Environmental Change Part B: Environmental Hazards*, 3(3):93–109, 2001.

[201] Katherine Donovan. Doing social volcanology: Exploring volcanic culture in Indonesia. *Area*, 42(1):117–126, 2010.

[202] S Bachri, J Stötter, M Monreal, and J Sartohadi. The calamity of eruptions, or an eruption of benefits? Mt. Bromo human-volcano system a case study of an open-risk perception. *Natural Hazards and Earth System Sciences*, 15(2):277–290, 2015.

[203] Supriyati Andreastuti, EkoTeguh Paripurno, Hendra Gunawan, Agus Budianto, Devy Syahbana, and John Pallister. Character of community response to volcanic crises at Sinabung and Kelud volcanoes. *Journal of Volcanology and Geothermal Research*, 2017.

[204] Karin Hansson, M. Danielson, and L. Ekenberg. A framework for evaluation of flood management strategies. *Journal of Environmental Management*, 86(3):465–480, 2008.

[205] Jennifer K. Poussin, W. J. Wouter Botzen, and Jeroen C. J. H. Aerts. Effectiveness of flood damage mitigation measures: Empirical evidence from French flood disasters. *Global Environmental Change*, 31:74–84, 2015.

[206] Bruno Merz, James Hall, Markus Disse, and Andreas Schumann. Fluvial flood risk management in a changing world. *Natural Hazards and Earth System Sciences*, 10(3):509–527, 2010.

[207] Giuliano Di Baldassarre, Alberto Viglione, Gemma Carr, Linda Kuil, JL Salinas, and G Blöschl. Socio-hydrology: Conceptualising human-flood interactions. *Hydrology and Earth System Sciences*, 17(8):3295–3303, 2013.

[208] National Transportation Safety Board. Grounding of the U.S. Tankship Exxon Valdez on Bligh Reef, Prince William Sound near Valdez, Alaska, March 24, 1989. Marine Accident NTSB/MAR-90-04, National Transportation Safety Board, Washington, DC, 1990.

[209] Elizabeth R. Millard. Anatomy of an oil spill: The Exxon Valdez and the Oil Pollution Act of 1990. *Seton Hall Legislation Journal*, 18:331–370, 1993.

[210] National Response Team (NRT). The Exxon Valdez Oil spill: A report to the president. Technical Report OSWER89VALDZ, National Response Team, US Environmental Protection Agency, Washington DC, 1989.

[211] Art Davidson. *In the Wake of the Exxon Valdez: The Devastating Impact of the Alaska Oil Spill*. Random House, Inc, New York, NY, 1990.

[212] John Keeble. *Out of the Channel: The Exxon Valdez Oil Spill in Prince William Sound*. Eastern Washington University, Cheney, WA, 10th Anniversary edition, 1999.

[213] J. Steven Picou and Duane A. Gill. The Exxon Valdez Oil Spill and Chronic Psychological Stress. *American Fisheries Society Symposium*, 18:879–893, 1996.

[214] John R. Harrald, Henry S. Marcus, and William A. Wallace. The Exxon Valdez: An assessment of crisis prevention and management systems. *INFORMS Journal on Applied Analytics*, 20(5):14–30, 1990.

[215] Christopher L. Dyer. Punctuated entropy as culture-induced change: The case of the Exxon Valdez oil spill. In Susanna Hoffman and Anthony Oliver-Smith, editors, *Catastrophe & Culture: The Anthropology of Disaster*, pages 159–185. School of American Research Press, Santa Fe, NM, 2002.

[216] Conrad Smith. News sources and power elites in news coverage of the Exxon Valdez oil spill. *Journalism Quarterly*, 70(2):393–403, 1993.

[217] Patricia Widener and Valerie J. Gunter. Oil spill recovery in the media: Missing an Alaska native perspective. *Society & Natural Resources*, 20(9):767–783, 2007.

[218] Daniel A. Farber. Probabilities behaving badly: Complexity theory and environmental uncertainty. *U.C. Davis Law Review*, 37(1):145–174, 2003.

[219] Henri Lefebvre. *The Urban Revolution*. University of Minnesota Press, Minneapolis, MN, 2003.

[220] Marcus J. Hamilton, Bruce T. Milne, Robert S. Walker, Oskar Burger, and James H. Brown. The complex structure of hunter-gatherer social networks. *Proceedings of the Royal Society B: Biological Sciences*, 274(1622):2195–2203, 2007.

[221] Wei-Xing Zhou, Didier Sornette, Russell A. Hill, and R. I. M. Dunbar. Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B: Biological Sciences*, 272(1561):439–444, 2005.

[222] Luís MA Bettencourt, José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B West. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences*, 104(17):7301–7306, 2007.

[223] Edward N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963.

[224] Marten Scheffer, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egbert H. van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.

[225] Peter M. Allen. *Cities and Regions as Self-Organizing Systems: Models of Complexity*. Environmental problems and social dynamics. Gordon and Breach, Luxembourg, 1997.

[226] Nigel Gilbert, Petra Ahrweiler, Pete Barbrook-Johnson, Kavin Preethi Narasimhan, and Helen Wilkinson. Computational modelling of public policy: Reflections on practice. *Journal of Artificial Societies and Social Simulation*, 21(1):14, 2018.

[227] Hiroyuki Shibusawa. A dynamic spatial CGE approach to assess economic effects of a large earthquake in China. *Progress in Disaster Science*, page 100081, 2020.

[228] Philip R. Berke, Jack Kartez, and Dennis Wenger. Recovery after disaster: Achieving sustainable development, mitigation and equity. *Disasters*, 17(2):93–109, 1993.

[229] Jean-Christophe Gaillard and Jessica Mercer. From knowledge to action: Bridging gaps in disaster risk reduction. *Progress in Human Geography*, 37(1):93–114, 2013.

[230] Claire B. Rubin, Martin D. Saperstein, and Daniel G. Barbee. Community Recovery from a Major Natural Disaster. Technical Report 41, Program on Environment and Behavior, Institute of Behavioral Science University of Colorado, Boulder, CO, 1985.

[231] Jessica L Liddell, Leia Y Saltzman, Regardt J Ferreira, and Amy E Lesen. Cumulative disaster exposure, gender and the protective action decision model. *Progress in Disaster Science*, 5:100042, 2020.

[232] Rosa Vicari, Ioulia Tchiguirinskaia, Bruno Tisserand, and Daniel Schertzer. Climate resilience in Paris: A network representation of online strategic documents released by public authorities. *Progress in Disaster Science*, 3:100040, 2019.

[233] Michael Schmitz, Julio J Hernández, Víctor Rocabado, Jean Domínguez, Cecilio Morales, Maxlimer Valleé, Kenny García, Javier Sánchez-Rojas, André Singer, Javier Oropeza, et al. The Caracas, Venezuela, seismic microzoning project: Methodology, results, and implementation for seismic risk reduction. *Progress in Disaster Science*, 5:100060, 2020.

[234] Sarah Henly-Shepard, Steven A. Gray, and Linda J. Cox. The use of participatory modeling to promote social learning and facilitate community disaster planning. *Environmental Science & Policy*, 45:109–122, 2015.

[235] Catharina Landström, Sarah J Whatmore, Stuart N Lane, Nicholas A Odoni, Neil Ward, and Susan Bradley. Coproducing flood risk knowledge: Redistributing expertise in critical ‘participatory modelling’. *Environment and Planning A*, 43(7):1617–1633, 2011.

[236] Zhanli Sun, Iris Lorscheid, James D Millington, Steffen Lauf, Nicholas R Magliocca, Jürgen Groeneveld, Stefano Balbi, Henning Nolzen, Birgit Müller, Jule Schulze, et al. Simple or complicated agent-based models? a complicated issue. *Environmental Modelling & Software*, 86:56–67, 2016.

[237] Robin L. Ersing and Kathleen A. Kost, editors. *Surviving disaster: the role of social networks*. Lyceum Books, Chicago, Ill, 2012.

[238] David G. Casagrande, Heather McIlvaine-Newsad, and Eric C. Jones. Social networks of help-seeking in different types of disaster responses to the 2008 Mississippi river floods. *Human Organization*, 74(4):351–361, 2015.

[239] Russell Dynes. Social capital: Dealing with community emergencies. *Homeland Security Affairs*, 2(2), 2006.

[240] Lewis M. Killian. An Introduction to Methodological Problems of Field Studies in Disasters. In *Methods of Disaster Research*, pages 49–93. Xlibris, Philadelphia, 2002.

[241] Michael Schweinberger, Miruna Petrescu-Prahova, and Duy Q. Vu. Disaster response on September 11, 2001 through the lens of statistical network analysis. *Social Networks*, 37:42–55, 2014.

[242] Qian Hu, Claire Connolly Knox, and Naim Kapucu. What have we learned since September 11, 2001? A network study of the Boston marathon bombings response. *Public Administration Review*, 74(6):698–712, 2014.

[243] Nicholas Geard, James M McCaw, Alan Dorin, Kevin B Korb, and Jodie McVernon. Synthetic Population Dynamics: A Model of Household Demography. *Journal of Artificial Societies and Social Simulation*, 16(1):8, 2013.

[244] Christopher L Barrett, Richard J Beckman, Maleq Khan, VS Anil Kumar, Madhav V Marathe, Paula E Stretz, Tridib Dutta, and Bryan Lewis. Generation and Analysis of Large Synthetic Social Contact Networks. In *Proceedings of the 2009 Winter Simulation Conference*, pages 1003–1014, Austin, TX, 2009.

[245] Kirill Müller and Kay W Axhausen. Population synthesis for microsimulation: State of the art. page 16, 2011.

[246] Sarah Wise. Using social media content to inform agent-based models for humanitarian crisis response. PhD Dissertation, George Mason University, Fairfax, VA, 2014.

[247] Zhijing Xu, Kathryn Glass, Colleen L. Lau, Nicholas Geard, Patricia Graves, and Archie Clements. A synthetic population for modelling the dynamics of infectious disease transmission in American Samoa. *Scientific Reports*, 7(1):16725, 2017. Number: 1 Publisher: Nature Publishing Group.

[248] Yi Zhu and Joseph Ferreira. Synthetic population generation at disaggregated spatial scales for land use and transportation microsimulation. *Transportation Research Record*, 2429(1):168–177, 2014. Publisher: SAGE Publications Inc.

[249] Guy H Orcutt. A New Type of Socio-Economic System. *Review of Economics and Statistics*, 39(2):116–123, 1957.

[250] Mark Birkin and Belinda Wu. A review of microsimulation and hybrid agent-based approaches. In Alison J. Heppenstall, Andrew T. Crooks, Linda M. See, and Michael Batty, editors, *Agent-Based Models of Geographical Systems*, pages 51–68. Springer, New York, NY, 2012.

[251] Johan Barthelemy and Philippe L. Toint. Synthetic population generation without a sample. *Transportation Science*, 47(2):266–279, 2013.

[252] David R. Pritchard and Eric J. Miller. Advances in population synthesis: fitting many attributes per agent and fitting to household and person margins simultaneously. *Transportation*, 39(3):685–704, 2012.

[253] W. Edwards Deming and Frederick F. Stephan. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4):427–444, 1940.

[254] Zengyi Huang and Paul Williamson. A comparison of synthetic reconstruction and combinatorial optimisation approaches to the creation of small-area microdata. Working paper, Department of Geography, University of Liverpool, 2001.

[255] Bilal Farooq, Michel Bierlaire, Ricardo Hurtubia, and Gunnar Flötteröd. Simulation based population synthesis. *Transportation Research Part B: Methodological*, 58:243–263, 2013.

[256] Poh Ping Lim. *Population synthesis for travel demand modelling in Australian capital cities*. PhD Thesis, The University of Queensland, 2020.

[257] William D. Wheaton, James C. Cajka, Bernadette M. Chasteen, Diane K. Wagener, Philip C. Cooley, Laxminarayana Ganapathi, Douglas J. Roberts, and Justine L. Allpress. Synthesized population databases: A US geospatial database for agent-based models. *Methods report (RTI Press)*, 2009(10):905, 2009.

[258] Paul Anderson, Bilal Farooq, Dimitrios Efthymiou, and Michel Bierlaire. Associations generation in synthetic population for transportation applications: Graph-theoretic solution. *Transportation Research Record*, 2429(1):38–50, 2014.

[259] Hao Wu, Yue Ning, Prithwish Chakraborty, Jilles Vreeken, Nikolaj Tatti, and Naren Ramakrishnan. Generating realistic synthetic population datasets. *ACM Transactions on Knowledge Discovery from Data*, 12(4):45:1–45:22, 2018.

[260] Yong Yang and Peter M Atkinson. Individual space–time activity-based model: A model for the simulation of airborne infectious-disease transmission by activity-bundle simulation. *Environment and Planning B: Planning and Design*, 35(1):80–99, 2008.

[261] Armando Geller and Scott Moss. Growing qawm: An evidence-driven declarative model of Afghan power structures. *Advances in Complex Systems (ACS)*, 11:321–335, 2008.

[262] Chris J. Kuhlman, V. S. Anil Kumar, Madhav V. Marathe, S. S. Ravi, and Daniel J. Rosenkrantz. Inhibiting diffusion of complex contagions in social networks: theoretical and experimental results. *Data Mining and Knowledge Discovery*, 29(2):423–465, 2015.

[263] William Rand and Roland T. Rust. Agent-based modeling in marketing: Guidelines for rigor. *International Journal of Research in Marketing*, 28(3):181–193, 2011.

[264] Arun Agrawal, Daniel G. Brown, Gautam Rao, Rick Riolo, Derek T. Robinson, and Michael Bommarito. Interactions between organizations and networks in common-pool resource governance. *Environmental Science & Policy*, 25:138–146, 2013.

[265] Bianica Pires and Andrew T. Crooks. Modeling the emergence of riots: A geosimulation approach. *Computers, Environment and Urban Systems*, 61:66–80, 2017.

[266] Samuel Thiriot and Jean-Daniel Kant. Generate country-scale networks of interaction from scattered statistics. In *Fifth ESSA Conference in Social Simulation 2008*, page 12, Brescia, Italy, 2008.

[267] Shah Jamal Alam and Armando Geller. Networks in agent-based social simulation. In Alison J. Heppenstall, Andrew T. Crooks, Linda M. See, and Michael Batty, editors, *Agent-Based Models of Geographical Systems*, pages 199–216. Springer, New York, NY, 2012.

[268] Lynne Hamill and Nigel Gilbert. Social circles: A simple structure for agent-based social network models. *Journal of Artificial Societies and Social Simulation*, 12(2):1–3, 2009. https://ideas.repec.org/a/jas/jasssj/2008-70-2.html.

[269] Meysam Alizadeh, Claudio Cioffi-Revilla, and Andrew Crooks. Generating and analyzing spatial social networks. *Computational and Mathematical Organization Theory*, 23(3):362–390, 2017. Publisher: Springer.

[270] Stephen Eubank, Hasan Guclu, V. S. Anil Kumar, Madhav V. Marathe, Aravind Srinivasan, Zoltán Toroczkai, and Nan Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 2004.

[271] Christopher L. Barrett, Richard J. Beckman, Maleq Khan, V. S. Anil Kumar, Madhav V. Marathe, Paula E. Stretz, Tridib Dutta, and Bryan Lewis. Generation and analysis of large synthetic social contact networks. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, pages 1003–1014, Austin, TX, USA, 2009. IEEE.

[272] US Census Burean. *2010 TIGER/Line Shapefiles*, 2010. https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2010layergroup=Roads.

[273] US Census Bureau. *Decennial Census (2010, 2000)*, 2000. https://www.census.gov/data/developers/data-sets/decennial-census.html.

[274] US Census Bureau. *County Business Patterns: 2010*, 2010. https://www.census.gov/data/datasets/2010/econ/cbp/2010-cbp.html.

[275] Jan Eeckhout. Gibrat's law for (all) cities. *American Economic Review*, 94(5):1429–1451, 2004.

[276] US Census Bureau Center for Economic Studies. *US Census Bureau Center for Economic Studies Publications and Reports Page*, 2019. https://lehd.ces.census.gov/announcements.html.

[277] USEPA Office of Environmental Information (OEI). *Educational Institutions, US, 2015, ORNL, SEGS*, 2015. https://edg.epa.gov/metadata/catalog/search/resource/details.page?uuid=%7B9C49AE4B-F175-43D0-BCC6-A928FF54C329%7D.

[278] Robin I. M. Dunbar. The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5):178–190, 1998.

[279] Mark EJ Newman and Duncan J Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4-6):341–346, 1999.

[280] RTI. *RTI U.S. Synthetic Household Population*, 2019. https://www.rti.org/impact/rti-us-synthetic-household-population.

[281] Andrew Crooks, Nicolas Malleson, Ed Manley, and Alison Heppenstall. *Agent-Based Modelling and Geographical Information Systems: A Practical Primer*. Sage, London, UK, 2019.

[282] Nigel Gilbert and Klaus Troitzsch. *Simulation for the Social Scientist*. Open University Press, Milton Keynes, UK, 2nd edition edition, 2005.

[283] Andrew Crooks, Alison Heppenstall, Nicolas Malleson, and Ed Manley. Agent-based modeling and the city: A gallery of applications. In Wenzhong Shi, Michael F. Goodchild, Mike Batty, Mei-Po Kwan, and Anshu Zhang, editors, *Urban Informatics*. Springer, New York, NY, 2021.

[284] Nicholas Magliocca, Elena Safirova, Virginia McConnell, and Margaret Walls. An economic agent-based model of coupled housing and land markets (CHALMS). *Computers, Environment and Urban Systems*, 35(3):183–191, 2011.

[285] Atesmachew Hailegiorgis, Andrew Crooks, and Claudio Cioffi-Revilla. An agent-based model of rural households' adaptation to climate change. *Journal of Artificial Societies and Social Simulation*, 21(4):4, 2018.

[286] Joshua M Epstein and Robert Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. MIT Press, Cambridge, MA, 1996.

[287] H Randy Gimblett, editor. *Integrating Geographic Information Systems and Agent-Based Modelling Techniques for Simulating Social and Ecological Processes*. Oxford University Press, Oxford, UK, 2002.

[288] Itzhak Benenson and Paul Torrens. *Geosimulation: Automata-Based Modelling of Urban Phenomena*. John Wiley & Sons, London, UK, 2004.

[289] Yan Yang, Liang Mao, and Sara S. Metcalf. Diffusion of hurricane evacuation behavior through a home-workplace social network: A spatially explicit agent-based simulation model. *Computers, Environment and Urban Systems*, 74:13–22, 2019.

[290] A. Yair Grinberger and Daniel Felsenstein. Dynamic agent based simulation of welfare effects of urban disasters. *Computers, Environment and Urban Systems*, 59:129–141, 2016.

[291] Richard J. Dawson, Roger Peppe, and Miao Wang. An agent-based model for risk-based flood incident management. *Natural Hazards*, 59(1):167–189, 2011.

[292] Sean Luke, Claudio Cioffi-Revilla, Liviu Panait, Keith Sullivan, and Gabriel Balan. MASON: A multiagent simulation environment. *SIMULATION*, 81(7):517–527, 2005.

[293] S. Luke, R. Simon, A.T. Crooks, H. Wang, E. Wei, D. Freelan, C. Spagnuolo, V. Scarano, G. Cordasco, and C. Cioffi-Revilla. The MASON Simulation Toolkit: Past, Present, and Future. In Davidsson P. and Verhagen H., editors, *Proceedings of the 19th International Workshop on Multi-Agent-Based Simulation*, pages 75–87, Stockholm, Sweden, 2018.

[294] ESRI. Shapefile technical description. *An ESRI white paper*, 4:1, 1998.

[295] P Russel Norvig. *Artificial Intelligence: A modern approach*. Prentice Hall, 2002.

[296] Samuel Glasstone, Philip J Dolan, et al. *The Effects of Nuclear Weapons*, volume 50. US Department of Defense, 1977.

[297] Paul M Torrens. Agent-based models and the spatial sciences. *Geography Compass*, 4(5):428–448, 2010.

[298] Sirui Liu, Pamela Murray-Tuite, and Lisa Schweitzer. Analysis of child pick-up during daily routines and for daytime no-notice evacuations. *Transportation Research Part A: Policy and Practice*, 46(1):48–67, 2012.

[299] N. Emel Ganapati. In good company: Why social capital matters for women during disaster recovery. *Public Administration Review*, 72(3):419–427, 2012.

[300] Sasha Rudenstine and Sandro Galea. *The causes and behavioral consequences of disasters: models informed by the global experience 1950-2005*. Springer, New York, NY, 2012. OCLC: ocn733237463.

[301] Gerd Gigerenzer and Wolfgang Gaissmaier. Heuristic decision making. *Annual review of psychology*, 62(1):451–482, 2011.

[302] Robert Jay Lifton and Kai Erikson. Survivors of nuclear war: psychological and communal breakdown. In Eric. Chivian, John P. Robinson, Jonathan R.H. Tudge, Nikolai P. Popov, and Vladimir G. Andreyenkov, editors, *International Physicians for the Prevention of Nuclear War. Last Aid: The Medical Dimensions of Nuclear War*. WH Freeman, San Francisco, CA, 1982.

[303] Eisei Ishikawa and David L Swain. *Hiroshima and Nagasaki: The Physical, Medical, and Social Effects of the Atomic Bombings: the Committee for the Compilation of Materials on Damage Caused by the Atomic Bombs*. Basic Books, New York, NY, 1981.

[304] Susan Cutter and Kent Barnes. Evacuation behavior and Three Mile Island. *Disasters*, 6(2):116–124, 1982.

[305] Xuan Bien Do. Fukushima nuclear disaster displacement: How far people moved and determinants of evacuation destinations. *International Journal of Disaster Risk Reduction*, 33:235–252, 2019.

[306] TJ Shields, KE Boyce, and N McConnell. The behaviour and evacuation experiences of wtc 9/11 evacuees with self-designated mobility impairments. *Fire Safety Journal*, 44(6):881–893, 2009.

[307] Delin Liu, Yue Li, Shurong Fang, and Yongling Zhang. Influencing factors for emergency evacuation capability of rural households to flood hazards in western mountainous regions of henan province, china. *International Journal of Disaster Risk Reduction*, 21:187–195, 2017.

[308] Edd Hauser, Sherry M. Elmes, and Nicholas J. Schwartz. Risk, preparation, evacuation and rescue. In *Natural Disaster Analysis after Hurricane Katrina*, pages 93–119. Edward Elgar, Cheltenham, U.K. ; Northampton, MA, 2008.

[309] Thomas E Drabek. Social processes in disaster: Family evacuation. *Social Problems*, 16(3):336–349, 1969.

[310] Robert M Stein, Leonardo Dueñas-Osorio, and Devika Subramanian. Who evacuates when hurricanes approach? the role of risk, information, and location. *Social Science Quarterly*, 91(3):816–834, 2010.

[311] Jerry Buckland and Matiur Rahman. Community-based disaster management during the 1997 Red River Flood in Canada. *Disasters*, 23(2):174–191, 1999.

[312] Thomas Ferris, Erick Moreno-Centeno, Justin Yates, Kisuk Sung, Mahmoud El-Sherif, and David Matarrita-Cascante. Studying the usage of social media and mobile technology during extreme events and their Implications for evacuation decisions: A case study of hurricane Sandy. *International Journal of Mass Emergencies and Disasters*, 34(2):204–230, 2016.

[313] Qi Wang and John E. Taylor. Patterns and limitations of urban human mobility resilience under the influence of multiple types of natural disaster. *PLoS one*, 11(1):e0147299, 2016.

[314] Yingxin Chen. Research on group emotional intelligence effects on group decisionmaking under emergency. *International Journal of Hybrid Information Technology*, 7(6):189–198, 2014.

[315] Arnold R Parr. Organizational response to community crises and group emergence. *American Behavioral Scientist*, 13(3):423–429, 1970.

[316] Ira Helsloot and Arnout Ruitenberg. Citizen response to disasters: A survey of literature and some practical implications. *Journal of Contingencies and Crisis Management*, 12(3):98–111, 2004.

[317] Robert Stallings. The structural patterns of four types of organizations in disaster. In Enrico Louis Quarantelli, editor, *Disasters: Theory and Research*, pages 87–103. Sage, New York, NY, 1978.

[318] Iraphne Childs. Emergence of new volunteerism: Increasing community resilience to natural disasters in Japan. *The Phoenix of Natural Disasters: Community Resilience*, pages 171–180, 2008.

[319] William R Voorhees. New Yorkers respond to the World Trade Center attack: An anatomy of an emergent volunteer organization. *Journal of Contingencies and Crisis Management*, 16(1):3–13, 2008.

[320] Adam J Liska, Tyler R White, Eric R Holley, and Robert J Oglesby. Nuclear weapons in a changing climate: Probability, increasing risks, and perception. *Environment: Science and Policy for Sustainable Development*, 59(4):22–33, 2017.

[321] Xiaoyi Yuan, Ross J Schuchard, and Andrew T Crooks. Examining emergent communities and social bots within the polarized online vaccination debate in twitter. *Social Media + Society*, 5(3):2056305119865465, 2019.

[322] Louise Ryan, Amanda Klekowski Von Koppenfels, and Jon Mulholland. 'the distance between us': A comparative examination of the technical, spatial and temporal dimensions of the transnational social relationships of highly skilled mmgrants. *Global Networks*, 15(2):198–216, 2015.

[323] Vedran Sekara, Arkadiusz Stopczynski, and Sune Lehmann. Fundamental structures of dynamic social networks. *Proceedings of the National Academy of Sciences*, 113(36):9977–9982, 2016.

[324] Petter Holme and Jari Saramäki. Temporal networks. *Physics Reports*, 519(3):97–125, 2012.

[325] Marc Barthélemy. Spatial networks. *Physics Reports*, 499(1-3):1–101, 2011.

[326] Eric C Jones and AJ Faas. *Social Network Analysis of Disaster Response, Recovery, and Adaptation*. Butterworth-Heinemann, 2016.

[327] A Heppenstall, N Malleson, and AT Crooks. Space, the final frontier": How good are agent-based models at simulating individuals and space in cities? systems, 2016, no. 4 (1) 9.

[328] Francesco Orsi and Davide Geneletti. Transportation as a protected area management tool: An agent-based model to assess the effect of travel mode choices on hiker movements. *Computers, Environment and Urban Systems*, 60:12–22, 2016.

[329] Andrew Rumbach, Carrie Makarewicz, and Jeremy Németh. The importance of place in early disaster recovery: a case study of the 2013 Colorado floods. *Journal of Environmental Planning and Management*, 59(11):2045–2063, 2016.

[330] Yaping Ma, Lihua Li, Hui Zhang, and Tao Chen. Experimental study on small group behavior and crowd dynamics in a tall office building evacuation. *Physica A: Statistical Mechanics and its Applications*, 473:488–500, 2017.

[331] Michael J Widener, Mark W Horner, and Sara S Metcalf. Simulating the effects of social networks on a population's hurricane evacuation participation. *Journal of geographical systems*, 15(2):193–209, 2013.

[332] Alison J Heppenstall, Nick S Malleson, Steve J Carver, Duncan J Quincey, Vern R Manville, et al. Modelling individual evacuation decisions during natural disasters: A case study of volcanic crisis in Merapi, Indonesia. *Geosciences*, 8(6):196, 2018.

[333] Bryan Lewis, Samarth Swarup, Keith Bisset, Stephen Eubank, Madhav Marathe, and Chris Barrett. A simulation environment for the dynamic evaluation of disaster preparedness policies and interventions. *Journal of Public Health Management and Practice: JPHMP*, 19(0 2):S42, 2013.

[334] Sean Luke, Robert Simon, Andrew Crooks, Haoliang Wang, Ermo Wei, David Freelan, Carmine Spagnuolo, Vittorio Scarano, Gennaro Cordasco, and Claudio Cioffi-Revilla. The MASON simulation toolkit: Past, present, and future. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation*, pages 75–86. Springer, 2018.

[335] Michael J North, Nicholson T Collier, Jonathan Ozik, Eric R Tatara, Charles M Macal, Mark Bragen, and Pam Sydelko. Complex adaptive systems modeling with Repast Simphony. *Complex Adaptive Systems Modeling*, 1(1):3, 2013.

[336] Nicholson Collier and Michael North. Parallel agent-based simulation with Repast for high performance computing. *Simulation*, 89(10):1215–1235, 2013.

[337] Federal Emergency Management Agency. Typed resource definitions-search and rescue resources. Working Paper 508-8, Federal Emergency Management Agency, Hyattsville, MD, 2005.

[338] John G Kemeny. *Report of the President's Commission on the Accident at Three Mile Island: The Need for Change: the Legacy of TMI.*, volume 41. The Commission, 1979.

[339] NA Beresford, S Fesenko, A Konoplev, L Skuterud, JT Smith, and G Voigt. Thirty years after the Chernobyl accident: What lessons have we learnt? *Journal of environmental radioactivity*, 157:77–89, 2016.

[340] Robert Axtell and Joshua Epstein. Agent-based modelling: Understanding our creations. *The Bulletin of the Santa Fe Institute*, 9:28–32, 1994.

[341] Sarah Wise, Andrew Crooks, and Michael Batty. Transportation in agent-based urban modelling. In Mohammad-Reza Namazi-Rad, Lin Padgham, Pascal Perez, Kai Nagel, and Ana Bazzan, editors, *Agent Based Modelling of Urban Systems*, Lecture Notes in Computer Science, pages 129–148, Cham, 2017. Springer International Publishing.

[342] William G Kennedy and Jeffrey K Bassett. Implementing a "fast and frugal" cognitive model within a computational social simulation. In *Proceedings of the Second Annual Meeting of the Computational Social Science Society of the Americas, http://citeseerx. ist. psu. edu/viewdoc/download*, 2011.

[343] Joshua M Epstein. Why model? *Journal of Artificial Societies and Social Simulation*, 11(4):12, 2008.

[344] Andrew Crooks, Christian Castle, and Michael Batty. Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems*, 32(6):417–430, 2008.

[345] Hazel R Parry and Mike Bithell. Large scale agent-based modelling: A review and guidelines for model scaling. In Alison Heppenstall, Andrew Crooks, Linda M. See, and Michael Batty, editors, *Agent-based models of geographical systems*, pages 271–308. Springer, New York, NY, 2012.

[346] Matteo D'Auria, Eric O Scott, Rajdeep Singh Lather, Javier Hilty, and Sean Luke. Distributed, automated calibration of agent-based model parameters and agent behaviors. In *AAMAS*, pages 1828–1830, 2020.

[347] Nicholson Collier and Michael North. Repast HPC: a platform for large-scale agent-based modeling. *Large-Scale Computing*, 10:81–109, 2012.

[348] James M Kendra and Tricia Wachtendorf. *American Dunkirk: The waterborne evacuation of Manhattan on 9/11*. Temple University Press, 2016.

[349] Kay W Axhausen, Andreas Horni, and Kai Nagel. *The multi-agent transport simulation MATSim*. Ubiquity Press, 2016.

A Listing of Research Outputs

The research output included academic advancement of the participants and the production of literature related to the project. The Principal Investigator was advanced to an associate professor position within the Computational and Data Science Department and the Co-PI was offered a full professorship at a different university. The project could support two graduate students at the same time and supported four graduate students for different amounts of time. Three completed their dissertation work shortly after the project ended. For readers wishing to find out more about this work, you are referred to the following publications:

Journal Papers:

- Burger, A., Oz, T., Kennedy, W.G. and Crooks, A.T. (2019), Computational Social Science of Disasters: Opportunities and Challenges, Future Internet, 11(5): 103. Available at `https://doi.org/10.3390/fi11050103`.
- Colletti, M., Hultquits, C., Kennedy, W.G., Cervone, G. (2017) Validating Safecast data by comparisons to a U. S. Department of Energy Fukushima Prefecture aerial survey. Journal of Environmental Radioactivity 171, pp 9-20.

Reviewed Conference Papers:

- Jiang, N., Burger, A., Crooks, A.T. and Kennedy, W.G. (2020), Integrating Social Networks into Large-scale Urban Simulations for Disaster Responses. Geosim '20: 3rd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation, Seattle, WA.
- Burger, A., Kennedy, W.G., Crooks, A.T., Jiang, N. and Guillen-Piazza, D. (2019), Modeling Social Networks in an Agent-Based Model of a Nuclear Weapon of Mass Destruction Event, The 2019 Computational Social Science Society of Americas Conference, Santa Fe, NM.
- Burger, A., Kennedy, W.G., Crooks, A.T., Jiang, N. and Guillen-Piazza, D. (2019), Modeling Society Reacting to a Nuclear Weapon of Mass Destruction Event, 2019 International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation, Washington DC.
- Lieto, A., Kennedy, W.G., Lebiere, C., Romero, O.J., Taatgen, N., West, R.L. (2018), Higher-level Knowledge, Rational and Social Levels Constraints on the Common Model of the Mind, paper presented at the AAAI Fall Symposium, Arlington, VA (18-20 Oct. 2018) and published in Elsevier's Procedia Computer Science, Vol 145, pp 757-764.
- Orr, M.G., Lebiere, C., Stocco, A., Pirolli, P., Pires, B., Kennedy, W.G. (2018), Multi-Scale Resolution of Cognitive Architectures: A paradigm for Simulating Minds and Society. In Social, Cultural, and Behavioral Modeling, LNCS 10899, Proceedings of the 11th International Conference, SBP-BRiMS 2018, pp3-15. Washington, DC. (*winner of the Best Paper Award*)
- Burger, A., Oz, T., Crooks, A.T. and Kennedy, W.G. (2017), Generation of Realistic Mega-City Populations and Social Networks for Agent-Based Modeling. The Computational Social Science Society of Americas Conference, Santa Fe, NM
- Briggs, T.W., Jr., Kennedy, W.G. (2016) Active Shooter: An Agent-Based Model of Unarmed Resistance. Proceedings of the 2016 Winter Simulation Conference (WSC), 3521-3531. T.M.K. Roeder, P.I. Frazier, R. Szechtman, E.Zhou, T. Huschka, and S.E. Chick, eds. Arlington, VA.

Invited Presentations:

- Kennedy, W.G. and Crooks, A.T. (2017), A Framework for Modeling Society Following a Nuclear WMD Event, Army Science Board Study on Mega Cities, 26th April, Arlington, VA.

Other Presentations/Conferences (Non Reviewed):

- Jiang, N., Burger, A., Guillen-Piazza, D., Crooks, A.T. and Kennedy, W.G. (2020), Modeling a Society's Reaction to a Disaster: An Agent Based Modeling Approach, The Association of American Geographers (AAG) Annual Meeting, 6th-10th, April, Denver, CO.
- Crooks A.T., Burger, A., Jiang, N., Guillen-Piazza, D. and Kennedy W.G. (2018), Decision-making in Times of Crisis: An Agent-Based Model of Disaster, The Association of American Geographers (AAG) Annual Meeting, 3rd-7th, April, Washington, DC.

- Burger, A., Kennedy, W.G. and Crooks, A.T. (2018), Generation of Realistic Routine Behavior in Mega-City Populations, The Computational Social Science Society of Americas Conference, 25th-28th October, Santa Fe, NM.
- Burger, A., Oz, T., Yuan, X., Crooks, A.T. and Kennedy, W.G. (2018), Agent-Based Models for Megacities and Social Networks in Disaster, International Sociological Association World Congress of Sociology, 15th-21st, July, Toronto, Canada.
- Crooks A.T., Burger, A., Yuan, X. and Kennedy W.G. (2018), The Generation and Application of Large Scale Synthetic Populations for Disease Outbreaks and Disasters, The Association of American Geographers (AAG) Annual Meeting, 10th-14th, April, New Orleans, LA.
- Crooks A.T., Kennedy W.G., Burger, A. Oz, T. and Heppenstall, A. (2017), Megacities through the Lens of Computational Social Science, The Association of American Geographers (AAG) Annual Meeting, 5th-9th, April, Boston, MA.

Articles and Interviews about this Work

- Waddell, K. (2017), What Happens If a Nuclear Bomb Goes Off in Manhattan?, The Atlantic, `https://www.theatlantic.com/technology/archive/2017/03/nuclear-bomb-manhattan-simulation/519585/`.
- Subsequently picked up by:
  - Bullen, J. (2017), Scientists to Simulate How 20 MILLION People Would React to Nuclear Bomb in New York City, The Mirror, Available at `http://www.mirror.co.uk/news/world-news/scientists-simulate-how-20-million-10054846`.
  - Rogers, J. (2017), Union City Blues: Scientists Simulate Nuclear Attack on New York, The Express, Available at `http://www.express.co.uk/news/world/781093/New-York-Manhattan-nuclear-attack-scientists-George-Mason`.
  - Waugh, J. (2017), World War 3: Scientists to Simulate Nuclear Bomb Going Off in New York, Metro, Available at: `http://metro.co.uk/2017/03/20/world-war-3-scientists-to-simulate-nuclear-bomb-going-off-in-new-york-6521225/`.
  - Wilkinson, J. (2017), What Happens if New York Gets Nuked? Scientists Simulate 20 MILLION People in Massive Computer System to See How State Would Respond to a Nuclear Attack, Daily Mail, Available at: `http://www.dailymail.co.uk/news/article-4327280/Scientists-simulate-nuclear-attack-New-York.html`.

PhD Dissertations resulting from this Research

This research helped fund three PhD students who all have now successfully defended their dissertations.

- Burger, A. (2020), Disaster Through the Lens of Complex Adaptive Systems: Exploring Emergent Groups Utilizing Agent Based Modeling and Social Networks, PhD Dissertation, George Mason University, Fairfax, VA.
- Oz, T. (2020), Collective Stress in the Digital Age, PhD Dissertation, George Mason University, Fairfax, VA.
- Yuan, X. (2020), Geo-textual Data Analytics Exploring Places and their Connections, PhD Dissertation, George Mason University, Fairfax, VA.

B Code Base

All the code developed for this project can be found at:

- For the code on the agent-based model [15] see: `https://github.com/AnnettaGB/NWMDinNYC`
- The source data and results data from the synthetic population generation are shared at `https://github.com/njiang8/Create_Synthetic_Population` and `https://osf.io/3vsaj/`).

C Sample of Research Outputs

Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S.E. Chick, eds.

ACTIVE SHOOTER: AN AGENT-BASED MODEL OF UNARMED RESISTANCE

Thomas W. Briggs

Department of Computational and Data Sciences
George Mason University
4400 University Drive
Fairfax, VA 20030, USA

William G. Kennedy

Krasnow Institute for Advanced Study
George Mason University
4400 University Drive
Fairfax, VA 20030, USA

ABSTRACT

Mass shootings unfold quickly and are rarely foreseen by victims. Increasingly, training is provided to increase chances of surviving active shooter scenarios, usually emphasizing "Run, Hide, Fight." Evidence from prior mass shootings suggests that casualties may be limited should the shooter encounter unarmed resistance prior to the arrival of law enforcement officers (LEOs). An agent-based model (ABM) explored the potential for limiting casualties should a small proportion of potential victims swarm a gunman, as occurred on a train from Amsterdam to Paris in 2015. Results suggest that even with a miniscule probability of overcoming a shooter, fighters may save lives but put themselves at increased risk. While not intended to prescribe a course of action, the model suggests the potential for a reduction in casualties in active shooter scenarios.

1 INTRODUCTION

Mass shootings unfold quickly and are rarely foreseen by victims. Mass shootings have occurred at a variety of locations including military installations and government buildings, public spaces including nightclubs (Orlando, FL), movie theaters (Aurora, CO), shopping malls, workplaces, religious facilities, and educational campuses (Littleton, CO; Blacksburg, VA; Newtown, CT).

The difficulty of preventing mass shootings has led to increased active shooter training. Law enforcement agencies have revised response tactics for active shooter situations following the Columbine high school shooting (Police Executive Research Forum 2014) and employers and public safety organizations have developed protocols including "Run, Hide, Fight" or "Avoid, Deny, Defend" for individuals in an active shooter situation. The implementation of these tiered strategies may benefit the individual who successfully flees or hides, but may subsequently put someone else at greater risk (e.g., by monopolizing a secure hiding spot) and may not substantively reduce the overall number of casualties in a mass shooting scenario.

In 2015, a presumed mass shooter on a Thalys train from Amsterdam to Paris was subdued by the rapid action of several men who engaged in hand-to-hand combat with the gunman. Two of the men were seriously injured—one shot, one severely cut—but both survived. No one was killed and the gunman was captured, despite being armed with an AKM rifle, a Luger pistol, and a box cutter.

Researching mass shootings presents obvious methodological challenges: conducting an experiment in which participants believe they are actually facing potential death from an active shooter is ethically intractable and could lead to actual harm (e.g., attempts to subdue shooter). While tactical drills such as those used by LEOs and military can simulate the mechanics of facing an active shooter, the explicit knowledge that one is in a simulation likely dampens neurophysiological responses and would hopefully preclude participants from improvising a lethal response against the individual acting as the mass shooter. Examining historical mass shootings is a valuable research technique, but there are known limitations on

978-1-5090-4486-3/16/$31.00 ©2016 IEEE

eyewitness accounts and certainly no possibility of altering the historical scenario in an attempt to influence outcomes. Agent-based modeling (ABM) is a logical choice to explore the potential impact of intended targets' behavior when encountering an active shooter since it harms no human subjects, can explicitly encapsulate behavioral rules, and offers the possibility of running the model under altered conditions to investigate outcomes. The present research uses ABM to investigate the degree to which the rapid action of a few individuals who physically confront a shooter might potentially limit the casualties in mass shooting scenarios.

2 BACKGROUND

2.1 Active Shooters and Mass Shootings

From 2000 to 2013, the U.S. FBI reported 160 active shooter incidents in which 486 were killed and 557 wounded, excluding the shooters (Blair and Schweit 2013). Any attempt to tabulate shooting incidents is ultimately definition-dependent and definitions are debated. The FBI defines an active shooter as "an individual actively engaged in killing or attempting to kill people in a populated area," noting that "implicit in this definition is the subject's criminal actions [must] involve the use of firearms." The definition of a *mass shooting* is based on that of mass murder, defined as four or more individuals killed during the same incident. An active shooter scenario may or may not qualify as a mass shooting, then, as fatalities depend on both the lethality of victims' wounds and relatively distal variables like the availability of advanced trauma care following the shooting. A potential drawback of using the mass murder definition is that it relies on quantified fatalities, so an active shooter incident in which many people are shot but fewer than four perish does not meet the threshold of mass shooting.

The FBI notes that both law enforcement and citizens have the potential to affect the outcome of an active shooter event (Blair and Schweit 2013). In the 104 active shooter incidents from 2000 to 2012, the shooter was stopped by victims in 17 incidents, by police in 32 incidents, and in 55 incidents, stopped on his own accord, committing suicide in 44 cases, surrendering in 6 cases, and leaving in 5 cases. (Blair, Martaindale, and Nichols 2014). Of the 17 incidents in which victims stopped the gunman, in 3 cases the active shooter was shot by armed victims.

2.2 Prior Agent-Based Models

Hayes and Hayes (2014) created several ABMs of mass shooting scenarios to test specific provisions of Senator Dianne Feinstein's proposed legislation to limit certain specific types of firearms. A model of the Aurora, CO movie theater shooting in 2012 and a generalized indoor model found that only a reduction in a firearm's rate of fire would have likely reduced the number of casualties in the Aurora shooting (Hayes and Hayes 2014). A school shooting model exploring the presence of armed school law enforcement officers (LEOs) and staff carrying concealed firearms suggested that either intervention would likely decrease response time in confronting the shooter and reduce casualties, though the model assumes that the shooter would be instantly neutralized upon entering a room in which a single armed individual is present (Anklam et al. 2015). This assumption may be overly optimistic in light of studies of shooting performance of law enforcement officers (Lewinski et al. 2015; Vickers and Lewinski 2012). Anklam et al. (2015) conclude that reducing the time-to-intercept of an active shooter will likely reduce casualties, but their school shooting model considered intercept possible only by armed individuals, with no distinction between LEOs and civilians.

No ABMs could be located that examined the potential role of unarmed resistance in an active shooter scenario.

3 METHOD

3.1 Agent-Based Model

Developed using NetLogo (Wilensky 1999), model implementation followed Wilensky and Rand's (2015) ABM design principle: start simply and build toward the question of interest. A crowd of agents is distributed on an open landscape (e.g., a large outdoor concert or rally) with no possible cover or concealment. Agents are unaware that a shooting is about to occur. A randomly-located shooter begins firing on the closest targets. Once the shooting begins, most agents flee from the shooter at their running speed. On reaching the outer perimeter of the simulation, fleeing agents are presumed safe and can no longer be targeted. A small proportion of agents, if close enough, try to tackle and subdue the shooter. The simulation ends if the shooter is subdued, when the shooter hits every possible target, and/or all targets have escaped. For parsimony, a fired shot can hit only one victim, no victim can be hit twice, and no lethality determination is made due to the many factors affecting outcomes of gunshot wounds.

3.2 Agents

Population. The agents in the current model possess a normally-distributed running speed sourced from the Hayes and Hayes (2014) ABM of active shooter scenarios: the distribution has a mean of 3.9 m/s and standard deviation 2.7 m/s. Agents are also assumed to have a cognitive delay required to recognize and process that a shooting has begun, after which they immediately run away from the shooter. While actual cognitive delay would likely differ for each individual, in the current model it is a constant such that the entire population simultaneously realizes that a shooting has begun. This parameter is user-adjustable and can be disabled if desired (i.e., set to 0 seconds).

Fighters vs Fleers. Some proportion of the agents are fighters. This proportion is set by the user and is expected to be very small relative to the population. Instead of fleeing from the shooter these individuals, like the individuals who subdued the gunman on the Thalys train, will attempt to tackle the shooter if/when they are close enough. Whether these individuals have military or law enforcement training or are simply extreme altruists is an open question beyond the scope of the current effort. The model simply assumes that some number of people – however few – might choose to endanger themselves in response to an active shooter. In this model, fighters run toward the active shooter, putting themselves at greater risk by closing the distance and increasing the likelihood of being hit by a consequently more accurate shot. The user sets the probability with which a fighter struggling with a shooter is likely to overcome the shooter on each tick. This is a global parameter: if the user gives a fighter a 1% chance of overcoming the shooter and three fighters struggle with a shooter, each fighter has precisely a 1% chance per tick of overcoming the shooter. In other words, there is currently no additional advantage when multiple fighters conduct a swarm attack and struggle with the shooter simultaneously, though this will be explored in future model extensions.

Shooter. User-adjustable parameters can be set to account for armament (magazine capacity and firearm effective range) and shooter ability (accuracy and field of view for targeting). For parsimony a shooter always targets the closest agent in (1) firearm effective range and (2) field of view, and will fire one round per second (tick). Firearm rate of fire is frequently debated. For parsimony, one round per second is fired in the current model. This rate of fire likely overestimates most shooters' ability to accurately target and fire but could represent indiscriminate firing into a crowd.

Whether or not the target is hit is probabilistic and depends on three factors: distance between shooter and target, the user-adjustable accuracy parameter, and the firearm's effective range. Firearm effective range is implemented in the current model as the range at which a 100% accurate shooter hits a human-sized target 50% of the time. This parameter allows users to approximate the type of firearm employed: most shooters will be accurate at greater distances with rifles than pistols and range can be set accordingly. The user-adjustable accuracy parameter allows the user to account for the human component of shooting accuracy. At 1.0, the shooter is 100% accurate at point-blank range and 50% accurate at the firearm's effective range. In actual firefights involving LEOs, many rounds miss their intended targets even at relatively close ranges, so a 1.0 accuracy setting is likely highly unrealistic, but is nevertheless available to the user (Lewinski et al. 2015). If a fired round misses the intended target, it continues traveling and may hit another agent if that agent is in the round's trajectory. In dense crowds, therefore, even an inaccurate shooter is capable of inflicting substantial casualties. The shooter continues to target and fire on each tick, either until subdued by fighters or until all potential targets have reached the perimeter of the landscape. In the current version of the model, the shooter does not move to pursue targets and remains in a single location for the duration of the simulation.

3.3 Initial Setup

The user adjusts the population size such that the desired physical crowd density is achieved. Density is important because it affects (1) the number of possible targets in the shooter's range and vision and (2) the likelihood that a shot that misses the intended target will wound another agent in the round's trajectory. The user also sets model parameters described above.

3.4 Model Action

On the first model tick, the shooter "activates," targeting the nearest individual in his field of view and firing. (To conceptualize field of view, imagine sweeping a wide-beam flashlight from side to side – everything in the cone made by the flashlight beam is in the field of view.) On each subsequent tick, the shooter takes the same action: target, then fire. When the shooter targets, he turns to directly face the targeted individual, changing his field of view. A shooter cannot see behind himself and can only see what is in his field of view. After the shooting begins and the cognitive delay time has elapsed, most agents will begin fleeing from the shooter. Fighters present will run toward the shooter and try to tackle him if close enough to reach in less than one second, a distance that varies depending on a fighter's unique running speed.

When a fighter reaches a shooter, a struggle begins and the shooter shifts his attention to the fighter. In reality, the likelihood of either a fighter overcoming a shooter or a shooter overcoming a fighter will depend on a substantial number of variables such as prior combat training, physical strength, weaponry, and assistance from others. As each of these can be vigorously debated, the user sets probabilities of success for both the shooter and the fighter. Probabilities are implemented on a per-tick basis. Calibration data for these probabilities could not be located, so it is suggested that the shooter should have a very high probability of overcoming the fighter (perhaps because the shooter also carries weapons intended for close-range combat, whether pistols or bladed weapons) and the fighter should have a low probability of overcoming the fighter due to the relative disadvantage in armament. Fighters who fail become victims (i.e., are wounded and incapacitated for the remainder of the simulation).

3.5 Model Output

In addition to a visual view of the unfolding scenario, the model tracks the number of rounds fired, the number of rounds that strike individuals, and the number of fighters struggling with a shooter at each tick.

3.6 Model Calibration

Parameter sweeps using NetLogo BehaviorSpace examined model sensitivity and differences in outputs. The parameters were varied as indicated in Table 1 and results are discussed in the next section.

Table 1: Model parameters with bold values indicating final stable model defaults.

Parameter	Values	Notes
population	500 1000 **5000** 7500	Agent population
%-who-fight	0.001 0.003 **0.005** 0.010	Percentage of agent population who are “fighters” rather than “fleers”
chance-of-overcoming-shooter	**0.01** 0.05 0.10	Per-tick probability of a fighter overcoming the shooter in a hand-to-hand struggle
shooters	**1**	Number of shooters
shooter-magazine-capacity	**10**	Rounds that can be fired before a magazine reload (shooters have unlimited magazines)
firearm-effective-range	30m **50m** 70m	Range at which a 100% accurate shooter will hit target 50% of the time; used in hit probability
shot-accuracy	0.5 **0.8** 1.0	Human factor in accuracy; combines with firearm-effective-range to determine hit probability of each shot
field-of-view	**180 degrees**	shooter’s field of view (see section 3.2)
shooter-chance-of-overcoming-fighter	**0.5**	Per-tick probability of shooter overcoming a fighter in a hand-to-hand struggle

3.7 Verification and Validation

Verification and validation are particularly challenging for the current model and topic. Though mass shootings occur, there is a dearth of detailed publicly available data and a large number of variables and unknowns that affect ultimate outcomes. Hayes and Hayes (2014) validated their model of the 2012 Aurora, CO movie theater shooting by calibrating the model such that, on average, a model run approximated the same number of casualties that actually occurred during the shooting. This is a laudable strategy, but one that is not easily employed in a generalized active shooter model. A shooter’s targeting strategy, weaponry, and accuracy are likely to have the greatest impact on casualties, followed by the behavior of intended victims (e.g., do intended victims make themselves easier or more difficult targets?). As mentioned in the introduction, conducting an experiment to test victim response to an active shooter is not practicable; it would be ethically impossible to create a true life-or-death situation in which individuals would respond with potentially lethal force. This model is inspired by the events on the Thalys train and also what is believed to have occurred on United Flight 93 on September 11, 2001, but these situations are extremely rare and ought not be considered representative. Each mass shooting is different, and caution should be employed making generalizations from one mass shooting to another. Subject-matter experts are invited to criticize the assumptions of the current model and suggestions are welcomed. Other modelers are encouraged to replicate or extend the current model.

The current ABM was subject to verification during the process of model construction using unit tests written into model code to ensure that a particular procedure is behaving as intended and that code was adequately debugged (Wilensky and Rand 2015).

Validation requires at least some correspondence between the model's behavior and the behavior of the target system (Gilbert and Troitzsch 2005). At the present stage of this research effort, invoking the oft-cited quote from George Box may prove helpful: "All models are wrong, but some are useful." The validation question, then, rests on whether or not the current model can be useful as platform for exploring the role of intended victims of an active shooter.

4 RESULTS

4.1 Overall

The current model suggests unarmed resistance to an active shooter may reduce overall casualties in an active shooter incident.

With default model parameter settings (as shown in Table 1), the shooter is subdued in 67 percent of experimental model runs and overall casualties are mean 30. This is a substantial reduction in casualties from the no fighter control condition in which mean casualties are 57. In the remaining 33 percent of model runs in which the shooter is not subdued, mean casualties are increased only slightly to 63, with a greater share of fighters among the casualties as a result of putting themselves in harm's way. Figure 1 plots casualties by simulation end time in 500 model runs in both the control and experimental conditions. The number of casualties sustained in each incident is directly related to time since the shooter has a sustained rate of fire of one round per second. In the experimental runs in which the shooter is subdued, mean time elapsed is 100 seconds, far less than in the control condition in which the simulation typically concludes at 255 seconds after which all remaining victims have escaped the perimeter.

Importantly, default model parameters were selected to be as conservative as possible, and the model and code are available upon request from the author for any user who wishes to set the parameters less or even more conservatively. In the absence of empirical data sources to calibrate the model, users are encouraged to consult relevant subject matter experts in choosing parameter settings.

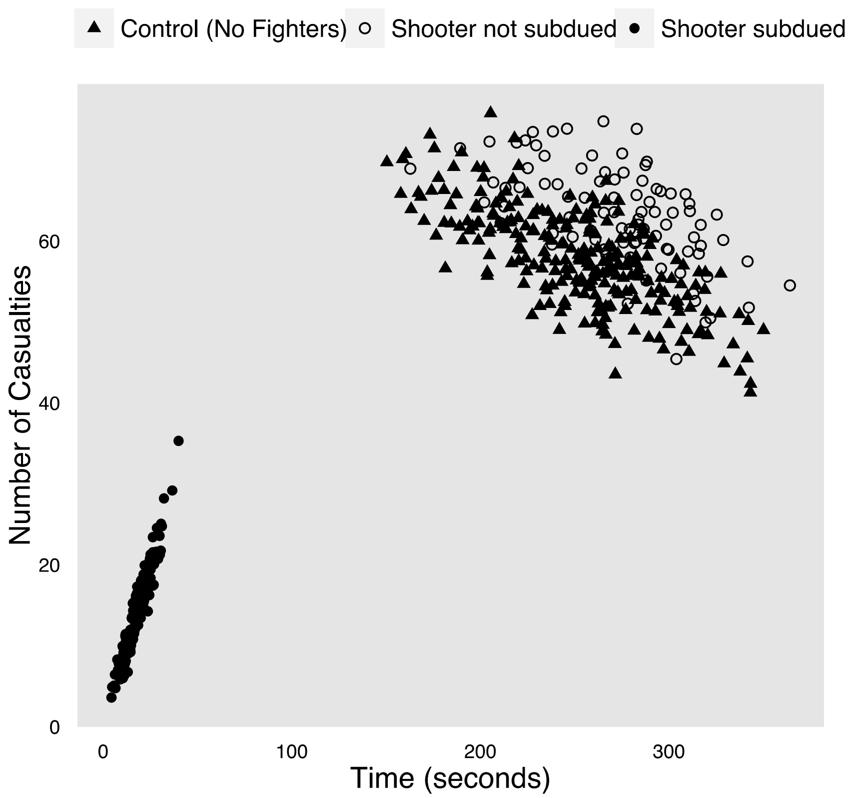


Figure 1: Casualties by simulation end time in control and experimental conditions (plot displays a random sample of 500 of 2000 model runs for ease of visibility).

4.2 Flee vs. Fight Proportion

Unsurprisingly, the greater the proportion of fighters in the population, the more likely the shooter will be subdued. If too few fight, there is little chance of overcoming the shooter. Varying the proportion of the population that fights changes the likelihood of overcoming the shooter. If only 0.1 percent fight, virtually no model runs result in subduing the shooter; if 0.4 percent fight, the shooter is subdued in about half of model runs, and if between 0.8 and 1 percent fight, the shooter is subdued in nearly all model runs.

4.3 Other Parameters

The current effort did not test rate of fire, since the Hayes and Hayes (2014) ABM demonstrated that reducing rate of fire would likely reduce casualties in an active shooter scenario. No appreciable difference in outcomes occurred by varying magazine capacity, since reload times of ~1 second (note that such a rapid reload time is possible by using a technique known as a “speed reload”) do not substantially reduce overall rounds fired. (Reloads may, however, present ideal opportunities to engage a shooter, though this was not tested with the current model.)

Firearm effective range was varied between 30 m, 50 m, and 70 m to explore potential differences between the use of pistols and rifles, the latter being more accurate at greater distances. Despite extensive media coverage of the use of semiautomatic rifles in mass shootings, the majority of mass shooters to date have used pistols. In runs in which the shooter is subdued, casualties are only slightly increased with the use of more accurate firearms since the majority of casualties occur initially at close range. When the shooter is not subdued and can continue firing on fleeing victims, casualties increase almost linearly, as might be expected.

4.4 Qualitative Observations

The greatest concentration of casualties will occur at the beginning of the simulation since victims only begin fleeing after realizing what is happening. Shooters will almost always possess an informational advantage over intended victims because only the shooter knows when and where he will open fire and his targeting strategy (if anything other than random or based on proximity).

Viewing the model visualization in real time illustrates that individuals who attempt to attack the shooter from a great distance are at a serious tactical disadvantage, particularly if they have a slow approach speed. By reducing the distance between themselves and the shooter, they increase the likelihood that they will be shot. This may suggest pursuing an avoid (run) or deny (hide) strategy unless structural features of the environment can shield would-be fighters from the shooter's sight and fire (e.g., rooms, corners, or other cover or concealment) and facilitate getting close enough for hand-to-hand combat with the shooter. Another important interpretation of this result is that LEO entry teams, moving slowly toward the shooter's location, would potentially be at great risk should a shooter stage an ambush.

5 DISCUSSION

5.1 Fighters will likely save lives but put themselves at increased risk

Attention is a scarce commodity, and every second that an active shooter struggles with a fighter is a second that he is not able to effectively target and fire upon another victim. The "Run" and "Hide" prescriptions are intended to occupy the shooter's time and attention: time spent by a shooter searching for available victims is time for law enforcement to arrive on the scene, form an entry team, and sweep for the shooter. Unfortunately, as suggested by incident reports for the Virginia Tech and Sandy Hook shootings, active shooters encountering harder targets like barricaded rooms will simply move on to softer targets. Further, when potential victims "hide" by huddling together in a room corner with little or no cover or concealment – like most victims at Sandy Hook Elementary – it may be even easier for a shooter to inflict maximal casualties with fewer rounds fired.

It is impossible to calculate precise odds of becoming a casualty in an active shooter scenario, regardless of whether an individual chooses to run, hide, or fight. However, it is the case that there is at least a nonzero probability of successfully overcoming a shooter, as demonstrated on the Thalys train and in 17 of the 104 cases studied by the FBI (Blair, Martaindale, and Nichols 2014). The present model suggests that even with a relatively low probability of success and no combined advantage from a coordinated group attack, overall casualties might be reduced if a small number individuals close enough to fight the shooter fight rather than flee.

5.2 Cautions and Guidelines for Interpretation

An important caveat of this work is that it is not intended to prescribe a course of action for individuals to specifically put themselves in harm's way. Most active shooter training emphasizes "Run, Hide, Fight" or "Avoid, Deny, Defend," and emphasis is placed on the order of those options. Trainees are told to "run if you can," "hide if you must," "fight if you have to," with the acknowledgement that each individual must make his or her own decision and there are no guaranteed outcomes.

However, active shooter training also contradicts prior training for hostage situations and armed robberies, which trained compliance with gunmen's demands to *prevent* violence. In mass shooting scenarios, calm cooperation may result in being shot.

The suggestion that untrained civilians engage armed attackers must be considered carefully. When shooters have been subdued in prior incidents, individuals with some form of combat training—either law enforcement or military—are typically involved. Two of the three Americans who subdued the gunman on the Thalys train had military training and one had just returned from deployment in Afghanistan. But even trained, armed LEOs responding to an active shooter can become victims, as was demonstrated when a shooter armed with a semiautomatic rifle attacked a Planned Parenthood facility in Colorado in 2015. Six of the responding LEOs were wounded and one, Officer Garrett Swasey, was killed. Whether one or more average citizens without training might subdue a gunman requires additional research. Though the principal and school psychologist at Sandy Hook were both killed by gunfire, the shooter was very underweight at only 112 pounds (50.8 kg) despite being six feet (1.83 m) tall. It is certainly *possible* that he could have been subdued in a hand-to-hand struggle had the staff been close enough to physically reach and engage him.

5.3 Limitations

Numerous limitations exist in this preliminary modeling effort.

The model does not give any combined advantage to multiple fighters who swarm attack a shooter. This likely underestimates the probability of success should multiple fighters engage the shooter as occurred on the Thalys train. One fighter might, for example, attempt to control the direction of a shooter's weapon while another fighter attempts to take the shooter to the ground by tackling the shooter's legs. (This type of swarm attack is exactly the technique that is typically emphasized in the "Fight" component of many active shooter trainings for civilians.)

The current model is low-fidelity in a number of respects. Both ballistics and hand-to-hand combat are modeled as probabilities. Additionally, agents, whether fleeing or fighting, do not communicate or interact with one another, nor do they have any cover or concealment in the open environment. Crowd behavior is not accounted for in the current model: faster agents simply run through slower agents.

This model does not address the cognitive and behavioral processes underlying heroic acts or acts of extreme altruism; the assumption is that at least some individuals are capable of such acts and will resist when faced with an imminent threat as in the incident on the Thalys train. The user is free to set the percentage of individuals likely to engage a gunman rather than flee.

Importantly, the current model does not represent ballistics with high fidelity. However, the model approximates shot accuracy and permits rounds to continue to travel beyond their intended target, possibly striking another person in the round's trajectory. Fired rounds do not discriminate, and physics ultimately determines when and where rounds will stop. (This is also relevant when considering armed response to an active shooter: trained LEOs may hit their intended targets 50 percent of the time, so an important aspect of modeling mass shooter scenarios is the potential collateral damage of various potential responses, including casualties by friendly fire.)

The current model also does not represent hand-to-hand combat with any fidelity. Any struggle will depend on the skills of the individuals involved and any weaponry available, either the shooter's or improvised by fighters.

A limitation of the current model is the lack of specific forensic information from prior mass shootings with which to validate the model. Presumably, such information exists but is not accessible by the general public. For example, precisely how close were the Americans to the gunman on the Thalys train in 2015? How close were the principal and school psychologist to the gunman at Sandy Hook Elementary when they confronted him in the hallway and were killed in the 2012 shooting? These are important data for model validation, especially for a higher-fidelity simulation.

5.4 Future Research

The current model serves as a starting point for future research efforts, including testing additional parameter combinations, variables, scenarios, and assumptions.

The notion of rapid collective action should be explored. Specifically, agents could be given the ability to communicate—even rapidly, as reportedly happened on the Thalys train—in making the decision to jointly attack a shooter. It may also be the case that there are only an infinitesimally small number of individuals who would attack an active shooter, but that others would join once that individual begins the struggle. In this sense, agents could be further divided into individuals who would attack, regardless, and a greater number of individuals who attack only when others do, invoking a threshold like Epstein's (2002) model of civil violence or Granovetter's (1978) model of collective behavior.

ACKNOWLEDGMENTS

The authors are grateful to Kenneth Comer for inspiring this modeling effort through his seminar on military agent-based modeling at George Mason University in 2015, to Dale Brearcliffe for comments and discussions on drafts of this paper, and to three anonymous peer reviewers for their thoughtful feedback and helpful suggestions.

REFERENCES

Anklam, C., A. Kirby, F. Sharevski, and J. E. Dietz. 2015. "Mitigating Active Shooter Impact: Analysis for Policy Options Based on Agent/computer-Based Modeling." *Journal of Emergency Management* 13 (3): 201–16. doi:10.5055/jem.2015.0234.

Blair, J. P., M. Hunter Martaindale, and Terry Nichols. 2014. "Active Shooter Events from 2000 to 2012." *FBI Law Enforcement Bulletin*. https://leb.fbi.gov/2014/january/active-shooter-events-from-2000-to-2012.

Blair, J. P, and K. W. Schweit. 2013. "A Study of Active Shooter Incidents, 2000-2013." https://hazdoc.colorado.edu/handle/10590/2712.

Epstein, J. M. 2002. "Modeling Civil Violence: An Agent-Based Computational Approach." *Proceedings of the National Academy of Sciences* 99 (Supplement 3): 7243–50. doi:10.1073/pnas.092080199.

Gilbert, G. Nigel, and K. G. Troitzsch. 2005. *Simulation for the Social Scientist*. 2nd ed. Maidenhead, England ; New York, NY: Open University Press.

Granovetter, M.. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83 (6): 1420–43.

Hayes, R., and R. Hayes. 2014. "Agent-Based Simulation of Mass Shootings: Determining How to Limit the Scale of a Tragedy." *Journal of Artificial Societies and Social Simulation* 17 (2): 5.

Lewinski, W. J., R. Avery, J. Dysterheft, N. D. Dicks, and J. Bushey. 2015. "The Real Risks during Deadly Police Shootouts Accuracy of the Naïve Shooter." *International Journal of Police Science & Management*, 117–27.

Police Executive Research Forum. 2014. *The Police Response to Active Shooter Incidents.*

Vickers, J. N., and W. Lewinski. 2012. "Performing under Pressure: Gaze Control, Decision Making and Shooting Performance of Elite and Rookie Police Officers." *Human Movement Science* 31 (1): 101–17. doi:10.1016/j.humov.2011.04.004.

Wilensky, U.. 1999. *NetLogo*. Center for Connected Learning and Computer-Based Modeling. Evanston, IL: Northwestern University. http://ccl.northwestern.edu/netlogo.

Wilensky, U., and W. Rand. 2015. *Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. Cambridge, Massachusetts: The MIT Press.

AUTHOR BIOGRAPHIES

THOMAS W. BRIGGS is a Ph.D. student in Computational Social Science at George Mason University. He holds a M.P.S. in Industrial/Organizational Psychology from the University of Maryland, Baltimore County. His email address is wsc16tbriggs@gmu.edu.

WILLIAM G. KENNEDY is Research Assistant Professor at the Center for Social Complexity in the Krasnow Institute for Advanced Study at George Mason University. He holds a Ph.D. in information technology from George Mason University. He is a retired Navy Captain with 30 years of service in submarines and the Naval Reserve and also a retired civil servant with 25 years of service with the U.S. Nuclear Regulatory Commission and the U.S. Department of Energy. His research interests include: integrating computational cognitive modeling and computational social science; cognitive plausibility; and cognitive robotics. He is an IEEE Life Member. His email address is wsc16wkennedy@gmu.edu.

Journal of Environmental Radioactivity 171 (2017) 9–20

Contents lists available at ScienceDirect

Journal of Environmental Radioactivity

journal homepage: www.elsevier.com/locate/jenvrad

Validating Safecast data by comparisons to a U. S. Department of Energy Fukushima Prefecture aerial survey

Mark Coletti [a, b, *], Carolynne Hultquist [a], William G. Kennedy [c], Guido Cervone [a, d]

[a] Dept. of Geography and Institute for CyberScience, Geoinformatics and Earth Observation Laboratory, The Pennsylvania State University, University Park, PA, United States
[b] Computational Sciences and Engineering Division, Geographic Information Science and Technology Group, Oak Ridge National Laboratory, Oak Ridge, TN, United States
[c] Center for Social Complexity, Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA, United States
[d] Research Application Laboratory, National Center for Atmospheric Research, Boulder, CO, United States

ARTICLE INFO

Article history:
Received 1 July 2016
Received in revised form
7 December 2016
Accepted 7 January 2017

Keywords:
Safecast
Volunteered geographic information
Fukushima Daiichi
Data validation

ABSTRACT

Safecast is a volunteered geographic information (VGI) project where the lay public uses hand-held sensors to collect radiation measurements that are then made freely available under the Creative Commons CC0 license. However, Safecast data fidelity is uncertain given the sensor kits are hand assembled with various levels of technical proficiency, and the sensors may not be properly deployed. Our objective was to validate Safecast data by comparing Safecast data with authoritative data collected by the U. S. Department of Energy (DOE) and the U. S. National Nuclear Security Administration (NNSA) gathered in the Fukushima Prefecture shortly after the Daiichi nuclear power plant catastrophe. We found that the two data sets were highly correlated, though the DOE/NNSA observations were generally higher than the Safecast measurements. We concluded that this high correlation alone makes Safecast a viable data source for detecting and monitoring radiation.

1. Introduction

Volunteered Geographic Information (VGI) provides alternatives to government and corporate sponsored sources for determining the impact of natural or man-made disasters via crowd-sourced measurements (Goodchild, 2007). Ordinary citizens persons with smartphones or handheld sensors can make observations of disaster related phenomena that can supplement data gathered from traditional remotely sensed sources and ground-based equipment. However, sensing platforms are expensive to deploy, operate, and maintain, whereas VGI equipment is typically owned and operated by volunteers for comparatively little cost. Also, these citizen-based observations can cover areas from perspectives difficult to achieve with official sources, and with a very high spatial and temporal resolutions, especially in urban areas.

In fact, while space- and air-borne remote sensing can achieve a very high spatial resolution, in the order of a few centimeters in different parts of the electromagnetic spectrum, and vehicles can be deployed to capture data from the ground, satellites are not always overhead and are limited by atmospheric opacity (clouds and pollution), planes cannot remain airborne indefinitely, and ground vehicles have limited operating ranges and times.

Moreover, individuals intelligently evaluate their surroundings to focus their equipment on interesting scenes, whereas government or corporate managed sensors mechanically scan the environment without consideration to what is being observed, which means that these government and corporate sources may require more post-processing and analysis to mine useful information.

The Safecast VGI project uses "citizens as sensors" (Goodchild, 2007) to produce publicly available collection of radiation levels by time and location. Safecast participants collect these radiation measurements as a public service as well as for awareness of their own radiation exposure, and can be used as a citizen-led early warning system to detect radioactive leaks and hot spots. On March of 2015 there were over 27 million logged observations from around the globe. About 75% of the observations originated in Japan, primarily in Fukushima, surrounding prefectures, and in

* Corresponding author. Dept. of Geography and Institute for CyberScience, Geoinformatics and Earth Observation Laboratory, The Pennsylvania State University, University Park, PA, United States.
E-mail addresses: colettima@ornl.gov (M. Coletti), cgh138@psu.edu (C. Hultquist), wkennedy@gmu.edu (W.G. Kennedy), cervone@psu.edu (G. Cervone).

http://dx.doi.org/10.1016/j.jenvrad.2017.01.005
0265-931X/

major Japanese cities. Up until 2013, virtually all measurements were confined to Japan (Bonner and Brown, 2015).

Unfortunately VGI fidelity can be questionable because of possible operator reporting bias, poor data quality (such as the inclusion of generated test data), and equipment reliability and accuracy (Flanagin and Metzger, 2008). Though the Safecast organization has taken steps to ensure that the hardware is properly tested and calibrated before shipping (Safecast, 2015a), it is possible for the volunteer to make mistakes in assembling the sensor, particularly if they are inexperienced with putting together complicated electronic equipment. Moreover, the user may ignore equipment operating instructions (Safecast, 2015c), which may reduce observation quality. Though Safecast employs moderators to vet newly uploaded data (Brown et al., 2016), it is still theoretically possible for low quality data to be added to the publicly available database.

To address this open issue, we compared Safecast radiation observations with a similar set of observations made by the U. S. Department of Energy jointly with the U. S. National Nuclear Security Administration (DOE/NNSA) (Lyons and Colton, 2012). We found that the two datasets were strongly correlated, but that the DOE/NNSA observations were generally higher than the corresponding Safecast values. Later, we explore possible explanations for these differences. However, given the high correlation between the two datasets, we conclude that the Safecast data has utility for measuring environmental radiation.

1.1. The 2011 Tohoku earthquake and tsunami

On March 11, 2011 at 5:46:24UTC a 9.0 magnitude earthquake occurred 130 km east of Sendai, Honshu, Japan at a depth of approximately 30 km near the Pacific and North American plate subduction zone (U. S. Geological Survey, 2011). Earthquake models showed that the fault moved upwards by 30 m–40 m over a 300 km by 150 km area with effects that were felt as far away as Korea, southeastern Russia, and China (U. S. Geological Survey, 2011). The plate shift was extreme enough to move the Earth's axis 25 cm and speed up its rotation by 1.8 μs per year (Chai, 2011). Moreover, the earthquake slid Honshu, the main island of Japan, 3.6 m to the east, while part of the Oshika Peninsula moved about 5.3 m towards the earthquake's epicenter (Norio et al., 2011).

The tsunami caused by this earthquake affected 20 different Pacific Rim countries with most of the damage occurring in Japan. It is estimated that the tsunami reached a peak of 38 m above mean sea level while penetrating up to 10 km inland (Norio et al., 2011). Over 300,000 buildings, 2000 roads, and 50 bridges were damaged or destroyed. There were also approximately 15,000 casualties, 5300 injured, and 4600 missing people due to the tsnunami (U. S. Geological Survey, 2011). The combined earthquake and tsunami had an estimated initial overall economic impact of up to 183 billion US Dollars (Norio et al., 2011).

1.2. The Fukushima Daiichi nuclear disaster

The 11 nuclear power plants in northeastern Japan automatically shutdown when the earthquake struck (Norio et al., 2011). In spite of these automatic safety procedures, the Fukushima Daiichi power plant suffered a level 7 catastrophic nuclear incident, the highest level on the International Nuclear and Radiological Event Scale (INES), due to earthquake and tsunami damage (Norio et al., 2011). The 5.7 m seawall at the Daiichi power plant was overcome by a 15 m high tsunami that flooded backup diesel generators and washed their fuel tanks into the ocean (Funabashi and Kitazawa, 2012), which meant that the power plant had no diesel generators to power the cooling systems (Nakamura and Kikuchi, 2011). In turn, this resulted in the partial meltdown of the reactor cores, which led to significant releases of radiation into the atmosphere and the ocean (Funabashi and Kitazawa, 2012; Nakamura and Kikuchi, 2011; Chino et al., 2011).

1.3. The advent of the Safecast project

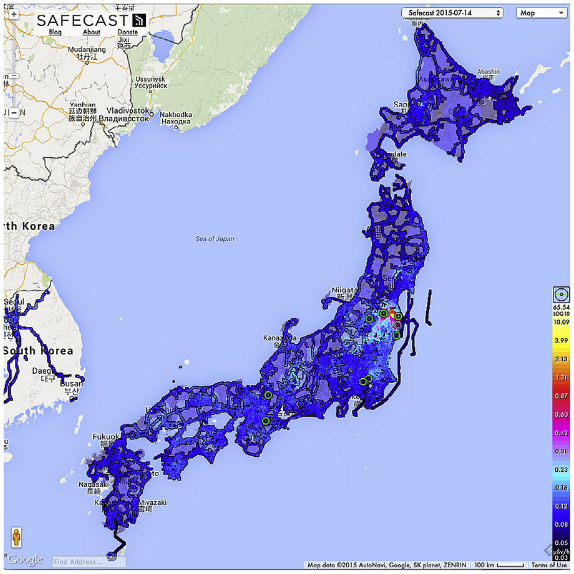
Motivated by the lack of reliable and publicly available information regarding the ongoing Fukushima Daiichi nuclear power plant disaster, that same month a group of hobbyists organized the Safecast project, which focused on providing the means for citizens to collect and share radiation observations. The Safecast project logged their first observations with handbuilt radiation detectors in April, 2011, just one month after the tsunami struck Japan (Brown et al., 2016). The Safecast project is internationally crowdfunded and crowdsourced with over 650 handheld units and several stationary sensors that have contributed more than 27 million radiation measurements as of March 2015 (Bonner and Brown, 2015). In May 2014 there were over 14 million Safecast observations within Japan, though there were also millions of observations from Korea, Iraq, the United States, and other locations.

Fig. 1 shows log-adjusted Safecast radiation observations for Japan, and depicts the radiation plume from the Fukushima Daiichi Nuclear Power Plant (FDNPP) spreading to the northwest about 50 km before turning and fading to the southwest. Most of the data follows the roadways because the hand held units are typically attached to car windows during observations, though there are also data gathered from ships off the east coast of the main island. The green circles indicate several permanent stationary sensors that also contribute observations to the Safecast database.

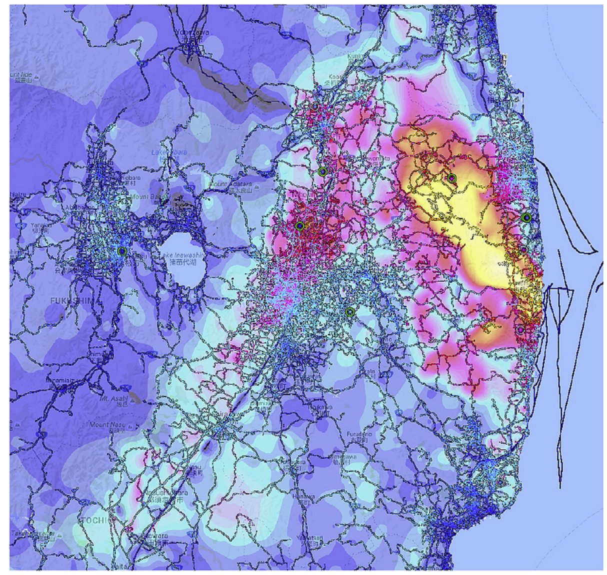
Safecast's current handheld radiation detector, the bGeigie Nano, is shown in Fig. 2, and is the fifth generation of their open source hardware design. It uses the LND 7317 radiation sensor, which is a 5.08 cm diameter pancake style radiation sensor that can detect alpha, beta, and gamma radiation using a Geiger-Müller tube filled with a mixture of neon and halogen gases (LND, Inc, 2011). The device also has a Global Positioning System (GPS) receiver to record the location of radiation readings. The Safecast detector records only the sensor output in counts per minute with time and location information and does not do any other manipulation of the saved data. For the display on the device, the counts per minute can be converted to either micro Sieverts per hour (μSv/h) or Becquerel per meter squared (Bq/m^2), both based on ^{137}Cs. Radiation observations are logged to a Secure Digital (SD) memory card, which can then be uploaded to the Safecast site; the data is freely available to the public via the Creative Commons CC0 license (Creative Commons, 2015). bGeigie Nano kits can be purchased online for roughly $450, and then assembled by users within a few hours (Safecast, 2015b; Brown et al., 2016). The device is typically deployed by attaching it to the outside of a vehicle's window and driving through areas of interest, though users can also collect observations from their bGeigie Nano while walking (Brown et al., 2016).

1.4. Safecast data Validity

The Safecast team has implemented quality control measures to ensure equipment accuracy. First, all the electronic components are factory tested before being shipped (Safecast, 2015a). Second, units are randomly selected, assembled, and undergo calibration tests at the Jülich Research Centre in Germany, QualTek in the US, and the International Atomic Energy Agency (IAEA) testing laboratory in Seibersdorf, Austria. The tested units have demonstrated ±10% accuracy, which is the typical Safecast performance (±15% is their maximum), which compares well to the normal industry



(a) Safecast observations for all of Japan.



(b) Safecast observations near the Fukushima Daiichi nuclear power plant.

Fig. 1. These are screenshots from the Safecast web site of log adjusted radiation observations in µSv/h for all of Japan and the vicinity near the FDNPP. Individual observations are overlaid on linearly interpolated values. Note that most observations follow along roadways, though the green circles represent permanent, stationary radiation sensors that also contribute to the Safecast network. The radiation plume from the Daiichi nuclear power plant can be seen following a northwesterly path for about 50 km before turning and fading southwest (Safecast, 2015d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

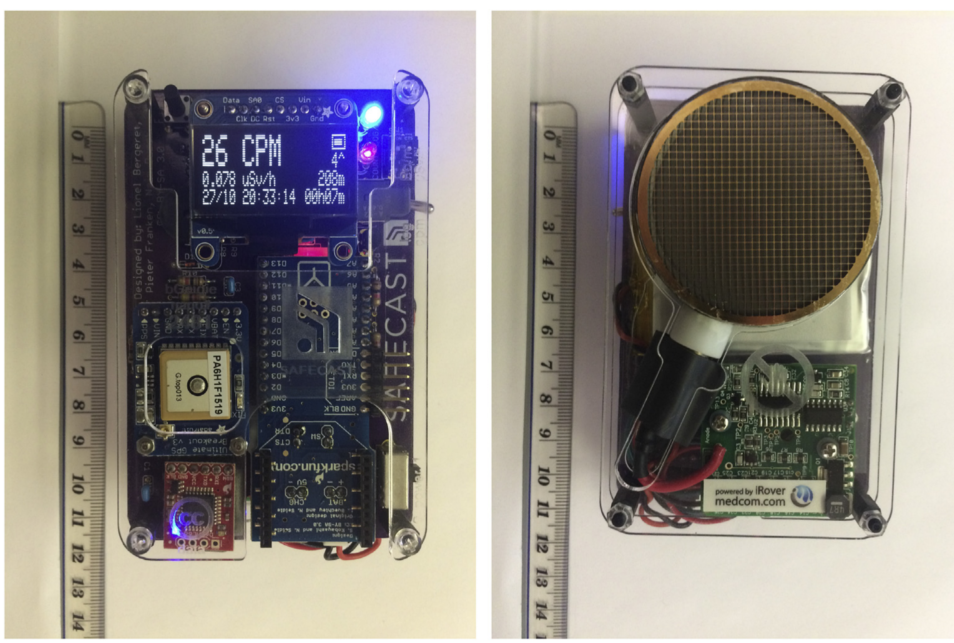


Fig. 2. These images show the front and back of the bGeigie Nano, and is the current handheld unit for contributing radiation observations to the Safecast database. This is the fifth generation of the bGeigie handheld sensor, which was released in 2013. (bGeigie Nano images courtesy of PSU's Geoinformatics and Earth Observation Laboratory).

calibration accuracy (Spinrad, 2011) and the recommended periodic check accuracy of ±10% for medical uses (Zanzonico, 2008). Third, all data is checked by a team of domain experts before being accepted into the database; however, approximately 0.1% of uploads required such scrutiny (Brown et al., 2016).

The kits are built by volunteers of varying technical ability, which means that there may be assembly errors that may have a negative impact on accuracy. Moreover, though there are guidelines for using the bGeigie Nano units to mitigate data accuracy problems (Safecast, 2015c), there is no guarantee that the users have followed those guidelines when deploying their devices, which also can have an impact on the unit's accuracy.

Given these concerns, to date there has been no in depth analysis of the accuracy of the millions of Safecast radiation measurements. It is necessary to be confident that Safecast observations are reliable and accurate if they are to be used to supplement authoritative data for decision making. One means of checking Safecast data is to compare it with a set of observations made from different equipment that come from a trusted source, which is the strategy we have taken in this work. In the next section we describe the details of this approach.

2. Material and methods

The Open Street Map (OSM) project is similar to Safecast in that it is a volunteer effort to freely provide global geospatial data. Individuals use GPSs, smartphones, or cameras to capture local spatial information that is then added to a central publicly accessible database. This database can then be used to create maps or to do route planning (Bennett, 2010). However, Volunteered Geographic Information (VGI) such as found with OSM can have problems with reliability, quality, and utility. That is, participants may be using faulty equipment or make observation errors that could negatively affect data quality (Flanagin and Metzger, 2008).

One way to validate VGI is to compare it with high fidelity data, such as from an authoritative and trusted source. OSM data for the London metropolitan area was compared to corresponding geospatial data from the British Ordnance Survey (BOS). London was selected because that is the first area OSM mapped, and so would have the oldest and therefore most reviewed data, as well as maximizing overlap between the two datasets (Haklay, 2010).

Safecast data is similar in nature to OSM's just by virtue of them both being VGI and, as such, just as there was an open question regarding OSM's quality since it was a form of VGI, the same question applied to Safecast. Therefore we took a similar approach to evaluating Safecast's observation fidelity by comparing Safecast observations with authoritative data. We chose to compare Safecast data with the DOE/NNSA aerial survey data gathered over the Fukushima Prefecture shortly after the nuclear disaster (Lyons and Colton, 2012) because these two datasets had significant overlapping spatiotemporal observations in that region.

3. Experimental

The DOE/NNSA dataset has 107,147 observations that cover roughly 20,000 km^2 over the Fukushima Prefecture for a period of five weeks, from April 2nd through May 9th, 2011. Given that the observations were made several hundred meters above sea level, the data values were corrected to what they would be 1 m above the ground presuming the ground or the air at this reference height is the reference. Also, since the observations were made on different days and the radiation from the elements ^{134}Cs and ^{137}Cs decay at different rates over time, to use that dataset you would have to take into consideration when the observations were made and which elements' radiation energy levels were measured to compensate for radioactive decay. Since this would be computationally cumbersome to do properly, the DOE/NNSA used the respective half lives of ^{134}Cs and ^{137}Cs to project forward all the observations to June 30, 2011 (Lyons and Colton, 2012).

The Safecast and DOE/NNSA data was compared by first clipping the Safecast data to the same geographical extent as the DOE/NNSA data, then considering only the Safecast observations for the same period — from April 2nd through June 30th — which resulted in 71,616 Safecast observations. Fig. 3 shows the distribution of observations between the two surveys for this period. We chose to emulate what was done for the DOE/NNSA aerial survey and extrapolated the remaining amount of Cs for the Safecast data to June 30, 2011, with the same assumption of a 1:1 ratio of the two Cs isotopes. The following formula was used to estimate the remaining amount of Cs radionuclides for June 30, 2011 (U. S. Occupational Safety & Health Administration, 2015):

$$A = A_i e^{-\left(0.693t/T_{1/2}\right)} \tag{1}$$

Where A is the activity at some time of interest, A_i is the activity at the initial time, t is the elapse time from the initial time to the time of interest, and $T_{1/2}$ is the given isotope's half life in the same units of time. The $T_{1/2}$ for ^{134}Cs is 2.06 years and ^{137}Cs is 30.17 years.

The DOE/NNSA used mR/h (miliRem per hour) as a unit of measure whereas Safecast used counts per minute (cpm), or the ionization events that the Geiger-Müller tube detects. The Safecast measurements were converted to use mR/h to facilitate comparison to the DOE/NNSA data using the following formula (Kozhuharov, 2014; Mallins, 2014; Dolezal, 2014):

$$1\ \text{mR/h} = \frac{1}{3340}\ \text{cpm} \tag{2}$$

Note that we did not make additional adjustments for the devices measuring different energy levels. The constant in Eq. (2) includes this conversion.

Now that the two datasets were for the same area and time period, and also used the same units for measuring radiation, the next step was to do the comparison between them. However, the two datasets had distinctly dissimilar spatial characteristics, which posed a challenge for doing a direct comparison. For example, Fig. 4 shows that the DOE/NNSA areal survey covered large swathes of territory in a gridlike pattern; by contrast, Fig. 1a shows that the bulk of the Safecast observations were made along roads.

To compare the two datasets we chose to follow a similar approach taken by the Japanese Nuclear Regulation Authority (JNRA) for rasterizing radiation measurements. The JNRA uses the following steps to gather and share their radiation measurements (Japanese Nuclear Regulation Authority, 2014):

1 project the data to a 2D coordinate system, such as Universal Transverse Mercator (UTM)
2 overlay a polygonal grid onto the point data
3 average the measurements for each grid cell
4 assign that average to a corresponding cell in a separate raster image with the same dimensions

Following these steps to compare the DOE/NNSA to Safecast measurements, we first projected the two datasets to UTM Zone 54, then used a polygon grid overlay comprised of 500 m^2 grid cells, and then averaged the radiation readings corresponding to each cell for both datasets. We then derived two raster images, one for the Safecast data, and the other for the DOE/NNSA, where each pixel value contained the corresponding average for each corresponding grid cell. We kept raster cells for which there were common set of observations to allow for direct comparisons between the two datasets. Fig. 5 shows the process of rasterizing and filtering grid cells for which there were observations from both the DOE/NNSA and Safecast.

4. Results

Fig. 6 shows the respective distributions of radiation observations between the DOE/NNSA aerial survey and Safecast measurements. Both datasets are similarly distributed with ≈80% of the observations being below 0.1 mR/h. However, the DOE/NNSA observations are slightly higher in value than the Safecast, and there are two DOE/NNSA readings that are much larger than all other observations.

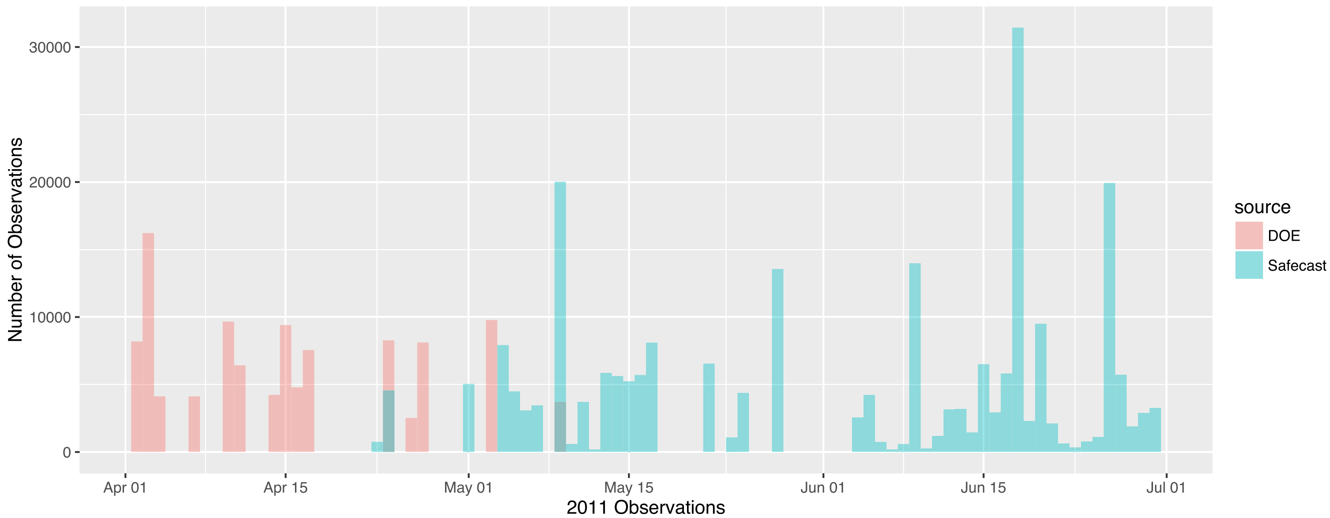


Fig. 3. This illustrates the number of samples gathered by day during the DOE/NNSA aerial survey of the Fukushima Prefecture from April 2nd through May 9th, 2011, and the Safecast observations made from April 24th, through June 30th, 2011. This shows that the bulk of the DOE/NNSA observations were made earlier in the same period than Safecast. However, both datasets had their respective observations extrapolated forward to June 30, 2011 for easy comparison.

Fig. 7 shows the respective spatial distributions of the DOE/NNSA and Safecast radiation measurements. There the two datasets appear visually to be highly correlated, which is supported by the pairwise $cor(\vec{v_d}, \vec{v_s})$ of 0.962, where $\vec{v_d}$ corresponds to the vector of DOE/NNSA observations, and $\vec{v_s}$ the Safecast. However, the non-peak DOE/NNSA values appear to be higher than the corresponding Safecast observations.

When the values of two identical sets of observations are plotted against each other, they normally align along the 45° diagonal. However, when the sets are not identical but just very similar, as when they are off by a constant, it is still possible to observe a strong linear relationship. Fig. 8a shows a scatter plot of Safecast measurements against the corresponding DOE/NNSA with the regression line with 95% confidence interval shown. The corresponding linear model has a p-value of less than 0.0001, and an adjusted R^2 of 0.9262. Table 1 shows the correspondingly low p-values and standard errors for the coefficients obtained through our statistical testing.

Results show that the DOE coefficient is less than one, providing supporting evidence that the DOE/NNSA values are generally higher with respect to Safecast values. This was saw earlier in Fig. 3. A Wilcoxon rank sum pairwise statistical test between the two datasets also supports the claim that the generally the DOE/NNSA data have higher values than much of the corresponding Safecast values ($p < 0.0001$). In other words, Safecast generally underestimates the radiation levels.

5. Discussion

While the Safecast and DOE/NNSA data were strongly correlated, the DOE/NNSA radiation measurements were generally higher. This may have several possible explanations. First, though we applied the same DOE/NNSA extrapolation procedure of projecting Safecast data to June 30th, 2011, the bulk of the Safecast measurements were made later than the DOE/NNSA, as shown in Fig. 3, and so extant potassium and iodine radionuclides would not have been as prevalent as for the DOE/NNSA to be detected by the bGeigie Nano units; moreover, any precipitation made before or during the Safecast observations would remove some of the water soluble Cs. Second, the observed differences could also be because the Safecast and DOE/NNSA observations were made by different sensor technology. That is, Safecast used a Geiger-Müller tube containing neon and halogen and the DOE and NNSA deployed thallium activated sodium iodide crystal-based detectors. Thirdly, another contributing factor to the higher DOE/NNSA readings could be that the Safecast measurements were predominately made from automobiles while the DOE/NNSA used a C-12 fixed wing aircraft, which meant that the DOE/NNSA measurements had to be extrapolated to 1 m above the ground, which was particularly challenging since the altitude had to be estimated given that aircraft's altimeter readings were inaccessible (Lyons and Colton, 2012), thus introducing a source of uncertainty. Fourth, given the strong linear relationship between the two datasets, another possibility is that the conversion factor that was used in Eq. (2) could be improved.

5.1. *Compensating for influence of early extant radioactive potassium and iodine*

If the existence of higher amounts of radioactive potassium and iodine were contributing factors in making the DOE observations greater than for Safecast, then we should observe that the largest values dominate the earliest measurements. Indeed, Fig. 9 shows that the highest recorded DOE radiation observations were made on the second day of flights.

Given that, the two datasets may be significantly closer in value if we drop the first two days of DOE observations. However, then the concern would be that the spatial distribution between the two datasets may change such that a fair comparison is no longer possible, but this concern may be mitigated if the areas corresponding to the dropped observations were later re-measured. The left sub-figure in Fig. 10 shows the subset of the DOE observations that are within 500 m of Safecast data that corresponds to the first two days — or days of highest recorded measurements; the right sub-figure shows all the remaining DOE measures, also within 500 m of Safecast data. This shows that we can drop the first two days of DOE measurements with little impact on the spatial representation of data since the earliest observed areas were measured again by the DOE.

Fig. 11 shows the regression between the DOE observations with those first two days of measurements removed compared once more to the corresponding Safecast data, and which looks almost

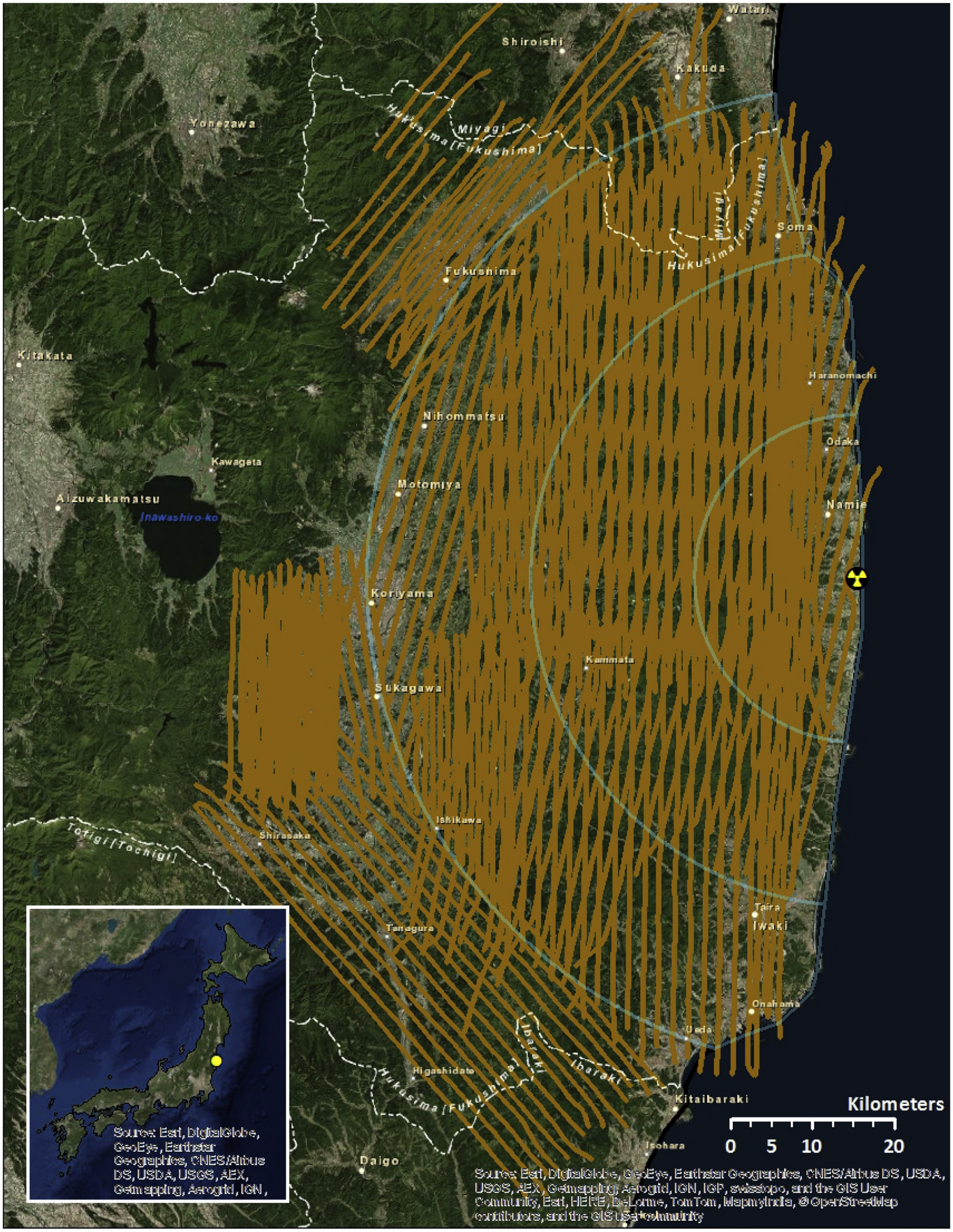


Fig. 4. DOE/NNSA radiation observations made over the Fukushima Prefecture from April 2nd through May 9th, 2011.

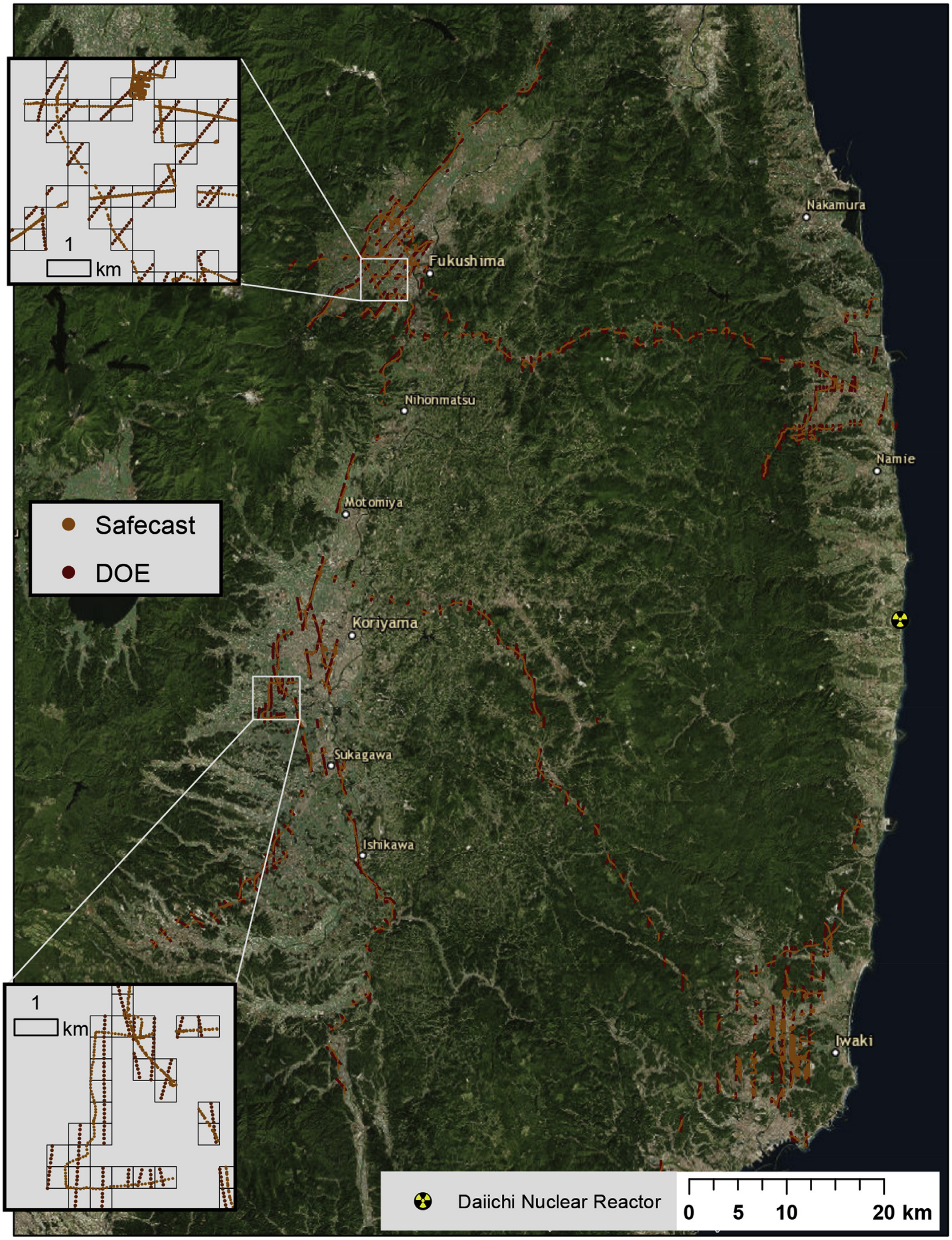


Fig. 5. The DOE/NNSA and Safecast datasets were point observations that were translated into a corresponding set of raster images to make comparisons easier. A grid of 500 m^2 cells was overlaid over the common area for the DOE/NNSA and Safecast observations as shown in the inset diagrams. A raster image for each set of observations was generated from the average of the respective observations for each grid cell, and grid cells for which there were no observations for both datasets were discarded. The location of the Fukushima Daiichi nuclear power plant is shown at the radiation hazard symbol along with 10 km concentric rings from the power plant shown up to 30 km.

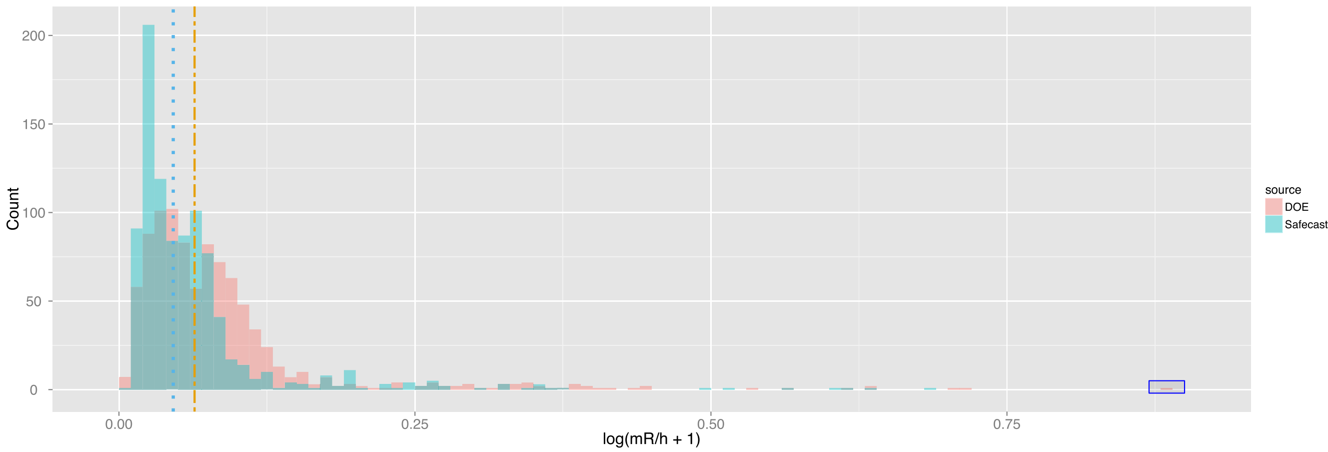
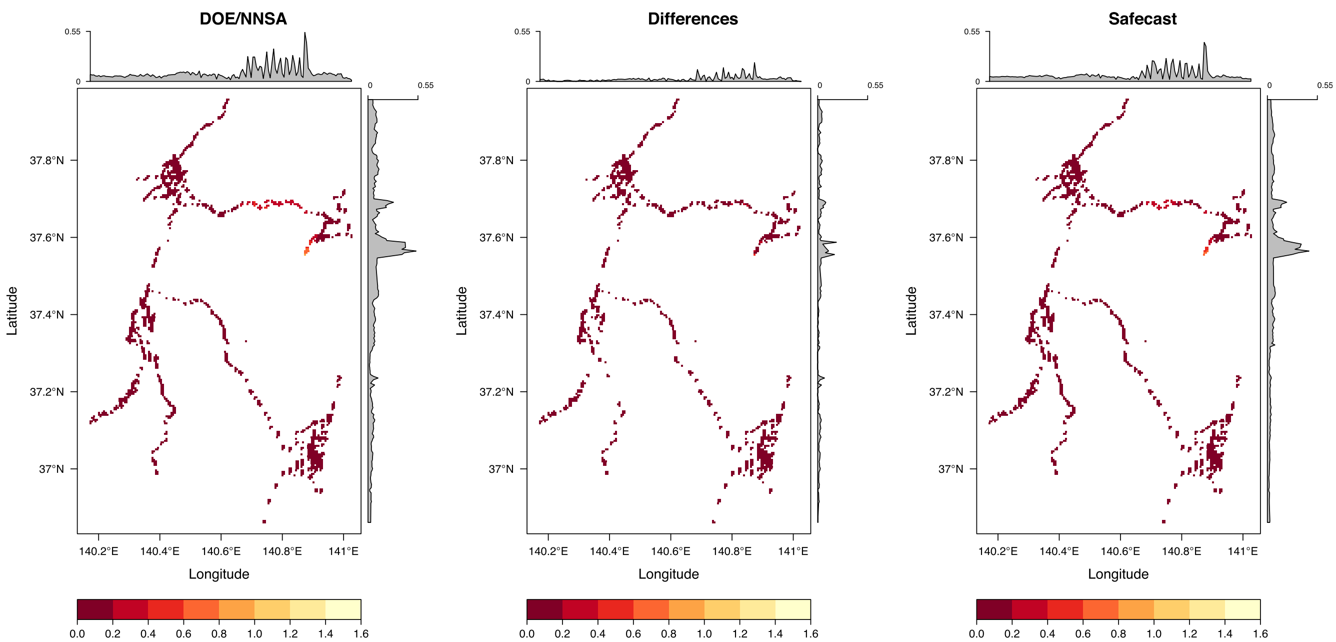


Fig. 6. This is a histogram of radiation intensities for DOE/NNSA and Safecast log adjusted observations in units of mR/h and with a bin size of 0.01. The number, or count, of observations for the corresponding bin values are shown, and indicates that the two sets of observations have similar distributions. However, the DOE/NNSA measurements are slightly larger than the corresponding Safecast observations as shown by their respective medians, with the gold line representing the DOE/NNSA median, and the blue dotted line for the Safecast. There are also two outliers of DOE/NNSA observations that are much higher than other measurements that are highlighted in the box on the right of the figure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(a) Level plot of the distribution of DOE/NNSA observations.

(b) Level plot of the differences between the DOE/NNSA and Safecast observations.

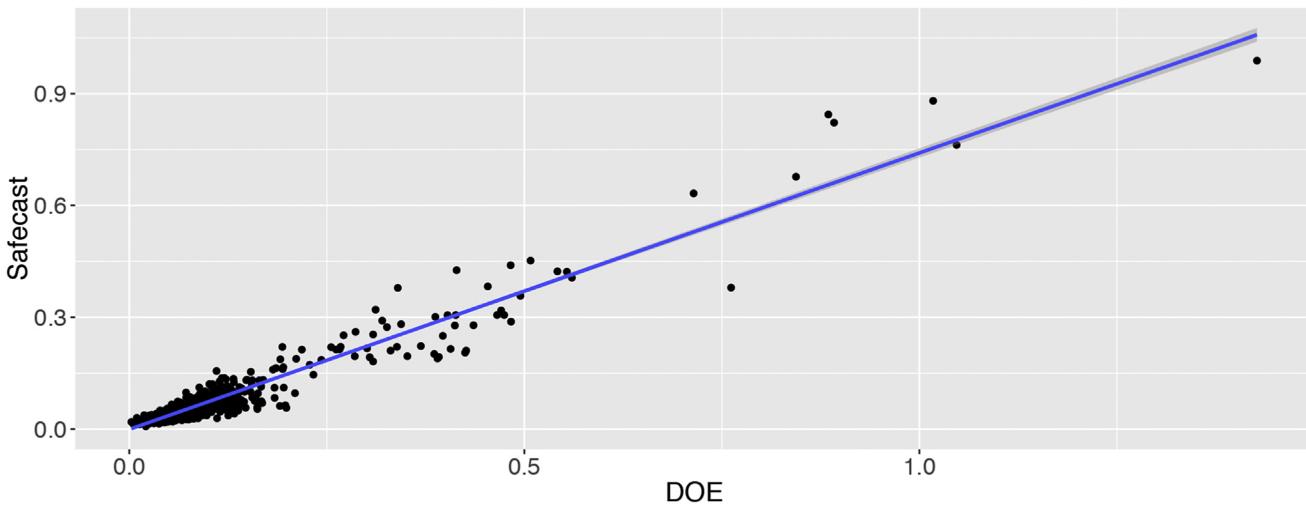
(c) Level plot of the Safecast data observations.

Fig. 7. These level plots depict the distribution of log adjusted radiation observations between the two datasets. The distributions appear to be highly correlated, though the DOE/NNSA non-peak observations are higher than the corresponding Safecast values, particularly for the areas of the highest level of measurements.

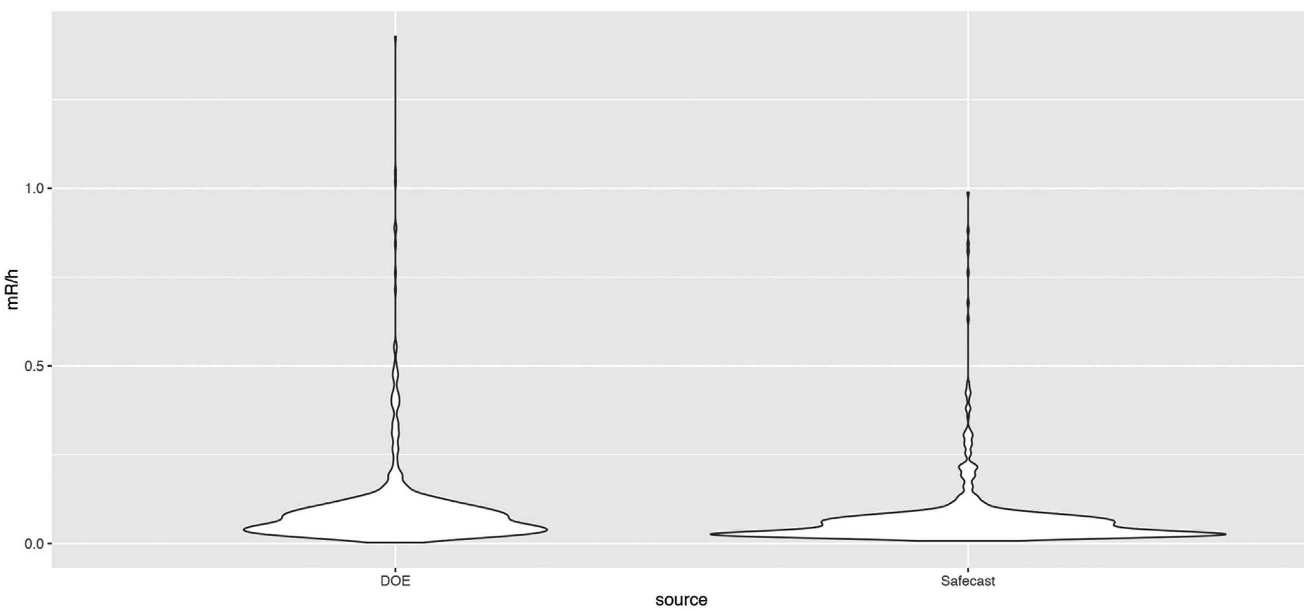
identical to the linear regression depicted in Fig. 8a. Table 2 corroborates this comparison in that it shows that the regression did not change much, though the adjusted R^2 did improve from 0.9262 to 0.9329. However, a pairwise Wilcoxon rank sum test shows that, overall, the DOE measurements are still statistically higher than the corresponding Safecast observations ($p < 0.0001$). Therefore, we conclude that removing the earliest, higher DOE measurements had little overall impact on the differences between the two datasets, and so one or more of the other possible explanations posed earlier may be the cause.

5.2. Summary

Regardless of the higher DOE/NNSA radiation values, even after removing the first two days of DOE observations containing the highest values, the two datasets are still strongly correlated. They both described the same relative regions of high vs. low areas of radioactive contamination. In this regard, Safecast has shown that it can be used to detect radioactivity, such as in scenarios for improving disaster response to radiation producing events.



(a) Linear regression model of DOE/NNSA vs. Safecast observations in units of mR/hr with linear regression 95% confidence interval region. That is, for every raster cell, the DOE/NNSA measurement is the x coordinate, and the Safeast the y. If the observations were identical, then the points would be in a 45° line. That the slope is less than 45° is another indicator that the DOE/NNSA values are somewhat higher.



(b) Violin plot that shows the differences between the DOE/NNSA and Safecast measurement distributions, and shows that the distribution of the Safecast observation values tend to be lower than that of the DOE/NNSA.

Fig. 8. Linear model of DOE/NNSA coordinates plotted against Safecast and a violin plot showing the respective measurement distributions. (a) Linear regression model of DOE/NNSA vs. Safecast observations in units of mR/h with linear regression 95% confidence interval region. That is, for every raster cell, the DOE/NNSA measurement is the x coordinate, and the Safeast the y. If the observations were identical, then the points would be in a 45° line. That the slope is less than 45° is another indicator that the DOE/NNSA values are somewhat higher. (b) Violin plot that shows the differences between the DOE/NNSA and Safecast measurement distributions, and shows that the distribution of the Safecast observation values tend to be lower than that of the DOE/NNSA.

Table 1
Summary statistics for coefficients for linear model of DOE/NNSA vs. Safecast coordinates shows correspondingly low p-values and standard errors.

	Estimate	Std. Error	t value	Pr (> |t|)
(Intercept)	−0.0007	0.0010	−0.73	0.4633
DOE	0.7418	0.0069	107.33	0.0000

6. Conclusions

Volunteers for the Safecast project use handheld sensors to gather radiation measurements that are later freely shared with the public. Unfortunately, the fidelity of this volunteer gathered data may be questionable given that it relies on participants of varying levels of training and from equipment that they assembled themselves. With this in mind, we validated Safecast data by comparing

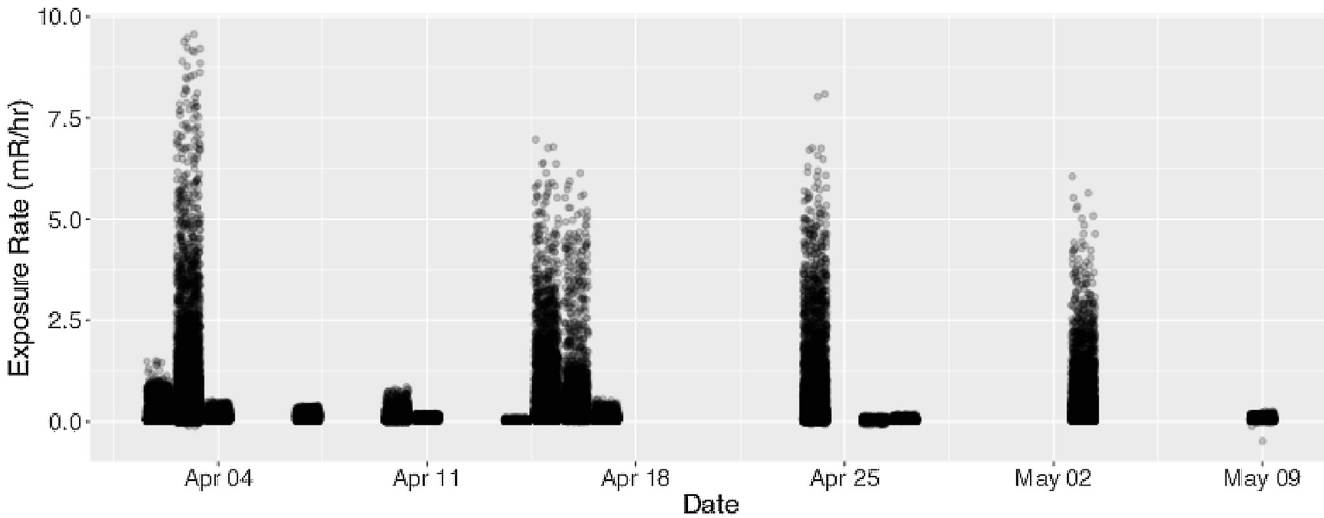


Fig. 9. From April 2nd through May 9th, 2011, the DOE/NNSA made 15 flights over the Fukushisma Prefecture to gather radiation observations, and this scatter plot aggregates those observations by each day. Note, some horizontal jitter and alpha transparency was applied to mitigate overplotting. The highest radiation observations were made on April 3, 2011.

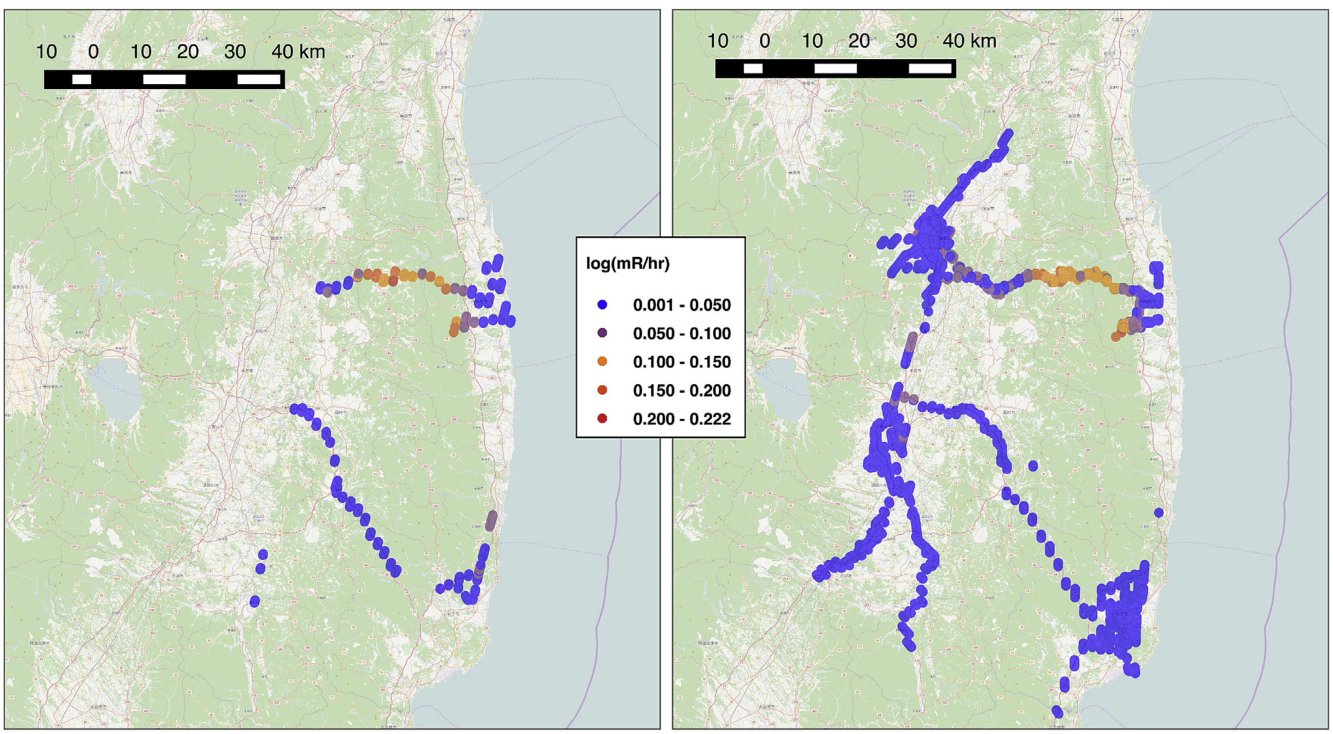


Fig. 10. Both figures depict the DOE radiation measurements made within 500 m of the corresponding Safecast observations. The left figure shows the 647 DOE observations made between April 2nd and 3rd, 2011, which also contained the highest measured values. The right shows the remaining 4533 observations, which were made between April 4th and May 9th, 2011. This shows that dropping the first two days of DOE observations does not have a significant impact on the overall spatial coverage.

it to U. S. Department of Energy's National Nuclear Security Administration observations made within the Fukushima Prefecture shortly after the March 2001 Daiichi nuclear power plant disaster occurred. We found that the two sets of observations were highly correlated, but that the DOE/NNSA measurements were somewhat higher. Nonetheless, despite the differences, we feel that the Safecast data is useful for public safety given that it identified similar regions of high radiation as did the DOE/NNSA.

One possible cause of the DOE/NNSA and Safecast data differences included significant periods of non-overlapping observations. One way to address that problem would be to perform similar types of comparisons as was used in this work between Safecast data and other authoritative datasets with observations made in the same span. Given that the Japanese government continues to regularly monitor radiation levels via aerial surveys of the Fukushima Prefecture, and likewise Safecast measurements

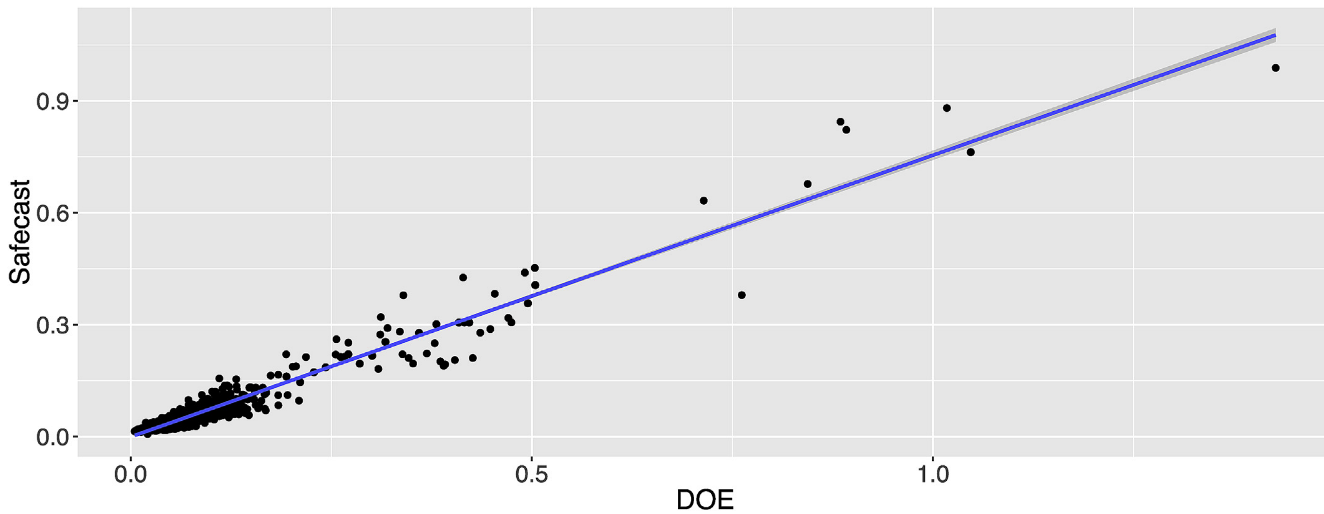


Fig. 11. Linear regression between DOE and Safecast observations with the first two days of DOE observations, which contained the highest values, dropped.

Table 2
Summary statistics for coefficients for linear model of DOE/NNSA vs. Safecast co-ordinates shows correspondingly low p-values and standard errors.

	Estimate	Std. Error	t value	Pr (> |t|)
(Intercept)	−0.0007	0.0010	−0.70	0.4846
DOE	0.7550	0.0070	107.79	0.0000

continue to be made in that same area, then similar studies can be made between those datasets. Likewise, other comparisons could be made where other authoritative data sufficiently overlaps with Safecast data.

Disclaimer

Oak Ridge National Laboratory is managed by UT-Batelle, LLC. This submission was written by one of the authors acting in their own independent capacity and not on behalf of UT-Batelle, LLC, or its affiliates or successors.

Acknowledgements

This research was funded by ONR grants N00014-13-1-0784 and N00014-14-1-0208. We would also like to thank Kalin Kozhuharov, Nick Dolezal, and Alex Mallins for their help with the Safecast cpm to mR/h unit conversions, and other participants on the Safecast mailing list for their feedback. We would also like to thank the reviewers for their valuable feedback.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jenvrad.2017.01.005.

References

Bennett, J., 2010. OpenStreetMap. Packt Publishing Ltd, Olton Birmingham, GBR, ISBN 978-1-84719-750-4.
Bonner, S., Brown, A., March 2015. The Safecast Report. http://blog.safecast.org/2015/03/the-safecast-report/.
Brown, A., Franken, P., Bonner, S., Dolezal, N., Moross, J., Jun. 2016. Safecast: successful citizen-science for radiation measurement and communication after Fukushima. J. Radiol. Prot. 36 (2), S82–S101. http://stacks.iop.org/0952-4746/36/i=2/a=S82?key=crossref.191ea617bad97e08b87e3a7ad72e6781.
Chai, C., March 2011. Japan's Quake Shifts Earth's axis by 25 Centimetres. http://www.webcitation.org/5x95t0CLU.
Chino, M., Nakayama, H., Nagai, H., Terada, H., Katata, G., Yamazawa, H., 2011. Preliminary estimation of release amounts of 131i and 137cs accidentally discharged from the Fukushima Daiichi nuclear power plant into the atmosphere. J. Nucl. Sci. Technol. 48 (7), 1129–1134.
Creative Commons, December 2015. Cc0 1.0 Universal (Cc0 1.0) Public Domain Dedication. http://creativecommons.org/publicdomain/zero/1.0/.
Dolezal, N., August 2014. Converting Safecast Measurements to mR/h. Personal Communication.
Flanagin, A.J., Metzger, M.J., 2008. The credibility of volunteered geographic information. GeoJournal 72, 137–148.
Funabashi, Y., Kitazawa, K., 2012. Fukushima in review: a complex disaster, a disasterous response. Bull. Atomic Sci. 68 (2), 9–21.
Goodchild, M., 2007. Citizens as sensors: the world of volunteered geography. GeoJournal 69 (4), 211–221.
Haklay, M., July 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environ. Plan. B Plan. Des. 37 (4), 682. http://ideas.repec.org/a/pio/envirb/v37y2010i4p682-703.html.
Japanese Nuclear Regulation Authority, November 2014. Monitoring Information of Environmental Radioactivity Level. http://radioactivity.nsr.go.jp/en/.
Kozhuharov, K., August 2014. Converting Safecast Measurements to mR/h. Personal Communication.
LND, Inc, 2011. LND 7317 Technical Specification. http://www.lndinc.com/products/17/.
Lyons, C., Colton, D., 2012. Aerial measuring system in Japan. Health Phys. 102 (5), 509–515. http://ovidsp.tx.ovid.com/sp-3.13.1a/ovidweb.cgi?QS2=434f4e1a73d37e8ce55b227ffaae5fa18cd3c69d19578a71a527e854f162eac1385894901cd57c91f11521e11845fd38621df082bb19136adabde3cd41232620f1cd461405c5f85f74d11bd4facb8785d15f9a37d1ab898c41c62fcbc765058d01351802ad9f9d83167e527efac970c938b1a231d62e47801e34f83b7474ee2cf8d039c25f095b4e1f27241008b85d0c915af226446df88bba1a6a66692c4c838a3ec32139abbb7e57951c4b3aa485abcb76efff18106aa0799563220c1576a7b755d7c1fdd244d725006875ee2f30a910623ffc05db0943975f5ba28fdd4ec09c8a4ab09ceb7f705a137300163a2019f71685d14e30ae9e90b5079d4c15f872c8b15bcaade70cba7f2d4a22d8bfe5db35bef8e2c94a6deb1d0184b5172026df82bf2040da67d81e6734c404ffe5cb0b6e3c0925965c668b3d7c741a3781bf57523fc406f7831767b68fdb1c6a3671b1b571c868e2c3740e6a196b35c6a53f8d4624b28dd287bafa9184325cd0e5389f96c9bc9cbc668d95cdc2a5c96956d8def233bca40aeb30f5.
Mallins, A., August 2014. Converting Safecast Measurements to mR/h. Personal Communication.
Nakamura, A., Kikuchi, M., 2011. What we know, and what we have not yet learned: triple disasters and the Fukushima Nuclear Fiasco in Japan. Public Adm. Rev. 71 (6), 893–899. http://dx.doi.org/10.1111/j.1540-6210.2011.02437.x.
Norio, O., Ye, T., Kajitani, Y., Shi, P., Tatano, H., 2011. The 2011 eastern Japan great earthquake disaster: overview and comments. Int. J. Disaster Risk Sci. 2 (1), 34–42. http://dx.doi.org/10.1007/s13753-011-0004-9.
Safecast, December 2015a. About Calibration and the Bgeigie Nano. http://blog.safecast.org/faq/about-calibration-and-the-bgeigie-nano/.
Safecast, December 2015b. Bgeigie Nano Kit. http://blog.safecast.org/bgeigie-nano/.
Safecast, December 2015c. Nano Operation Manual. https://github.com/Safecast/bGeigieNanoKit/wiki/Nano-Operation-Manual.

Safecast, August 2015d. Safecast maps. http://blog.safecast.org/maps/.
Spinrad, P., 2011. John Iovine: Geiger counter sanity check. Makezine. http://makezine.com/2011/04/25/john-iovine-geiger-counter-sanity-check/.
U. S. Geological Survey, March 2011. Magnitude 9.0-Near the East Coast of Honshu, Japan. http://earthquake.usgs.gov/earthquakes/eqinthenews/2011/usc0001xgp/.
U. S. Occupational Safety & Health Administration, December 2015. Ionizing Radiation Equations. https://www.osha.gov/SLTC/radiationionizing/introtoionizing/ionizingradequations.html.
Zanzonico, P., 2008. Routine quality control of clinical nuclear medicine instrumentation: a brief review. J. Nucl. Med. (49)

Generation of Realistic Mega-City Populations and Social Networks for Agent-Based Modeling

Annetta Burger
Department of Computational and Data Sciences
George Mason University, USA
aburger2@gmu.edu

Talha Oz
Department of Computational and Data Sciences
George Mason University, USA
toz@gmu.edu

Andrew Crooks
Department of Computational and Data Sciences
George Mason University, USA
acrooks@gmu.edu

William G. Kennedy
Center for Social Complexity
George Mason University, USA
wkennedy@gmu.edu

ABSTRACT

Agent-based modeling is a means for researchers to conduct large-scale computer experiments on synthetic human populations and study their behaviors under different conditions. These models have been applied to questions regarding disease spread in epidemiology, terrorist and criminal activity in sociology, and traffic and commuting patterns in urban studies. However, developing realistic control populations remains a key challenge for the research and experimentation. Modelers must balance the need for representative, heterogeneous populations with the computational costs of developing large population sets. Increasingly these models also need to include the social network relationships within populations that influence social interactions and behavioral patterns. To address this we used a mixed method of iterative proportional fitting and network generation to build a synthesized subset population of the New York megacity and region. Our approach demonstrates how a robust population and social network relevant to specific human behavior can be synthesized for agent-based models.

KEYWORDS

Agent-based Models, Geographical Systems, Population Synthesis, Social Networks, Mega-city

CSS '17, October 19–22, 2017, Santa Fe, NM, USA

DOI: 10.1145/3145574.3145576

1 INTRODUCTION

Agent-based models (ABMs) are increasingly being used to study complex systems involving human and environment interactions such as in the areas of epidemiology, transportation, migration, climate change, and urban studies [1], yet the social networks that inform and influence human interactions remain largely absent from these models. Creating robust synthetic populations with their social networks remains a challenge in agent-based models. Traditional population synthesis methods of synthetic reconstruction and combinatorial optimization involve generating the population by fitting individual agents into set distributions of attributes based on contingency tables from demographic statistical or survey data. These distributions of attributes do not extend to the social network tie information that may be latent in the demographic data or captured in social media data. The proposed method addresses the gap between current population synthesis and social network analysis with a set of algorithms to generate synthetic social networks for agent-based models.

Our study uses an agent-based model to simulate human behavior in the event of a nuclear explosion in New York City. Both the population and social networks are synthesized to represent the actors and relationships relevant to human behavior during the impact phase of a no-warning disaster or emergency. During the impact time phase, individual behavior is influenced by immediate family and group cohorts. The model includes New York City and the extended region, shown in Figure 1, to represent a mega-city and its surrounding area and allow migration flows of large numbers of individuals away from the urban center. Until empirical data for richer patterns of life are identified and addition modes of transportation are integrated into the model, our agents are restricted to travel along roads. Within the 262 x 234 km area 22,795,866 people travel along 225,906 km

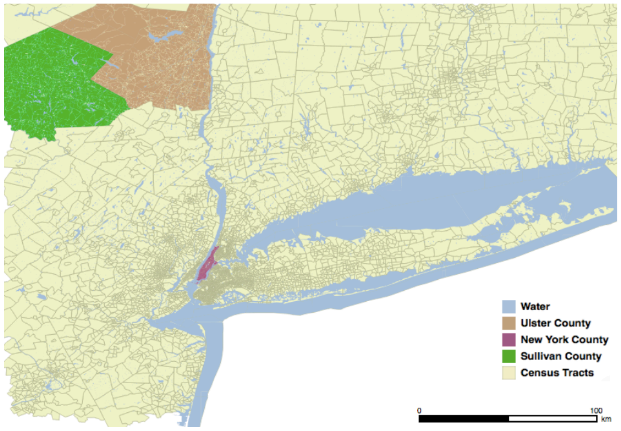


Figure 1: **Map of the New York megacity region highlighting Ulster, Sullivan and New York (Manhattan) counties.**

of roads between home, work, and school. The control population is based on available data including household demographic profiles, school and work place locations, and commuting distances. Because current social theory and empirical data emphasizes the importance of social networks for information and decision-making, the control population is generated with connected household networks and small-world network ties to represent relationships between co-workers and school. Family ties are known to have significant influence on human behavior and decision-making in disasters [2, 3, 4], and at the time of an emergency humans form ad hoc groups to improve their chances for survival [5, 3, 6, 7, 8]. In this paper we present a mixed method of iterative proportional fitting and social network generation to create an experimental control population for two New York counties, Ulster and Sullivan, and discuss our findings.

In the remainder of this paper we first provide background information on population and social network synthesis for agent-based models (Section 2). We then describe our methodology in Section 3, before discussing our validation results in Section 4. Finally we briefly provide a summary of our research and discuss areas of further work (Section 5).

2 BACKGROUND

Corresponding with the increasing use of agent-based models as a method of study in economics, sociology, ecology and other social science disciplines, new techniques for generating realistic model populations are being developed to leverage existing data sets and improve the representation of individual study subjects for simulation. Agent-based models provide the opportunity to study how the heterogeneous characteristics of individuals within a population generate patterns of behavior, however synthesis of these individual models is dependent on aggregate data sets, limited survey data or statistical representations of a population. These data do not reflect real-world heterogeneity and rarely represent social relationships beyond family or household ties. By exploiting improvements in computational techniques for population synthesis and social networks, algorithms can be integrated into existing population synthesis methods. The integration of current population synthesis methods and available social network algorithms provide an opportunity to study how population heterogeneity and multipartite social networks create patterns of human behavior within an experimentally defined context.

Methods for population synthesis in agent-based models originate from microsimulation techniques and involve a two-step process of fitting a population to a set of relevant attributes and constraints and then generating individual units on the fitted population [9, 10, 11]. Broadly, these can be categorized into sample-based and sample-free methods [12]. The sample-based methods can be broken into two categories [13, 14]. These are (1) *synthetic reconstruction* (SR) [e.g. 15]; and (2) *combinatorial optimization* (CO) or reweighting [e.g. 16]. The first, SR, involves obtaining the joint-distribution of relevant attributes and using Iterative Proportional Fitting (IPF) [17] with the sample population used to create a fitted population and generate individual units on that population. The second, CO, involves creating a population and modifying it with the sample population until it meets a threshold of required constraints [9, 18]. Synthetic Populations and Ecosystems of the World (SPEW) [19], RTI U.S. Synthetic Household Population [20, 21] and Virginia Bioinformatics Institute Synthetic Data [22, 26] are examples of populations built with SR and used in agent-based models for the study of infectious disease, and PopGen [23, 24] is an example of a synthetic population built for urban planning and analysis of transportation, routes, activities, vehicles, emissions and land-use. In sample-free methods, individual units are picked for a household or other grouping from the whole population as it progressively shrinks [13, 25].

Previous work on population synthesis for agent-based models has not directly accounted for relationships between individuals. However, indirectly these relationships are represented by information on whether individuals occupy the same household or have family members or social contacts in common. A growing body of agent-based modeling work use synthetic social networks . These can be found in a diverse set of topics including epidemiology [26, 27, 28, 29, 30], power structures [31], diffusion in networks [32, 33, 34), common pool resource governance [35], information sharing [36], rumor and riots [37], evacuation [9] and, safety-nets in socioeconomics [38].

Networks in ABMs have been used to generate and represent social, geographical and cognitive (semantic) spaces [38]. Social networks are commonly generated from available network algorithms representing regular lattice, random, small-world/scale-free (preferential attachment) networks [34, 36]. Gilbert and Hamill [37] created an algorithm that generates network ties based on social reach measures. Another common technique is to derive a synthesized network by leveraging existing datasets like those available in social media such as Twitter or Wikipedia [32, 9] or in demographic statistics [38]. In Agrawal, et al., [35] the ABM's social network of households was derived from Moore-neighbor households and a fraction of non-adjacent households. More rarely, social networks are developed endogenously from the model's internal processes. Social networks in ABMs can emerge dynamically as agents interact and decide to form or cut social ties. In past work, agents in these ABMs have decided to change network ties based on expected payoffs [39] or social influence interactions within the model's physical and social spaces [37].

Whether the social networks are derived exogenously from existing data or endogenously from processes within the model, these ABM social networks are typically limited by the availability of network data. The networks are derived from datasets of virtual social ties, such as social media or email, or aggregated data, as found in statistics, rather than empirically known ties formed in the physical world. We show in this paper that a synthesized social network can be realistically generated from empirical data by leveraging social demographics and known spatial characteristics such as home, school and work locations using heuristics. In the following section we discuss our method for synthesizing a population that explicitly captures social networks from empirical data.

3 METHODOLOGY

To demonstrate our approach we generated a control population for the two sparsely populated New York state counties of Ulster and Sullivan using a mixed method of population synthesis, derived from Barthelemy and Toint [13] and Wise [9], and network tie assignments. The algorithms were coded in Python to leverage existing Python libraries for mapping and data processing. Our method includes the following steps:

1. Creating a spatial environment with road network and places for homes, work and school,
2. Generating individual agents organized into households,
3. Assigning individual agents work and school daytime locations, and
4. Creating individual networks representing group membership in a family or other household type and either a work or school cluster.

This process is shown in Fig. 2. Method code is shared here: http://nbviewer.jupyter.org/gist/oztalha/a1c167f3879c5b95f721acef791c8111

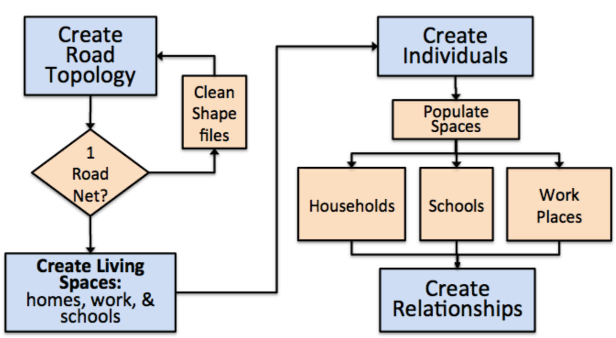


Figure 2: Method for population and social network synthesis.

The basic modeling environment is created with a transportation layer built from road network data provided by 2010 U.S. Census data, http://www.census.gov/cgi-bin/geo/shapefiles/index.php [40]. Tiger Shape files were used for the primary and secondary road systems of counties in New York State. These networks were merged to create a single giant connected component road network file for Ulster and Sullivan counties as shown in Fig. 3. To clean the road data file and create a network topology, we used GRASS (Geographic Resources Analysis Support System) C++ code libraries (also available in QGIS software). The process included simplifying lines, snapping lines to points, breaking lines at each intersection, removing duplicate geometric features, and removing small angles between lines at nodes.

Figure 3: Giant connected component road network for Ulster and Sullivan counties.

In the second step is a process of synthetic reconstruction (SR) in which we create an agent for every person within every census tract and assign their sex and age based on information from the U.S. 2010 Census data, https://www.census.gov/geo/maps-data/data/tiger-data.html [41]. The agents are grouped into households based on the household types present within a tract and on normal (Gaussian) distributions. The U.S. Census categorizes households into 10 types: husband-and-wife families, male/female/nonfamily householders, households with a child less than 18, and single householders over 65 and group quarters. Group quarters can be institutional (e.g. correctional facilities for adults, juvenile facilities, nursing facilities/skilled-nursing facilities) or non-institutional (e.g. college/university student housing, military quarters).

Home, work and school places are assigned in the third step. As we do not have the exact home locations or detailed land parcel information, houses are placed on local roads at least 50m apart or on top of each other when area population density is high (e.g., representing apartment complexes). The number of houses in the model is the number of occupied housing units in a census tract. Work places are randomly placed either onto secondary roads at ~20m apart or at local road intersections. We presume that in general zoning restricts businesses to secondary roads with the exception of institutions like religious centers that may be located on residential roads. No buildings are placed on primary roads as these are divided, limited-access highways [40]. County level business establishment counts (and binned-sizes) from the U.S. Census Bureau's County Business Patterns https://www.census.gov/data/datasets/2010/econ/cbp/2010-cbp.html [42] were disaggregated to the tract level and distributed in proportion to population size. We used a lognormal distribution within census tracts based on findings that job size distributions in U.S. cities are lognormal [43]. The number of work places and commuting patterns were derived from the U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) dataset and are found at https://lehd.ces.census.gov/data/lodes/LODES7/ [44]. After we aggregated the information to the tract-level, we assigned work-age agents to a random building location within a tract based on the origin-destination statistics. Data for school locations are extracted from the Educational Institution dataset retrieved from the US Environmental Protection Agency (EPA) Office of Environmental Information (OEI) (Education, US, 2015, ORNL Freedom, SEGS), https://geodata.epa.gov/arcgis/rest/services/OEI/ORNL_Education/MapServer [45]. The dataset contains geographic coordinates of educational institutions, enrollments, grade levels, and start and end grades of each institution. We assign school-age agents to the nearest available school location within a tract. School-age agents are sorted into schools based on grade and enrollment levels. In Fig. 4 we show representative example of home, work and school location for one census tract within our study area.

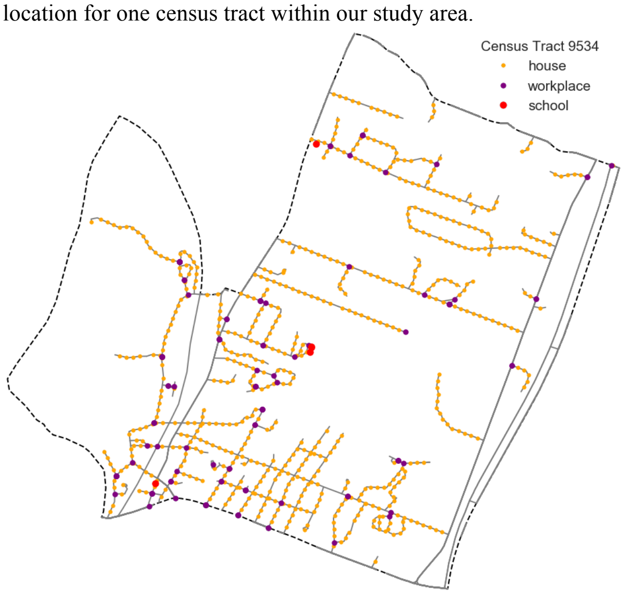


Figure 4: Mapped locations on Census Tract 9534.

Immediate family and group cohorts expected to be present at the time of the disaster event are represented with social network ties. In the remaining step we create social groups based on living in the same household, working in the same workplace, or attending the same school. Individuals receive a link to each agent located in the same household, work or school place. If the group size of a household, work or school is greater than 5 [46], a Newman-Watts-Strogatz [47] small-world network is generated. The resulting ties create individual and household multilayer networks and allow for simulation of the influence family members and group cohorts have on individual behavior. Fig. 5 shows an example of the multilayer network within a home including an individual's familial ties within their household and proximity ties to people at work and school.

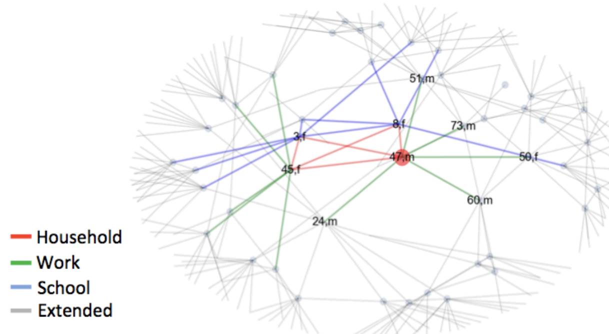


Figure 5: Sample social networks of an individual within a household.

4 RESULTS

The mixed method using 2010 U.S. Census data resulted in a population living, working, and going to school at locations on secondary and local roads within the road network. To ensure model realism, the baseline environment and representations for the synthetic population were derived from empirical data sets. (See Table 1.) For verification, a subset of three attribute

Table 1: Model characterization and empirical datasets

Model	Characterization	Dataset	Granularity	Source
Space	Road networks	2010 U.S. Census Tiger Shape Files	Primary, Secondary and Local Roads	http://www.census.gov/cgi-bin/geo/shapefiles/index.php
Population	Household demographics	2010 U.S. Census Tracts	Census Tracts	https://www.census.gov/geo/maps-data/data/tiger-data.html
Schools	Schools	U.S. Environmental Protection Agency (EPA) Office of Environmental Information (OEI) (Education, US, 2015, ORNL Freedom, SEGS)	Geolocation Coordinates	https://geodata.epa.gov/arcgis/rest/services/OEI/ORNL_Education/MapServer
Workplaces	Establishment Sizes	2010 U.S. Census County Business Patterns: Complete County File	County	https://www.census.gov/data/datasets/2010/econ/cbp/2010-cbp.html
Commutes	Commuting flows by tract	U.S. Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES)	Census blocks aggregated to tracts	https://lehd.ces.census.gov/data/lodes/LODES7/

measures not used explicitly in generating the synthetic population are shown: the average family size, the number of households with minors (under 18), and the number of household with seniors (over 65). We compared these for the synthesized population to the actual population (census tracts) as shown in Fig. 6. The percentage error is calculated as 100 * (Synthesized – Actual) / Actual for each census tract. The error for family size is under 4 percent for each tract, indicating that the synthesized population varies only slightly from the actual. The error for Households with minors ranged from -5 to 35 percent, and the error for Households with seniors ranged from -40 to 22 percent. The larger spread and error numbers for minors and seniors may be a result of varying numbers of seniors and minors housed in group quarters like juvenile facilities or nursing facilities. The outlier household with minor in Figure 6 at ~120 percent above the baseline is located in the census tract of State University of New York (SUNY) at New Paltz campus.

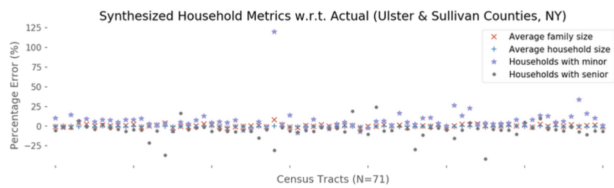


Figure 6: **Percentage error of synthesized population for each census tract.**

The social network resulting from our method consists of 232,096 nodes representing the total population of Ulster and Sullivan counties, and 736,757 ties representing relationships derived from living in the same household or going to the same work or school place. The majority of edges consist of household ties, and school ties represent the smallest portion of edges as shown in Fig. 7 for one specific census tract. Fig. 8 shows the degree distributions for the combined network and each of its edge types. The multi-layer network represents one layer of ties created to represent household relationships, and one to represent relationships present at daytime locations. In the household layer consists of individuals in cliques ranging from 0 to 10 ties with the majority of the population in small groups of 2 to 4 individuals. Households with only one person represent singles living alone, and their only relationships are work related. We also see that there are a few nodes in the workplaces with degrees ranging from 1 to 3 due to the small size of some work places. Because schools are occupied by groups of students, as expected we find no isolates.

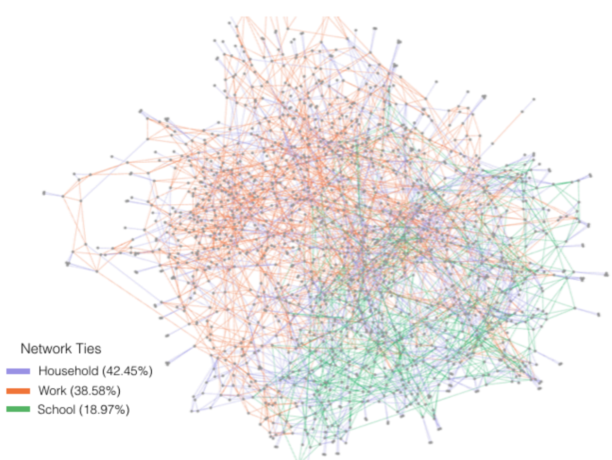


Figure 7: **Synthesized network of a census tract.**

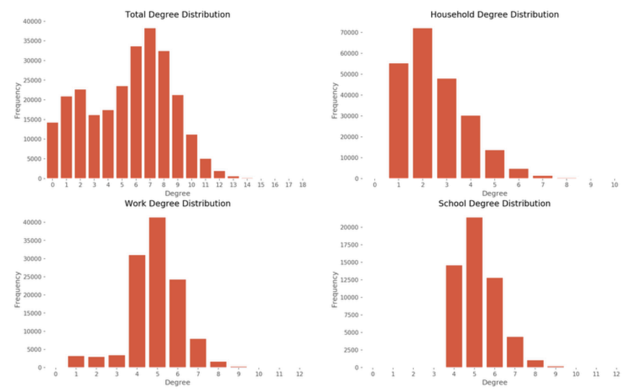


Figure 8: **Network degree distributions.**

5 CONCLUSIONS

We have presented methods to generate a synthesized population from available census data and to generate synthesized social networks for these agents. Synthesized populations that represent social networks allow for agent interactions at both the individual and the household levels in agent-based models. We developed a mixed method of population and social network generation from empirical data to represent the population in Ulster and Sullivan counties in New York State. Verification of the population synthesis techniques resulted in relatively low error rates as compared with unutilized household attribute measures from the US Census. The synthesized network represents multilayer patterns, and information dissemination and decision-making in the context of disasters or purchasing decisions. These synthesized networks no longer restrict agent-based models to simulations of interactions based only on physical proximity connections (i.e. adjacent cells), rather they allow distant and multi-layer network connections and interactions to impact agent that replicates aggregate statistical descriptions from empirical data, and yet maintains the anonymity of personal information. Techniques for anonymizing data are critical to the utilization of big data sets in simulation as the agent-based modeling community builds models at higher resolution and closer to real-world conditions.

In our next steps we will scale the synthesized population from two rural counties to the denser area of mid-town Manhattan and the greater New York megacity region. The population will be implemented in an agent-based model to create a baseline simulation of individuals interacting as part their daily behavioral patterns. Cognitive frames with decision-tree heuristics will support the agents having differentiated emergency responses and behavior based on historical data, and these will be exercised in the event of a nuclear detonation and its physical effects. A smaller area model of mid-town Manhattan will be expanded for higher resolution behavior to include additional modes of travel such as walking, subway, railway and bus and additional locations such as fire stations and hospitals. The current model includes network ties that are relevant during the evacuation phase of a disaster. To realistically capture support networks in times of disaster beyond the impact phase, the model will be expanded to include extended family such as grandparents, aunts and uncles, and siblings.

ACKNOWLEDGMENTS

This work was supported by the Center for Social Complexity at George Mason University and the Defense Technology Research Agency (DTRA) under Grant number HDTRA1-16-0043. The opinions, findings and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the sponsors.

REFERENCES

[1] Alison J. Heppenstall, Andrew T. Crooks, Linda M. See and Michael Batty (Ed.). 2012. Agent-based models of geographical systems. Springer, Dordrecht.

[2] Orjan Hultaker. 1983. Family and Disaster. International Journal of Mass Emergencies and Disasters 1, 2 (March 1983), 7–18.

[3] James D. Goltz, Lisa A. Russell, and Linda B. Bourque. 1992. Initial Behavioral Response to a Rapid Onset Disaster: A Case Study of the October 1, 1987 Whittier Narrows Earthquake. International Journal of Mass Emergencies and Disasters 10, 1 (March 1992), 43–69.

[4] Sirui Liu, Pamela Murray-Tuite, and Lisa Schweitzer. 2012. Analysis of child pick-up during daily routines and for daytime no-notice evacuations. Transportation Research Part A: Policy and Practice 46, 1 (January 2012), 48–67. DOI: https://doi.org/10.1016/j.tra.2011.09.003.

[5] Allen H. Barton. 1970. Communities in Disaster: A Sociological Analysis of Collective Stress Situations. Doubleday.

[6] Amanda Ripley. 2008. The unthinkable: who survives when disaster strikes and why (1st ed.). Crown Publishers, New York.

[7] N. Emel Ganapati. 2009. Rising from the Rubble: Emergence of Place-Based Social Capital in Golcuk, Turkey. International Journal of Mass Emergencies and Disasters 27, 2 (August 2009), 127–166.

[8] Sasha Rudenstine and Sandro Galea. 2012. The causes and behavioral consequences of disasters: models informed by the global experience 1950-2005. Springer, New York,

[9] Sarah Wise. 2014. Using Social Media Content to Inform Agent-based Models for Humanitarian Crisis Response. George Mason University. Retrieved September 3, 2015 from http://digilib.gmu.edu/xmlui/handle/1920/8879.

[10] Kirill Müller and Kay W. Axhausen. 2010. Population synthesis for microsimulation: State of the art. Retrieved July 12, 2017 from http://www.strc.ch/2010/Mueller.pdf.

[11] Mark Birkin and Belinda Wu. 2012. A Review of Microsimulation and Hybrid Agent-Based Approaches. In Agent-Based Models of Geographical Systems. Springer, Netherlands, 51–68.

[12] Maxime Lenormand and Guillaume Deffuant. 2013. Generating a Synthetic Population of Individuals in Households: Sample-Free Vs Sample-Based Methods. JASSS 16, 4 (2013), 12.

[13] Johan Barthelemy and Philippe L. Toint. 2013. Synthetic Population Generation Without a Sample. Transportation Science 47, 2 (May 2013), 266–279. DOI: https://doi.org/10.1287/trsc.1120.0408.

[14] David R. Pritchard and Eric J. Miller. 2012. Advances in population synthesis: fitting many attributes per agent and fitting to household and person margins simultaneously. Transportation 39, 3 (May 2012), 685–704. DOI: https://doi.org/10.1007/s11116-011-9367-4.

[15] Alan Geoffrey Wilson and Carol E. Pownall. 1976. A new representation of the urban system for modelling and for the study of micro-level interdependence. Area (1976), 246–254.

[16] Peter Williamson, Michael Birkin, and Phillip H. Rees. 1998. The Estimation of Population Microdata by Using Data from Small Area Statistics and Samples of Anonymised Records. Environment and Planning A 30, 5 (1998), 785–816.

[17] W. Edwards Deming and Frederick F. Stephan. 1940. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. The Annals of Mathematical Statistics 11, 4 (1940), 427–444.

[18] Zengyi Huang and Paul Williamson. 2001. A comparison of synthetic reconstruction and combinatorial optimisation approaches to the creation of small-area microdata. Department of Geography, University of Liverpool (2001). Retrieved July 12, 2017.

[19] Bill Eddy, Shannon Gallagher, Lee Richardson, and Sam Ventura. 2017. Synthetic Populations and Ecosystems of the World, Version 1.2.0. Retrieved September 15, 2017 from http://www.stat.cmu.edu/~spew/assets/spew_

[20] Richard J. Beckman, Keith A. Baggerly, and Michael D. McKay. 1996. Creating Synthetic Baseline Populations. *Transportation Research Part A: Policy and Practice* 30, 6 (November 1996), 415–429.

[21] 2014. US Synthetic Population 2010 Version 1.0. Retrieved September 22, 2017 from http://www.epimodels.org/10_Midas_Docs/SynthPop/2010_synth_pop_ver1_quickstart.pdf

[22] *Synthetic Data Products for Societal Infrastructures and Proto-Populations: Data Set 3.0*. Network Dynamics and Simulation Science Laboratory, Virginia Polytechnic Institute and State University, Blacksburg, VA. Retrieved September 22, 2017 from http://ndssl.vbi.vt.edu/Publications/ndssl-tr-07-010.pdf

[23] Xin Ye, Karthik Konduri, Ram M. Pendyala, Bhargava Sana, and Paul Waddell. 2009. A methodology to match distributions of both household and person attributes in the generation of synthetic populations. In *88th Annual Meeting of the Transportation Research Board, Washington, DC*. Retrieved September 22, 2017 from http://www.scag.ca.gov/Documents/PopulationSynthesizerPaper_TRB.pdf

[24] Theo Arentze, Harry Timmermans, and Frank Hofman. 2007. Creating Synthetic Household Populations: Problems and Approach. *Transportation Research Record: Journal of the Transportation Research Board* 2014, (December 2007), 85–91. DOI:https://doi.org/10.3141/2014-11

[25] Floriana Gargiulo, Sônia Ternes, Sylvie Huet, and Guillaume Deffuant. 2010. An Iterative Approach for Generating Statistically Realistic Populations of Households. PLoS ONE 5, 1 (January 2010), e8828. DOI: https://doi.org/10.1371/journal.pone.0008828.

[26] S. Eubank, H. Guclu, V.S. Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. 2004. Modeling Disease Outbreaks in Realistic Urban Social Networks. Nature 429, 6988 (May 2004), 180–4. DOI: https://doi.org/10.1038/nature02534.

[27] Chung-Yuan Huang, Chuen-Tsai Sun, Ji-Lung Hsieh, and Holin Lin. 2004. Simulating SARS: Small-World Epidemiological Modeling and Public Health Policy Assessments. Journal of Artificial Societies and Social Simulation 7, 4 (October 2004). Retrieved August 2, 2017 from http://jasss.soc.surrey.ac.uk/7/4/2.html.

[28] Jill Bigley Dunham. 2005. An Agent-Based Spatially Explicit Epidemiological Model in MASON. Journal of Artificial Societies and Social Simulation 9, 1 (January 2005). Retrieved August 2, 2017 from http://jasss.soc.surrey.ac.uk/9/1/3.html.

[29] Yong Yang and Peter M Atkinson. 2008. Individual Space – Time Activity-Based Model: A Model for the Simulation of Airborne Infectious-Disease Transmission by Activity-Bundle Simulation. Environment and Planning B: Planning and Design 35, 1 (February 2008), 80–99. DOI: https://doi.org/10.1068/b32162.

[30] Chung-Yuan Huang, Yu-Shiuan Tsai, and Tzai-Hung Wen. 2010. A Network-based Simulation Architecture for Studying Epidemic Dynamics. SIMULATION 86, 5–6 (May 2010), 351–368. DOI: https://doi.org/10.1177/0037549709340733.

[31] Armando Geller and Scott Moss. 2008. Growing qawm: An evidence-driven declarative model of Afghan power structures. Advances in Complex Systems 11, 02 (2008), 321–335.

[32] William Rand, Jeffrey Herrmann, Brandon Schein, and Neža Vodopivec. 2015. An Agent-Based Model of Urgent Diffusion in Social Media. JASSS 18, 2 (2015), 1.

[33] Chris J. Kuhlman, V. S. Anil Kumar, Madhav V. Marathe, S. S. Ravi, and Daniel J. Rosenkrantz. 2015. Inhibiting diffusion of complex contagions in social networks: theoretical and experimental results. Data Mining and Knowledge Discovery 29, 2 (March 2015), 423–465. DOI: https://doi.org/10.1007/s10618-014-0351-4.

[34] William Rand and Roland T. Rust. 2011. Agent-based modeling in marketing: Guidelines for rigor. International Journal of Research in Marketing 28, 3 (September 2011), 181–193. DOI: https://doi.org/10.1016/j.ijresmar.2011.04.002.

[35] Arun Agrawal, Daniel G. Brown, Gautam Rao, Rick Riolo, Derek T. Robinson, and Michael Bommarito. 2013. Interactions between organizations and networks in common-pool resource governance. Environmental Science & Policy 25, (January 2013), 138–146. DOI:

https://doi.org/10.1016/j.envsci.2012.08.004.
[36] Robert L. Goldstone, Michael E. Roberts, Winter Mason, and Todd Gureckis. 2008. Collective search in concrete and abstract spaces. In Decision modeling and behavior in complex and uncertain environments. Springer, 277–308. Retrieved July 21, 2017 from http://link.springer.com/chapter/10.1007/978-0-387-77131-1_12.
[37] Bianica Pires and Andrew T. Crooks. 2017. Modeling the emergence of riots: A geosimulation approach. Computers, Environment and Urban Systems 61, (January 2017), 66–80.
[38] Shah Jamal Alam and Armando Geller. 2012. Networks in Agent-Based Social Simulation. In Agent-Based Models of Geographical Systems. Springer, 199–214.
[37] Lynne Hamill and Nigel Gilbert. 2009. Social Circles: A Simple Structure for Agent-Based Social Network Models. Retrieved June 21, 2017 from http://jasss.soc.surrey.ac.uk/12/2/3.html.
[38] Samuel Thiriot and Jean-Daniel Kant. 2008. Generate country-scale networks of interaction from scattered statistics. In The fifth conference of the European social simulation association, Brescia, Italy. Retrieved July 26, 2017 from http://samuelthiriot.res-ear.ch/IMG/pdf/thiriot_2008_3.pdf.
[39] Josep M. Pujol, Andreas Flache, Jordi Delgado, and Ramon Sangüesa. 2005. How Can Social Networks Ever Become Complex? Modelling the Emergence of Complex Networks from Local Social Exchanges. Retrieved July 26, 2017 from http://jasss.soc.surrey.ac.uk/8/4/12.html.
[40] US Census TIGER/Line® Shapefiles. Retrieved August 9, 2017 from https://www.census.gov/cgi-bin/geo/shapefiles/index.php.
[41] US Census Bureau Geography. TIGER/Line® with Data. Retrieved August 9, 2017 from https://www.census.gov/geo/maps-data/data/tiger-data.html.
[42] US Census Bureau. County Business Patterns: 2010. Retrieved August 9, 2017 https://www.census.gov/data/datasets/2010/econ/ cbp/2010-cbp.html.
[43] Jan Eeckhout. 2004. Gibrat's law for (all) cities. *The American Economic Review* 94, 5 (2004), 1429–1451.
[44] US Census Index of /data/lodes/LODES7. Retrieved August 9, 2017 from https://lehd.ces.census.gov/data/lodes/LODES7/
[45] US Environmental Protection Agency (EPA) Office of Environmental Information (OEI)/ Oak Ridge National Labs (ORNL)_Education (MapServer). Retrieved August 9, 2017 https://geodata.epa.gov/arcgis/rest/services/OEI/ ORNL_Education/MapServer.
[46] Robin Dunbar. 1998. The Social Brain Hypothesis. Evolutionary Anthropology 6, 5 (1998), 178–190.
[47] Mark EJ Newman and Duncan J. Watts. 1999. Renormalization group analysis of the small-world network model. Physics Letters A 263, 4 (1999), 341–346.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 145 (2018) 757–764

Procedia Computer Science

www.elsevier.com/locate/procedia

Postproceedings of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2018 (Ninth Annual Meeting of the BICA Society)

Higher-level Knowledge, Rational and Social Levels Constraints of the Common Model of the Mind

Antonio Lieto[a,*], William G. Kennedy[b], Christian Lebiere[c], Oscar J. Romero[d], Niels Taatgen[e], Robert L. West[f]

[a] *University of Turin, Dept. of Computer Science, and ICAR-CNR (Palermo), Italy*
[b] *Center for Social Complexity, George Mason University, USA*
[c] *Carnegie Mellon University, Dept. of Psychology, USA*
[d] *Carnegie Mellon University, Dept. of Machine Learning, USA*
[e] *AI and Cognitive Engineering Institute, University of Groningen, The Netherlands*
[f] *Institute of Cognitive Science, Carleton University, Canada*

Abstract

We present the input to the discussion about the computational framework known as Common Model of Cognition (CMC) from the working group dealing with the knowledge/rational/social levels. In particular, we present a list of the higher level constraints that should be addressed within such a general framework.

Keywords: Unified Cognitive Architectures; Computational Models of Cognition; Common Model of the Mind; Cognitive Constraints; Knowledge Level; Rational Level; Social Level.

1. Introduction

In his famous 1982 paper, Allen Newell [23, 24] introduced the notion of knowledge level to indicate a level of analysis, and prediction, of the rational behavior of a cognitive artificial agent. This analysis concerns the investigation about the availability of the agent knowledge, in order to pursue its own goals, and is based on the so-called Rationality Principle (an assumption according to which "an agent will use the knowledge it has of its environment to achieve its goals" [23, p. 17]. By using the Newell's own words: "To treat a system at the knowledge level is to treat it as having

* Corresponding Author: Antonio Lieto
E-mail address: lieto@di.unito.it

10.1016/j.procs.2018.11.033

some knowledge, some goals, and believing it will do whatever is within its power to attain its goals, in so far as its knowledge indicates" [23, p. 13].

In the last decades, the importance of the knowledge level has been historically and systematically downsized by the research area in cognitive architectures (CAs), whose interests have been mainly focused on the analysis and the development of mechanisms and the processes governing human and (artificial) cognition. The knowledge level in CAs, however, represents a crucial level of analysis for the development of such artificial general systems and therefore deserves greater research attention [18]. In the following, we will discuss areas of broad agreement and outline the main problematic aspects that should be faced within a Common Model of Cognition [13]. Such aspects, departing from an analysis at the knowledge level, also clearly impact both lower (e.g. representational) and higher (e.g. social) levels.

2. Areas of Agreement

The analysis at the knowledge level is directly involved in, at least, three of the four dimensions considered within the Standard Model of the Mind [13] (later renamed Common Model of Cognition – CMC). In particular, it concerns the issues related to: i) the Structure and Processing mechanisms, ii) the Memory and Content of the CMC, as well as iii) its Learning processes. Concerning the first element, there is an agreement about the architectural necessity regarding the distinction between a Long-Term Declarative Memory and a Procedural one, as well as the necessity of a working memory module operating as a control interface between the Procedural module and other modules such as Declarative Memory and the Perception/Motor modules. Also, the cognitive cycle assumption [13] (with both serial and parallel information processing mechanisms between/within modules), seems perfectly compatible with the above mentioned Rationality Principle through which it is possible to evaluate the agent intelligent behavior.

For what concerns the Memory and Content issues, the integration of hybrid symbolic-subsymbolic representations and processing - and the inclusion of relevant metadata like frequency, recency, similarity, activation etc - represents the main element of difference with respect to the classical early symbolic CAs. The fact that such integration is necessary, in order to build integrated intelligent agents able to interact in the real world, is widely accepted. Some issues concern the way in which such integration can be obtained and novel solutions to address some representational problems of the Declarative Memories have been proposed and will be discussed in the sections below.

Finally, for what concerns the learning part, the facts that: i) all types of long-term knowledge are learnable, ii) learning is an incremental processes typically based on some form of a backward flow of information through internal representations of past experiences and, iii) learning over longer time scales is assumed to arise from the accumulation of learning over short-term experiences, also seem to be accepted constraint elements. These elements are also explicitly grounded in, and compliant with, the Anderson's Decomposition Thesis [1][1] based on the schema between different time-scales, types of operations and bands of cognition proposed by Allen Newell.

Figure 1 reproduces an extended version of the original four bands of cognition schema proposed by Newell [24] including the Biological, the Cognitive, the Rational, and the Social band. In the Newell framework, each band captures different types of human experience and represents different types of information processing mechanisms required to describe the levels within them. In particular: the neural band is described in terms of cellular biology, the cognitive band in terms of symbolic information processing, the rational band in terms of knowledge, reasoning and goals, and the social band in terms of distributed, multi-agent processing. The elements discussed so far in the CMC are, as for the entire enterprise of the cognitive architectures, mainly focused on the deliberate act level of the Cognitive band. In the following we provide additional elements of discussion for what concern both the Rational and the Social Band. Discussions about the lower band (e.g. the Biological one) are out of the scope of the present contribution.

[1] According to such thesis learning at the highest Band (the social one) can be reduced to learning occurring at lower bands. In general this thesis suggests that there is good evidence that high level tasks can be decomposed and understood at the micro-cognitive level, and that improvements at the micro-cognitive level can create improvements as measured at higher levels.

Levels of “Cognition”

t (sec)	*Time Terms*	*Band*	*System*
$10^{11\text{-}13}$	10^4-10^6 years	Evolutionary	Archeology
10^{10}	Millennia	Historical	Written History
10^9	~50 years	Historical	Personal history
10^8	Years	Historical	(Expertise)
10^7	Months	Social	(Expertise)
10^6	Weeks	Social	Culture
10^5	Days	Social	Culture
10^4	Hours	Rational	Task
10^3	10 min	Rational	Task
10^2	Minutes	Rational	Task
10^1	10 sec	Cognitive	Unit task
10^0	1 sec	Cognitive	Operations
10^{-1}	100 ms	Cognitive	Deliberate act
10^{-2}	10 ms	Biological	Neural circuit
10^{-3}	1 ms	Biological	Neuron
10^{-4}	100 μs	Biological	Organelle

Fig. 1. Extended version of the Newell’s Time Scales and the different Action Bands.(Green is Newell’s Figure 3.3, blue is Newell’s Figure 3.14, black added by W.G. Kennedy).

3. Rational Band

Newell equates the rational band with the knowledge level. The knowledge level refers to the level at which knowledge becomes abstract and can be treated largely independently from the physical systems that process it. Currently, the position according to which there is no need of including, in the CMC, specialized architectural modules to perform activities belonging to the rational band (e.g. planning, language processing and Theory of Mind) is majoritarian. The underlying assumption is that all such activities should arise based on the composition of processes executed during different cognitive cycles according to specific computational models. Despite this position, however, few cognitive architectures have formulated computational approaches to this effect, in particular to the Theory of Mind (ToM). Among these few works, there is that one by [27] where the SIGMA cognitive architecture [33] is used to demonstrate two distinct mechanisms (automatic processing vs. controlled reasoning) for ToM using as an example several single-stage simultaneous-move games, in particular, the well-known Prisoners Dilemma. Authors left open the possibility of using SIGMA’s learning capability to allow the agents to learn models of each other in a repeated game setting. ACT-R [2] has also been used to build several models of false belief and second-order false belief task (answering questions of the kind ‘Where does Ayla think Murat will look for chocolate?’) to assess whether children have a ToM [42]. Additionally, in [40] several scenarios were set up using ACT-R to show how ToM ability can improve the quality of interaction between a robot and a human by predicting what a person will do in different situations; e.g., that a person may forget something and may need to be reminded or that a person cannot see everything the robot sees.

Despite such efforts, however, the modeling attempts of such rational aspects remain still limited (while there are many more models developed for other phenomena concerning, for example, planning, or natural language processing etc.). An alternative proposal [44] to the view of creating specific computational models for cognitive phenomena occurring at the rational band, suggests to adopt additional schema to organize the activity at the higher bands to complement those proposed in Newell [24]. Such proposal will be discussed in more detail in the next sections since it also affects the modeling of phenomena occurring not only at the Rational band but also at the Social one and beyond.

In particular: in the section entitled "Knowledge Problems at the Rational Band", we will focus on some current limitations of the knowledge level of CAs and will show how this level of analysis also suggests some reflections for the CMC concerning the underlying representational assumptions to adopt/integrate within the Declarative Memory of cognitive agents. In the following section, instead, we introduce some of the main issues concerning the Social Band.

4. Social Band

Events happening at the Social Band can take place over longer time scales (days, weeks, and months). However, they are composed of social events happening at much shorter time scales (seconds, minutes, and hours) and, as evidenced by Anderson [1] may be supported by cognitive and rational processing in individuals at lower time scales. To discuss the potentially controversial topic of motivations, some of our needs, such as physiological needs, come from the bottom up and may explain why we developed high level cognition, i.e., the ability to solve those problems. The social level can also result in goals that people are intrinsically motivated to pursue and our cognition needs to be able to account for that. Our biological and social needs provide goals and may have resulted in innate cognitive capabilities. For example, we seem to have some innate capabilities associated with social cognition, such as perceiving other people, processing direction of gaze, determining intentions of others, limitations on knowing others as individuals (Dunbar's Number), and the topic of Theory of Mind. Related areas such as cooperation, trust, collective action are tasks or behaviors that arise at the social level. These social interactions may influence the CMC at the lower levels, not simply above the rational level.

Concerning the modeling of macro-scale or macro-cognition events there are, as mentioned, two different perspectives currently debated in the literature. On the one hand such elements are seen as too high level to be included within the minimal information processing mechanisms of a general cognitive architecture and, as such, are left to specific computational models to be developed on the top of such architectures. ACT-R, for example, has been used to study social behaviors and distributed collective decision-making processes which must balance diverse individual preferences with an expectation for collective unity. Romero and Lebiere [30, 31] proposed a multi-agent approach where cognitive agents have to reach global consensus while opposing tensions are generated by conflicting incentives, so agents have to decide whether to follow the most influential agent, follow the majority, negotiate with others, come to an agreement when conflicting interests are present, or keep a stubborn position.

Alternatively, the PolyScheme cognitive architecture [6] applies ToM to perspective taking in the human-robot interaction scenario, that is, the robot can model the scene from the human's perspective and use this information to disambiguate the command when moving in a scenario with multiple occluding elements [41]. However, although there are computational models of social interactions in a practical sense, there are no current CAs supporting research in social cognition, at best there are frameworks (as examples, BDI [28, 29] and PECS [43, 38]). An alternative position with respect to the current majoritarian view is proposed in [44] and [26]. These authors suggest the inclusion of macro-cognitive architectural elements that should specify the information processing mechanisms allowing to determine complex behavior at both the rational and social level bands and other constraints from the social band. These elements will be discussed in the section 'Problems for Macro-Cognition'.

5. Higher Bands

As Newell observed: "It is not clear that there actually are any higher bands". We agree that currently there is not evidence for a system level above the social level for cognition of individuals or for the cognition of groups of people/agents. For additional discussion on this point, we remind to the section 3.11 of Newell's Unified Theories of Cognition [24].

6. Knowledge Problems at the Rational Band

From a knowledge processing perspective, one of the main problem concerning the knowledge level of CAs is that, currently, the CAs are not able to deal with wide and complex knowledge bases that can be, even slightly, comparable (for what concerns both the size and heterogeneity of the handled knowledge[2]) to the amount of knowledge heuristically managed by humans. The limited size of the knowledge bases processed by the cognitive architectures was already acknowledged by Newell as a functional problem to address [24]. More recently, the content limit has been newly pointed out in literature [25] and some solutions for filling this "knowledge gap" have been proposed.

The ACT-R architecture, for example, has been semantically extended with an external ontological content coming from three integrated semantic resources composed by the lexical databases WordNet [21], FrameNet [23] and by a branch of the top level ontology DOLCE [20] related to the event modeling. In this case (see [25]), the amount of semantic knowledge selected to extend the ACT-R declarative memory only concerned the ontological knowledge about the events. While this is a reasonable approach in an applied context, it still does not allow to test the general cognitive mechanisms of a CA on general, multi faceted and multi-domain, knowledge. Therefore it does not allow to evaluate, *strictu sensu*, to what extent the designed heuristics allowing to retrieve and process, from a massive and composite knowledge base, conceptual knowledge can be considered satisfactory with respect to the human performances.

More recent works have tried to completely overcome the size problem of the knowledge level. To this class of works belongs one proposed by Salvucci [34] aimed at enriching the knowledge model of the Declarative Memory of ACT-R with a world-level knowledge base such as DBpedia (i.e. the semantic version of Wikipedia represented in terms of ontological formalisms) and a previous one proposed in [3] presenting an integration of the ACT-R Declarative and Procedural Memory with the Cyc ontology [15] (one of the widest ontological resources currently available). Both the wide-coverage integrated ontological resources, however, represents conceptual information in terms of classical symbolic structures and encounter the standard problems affecting this class of formalisms concerning the representation and reasoning on common-sense knowledge. (see [18] for a detailed treatment of this aspect). With respect to the size problem, the knowledge level is also problematic for the Soar [14] and the SIGMA [32, 33] CAs. Both architectures, in fact, do not currently allow to endow agents with general knowledge. For Soar, this problem is acknowledged in [14] but there is no available literature attesting progress in this respect [3]. A possible alternative solution that, in this perspective, is suitable to account for both the size problem and typicality (or common-sense) effects in conceptualization has been proposed in DUAL-PECCS [19]: a system that has been successfully employed to extend the Declarative Memory of diverse CAs and that combine, on a large scale, both common-sense representation and reasoning with standard ontological semantics. The main merit of such proposal lies in the adoption of the representational component of Conceptual Spaces [9] integrated with other neural-symbolic formalisms. The benefits coming from the integration of the Conceptual Spaces framework as an intercommunication layer between different types of representations in a general cognitive architecture has been recently pointed out in [17] and, with respect to the CMC, it has been acknowledged both in [7] and [11]. Recently, within the representation framework adopted in such system, it has been proposed a unifying categorization algorithm able to reconcile all the different theories of typicality about conceptual reasoning available in the psychological literature (i.e. prototypes, exemplars and the theory-theory, see [16]). Another interesting approach that aims at synchronizing beliefs and truth values among multiple domains to provide a unified treatment of these various forms of knowledge is described in [35]. In general, going towards unified representations and reasoning procedures seems to be a reasonable path to explore within the CMC research efforts.

Another important issue to consider in this dimension of analysis concerns whether the CMC should make a distinction between the knowledge about the 'self' and knowledge about 'others'. If so, a plausible solution would be that knowledge about the self is maintained by specialized memory systems, for instance, autobiographical declarative memories (both semantic and episodic) while beliefs about others are maintained by separate declarative memories. Likewise, procedural memory would contain distinct knowledge about actions allowing the agent to pursue both individual and collective goals, and those actions would be competing against each other inside the cognitive agent's

[2] The heterogeneity issue concerns the problem of representing different types of conceptual knowledge in a cognitive agent, including the common-sense one. Handling common-sense knowledge representation and reasoning mechanisms, however, still represents an unsolved problem to deal with. This aspect is problematic not only for the rational band processes but also for the information processing mechanisms occurring at Cognitive Band (in particular those involving the operations and unit task levels, see Figure 1).

[3] There are, however, attempts to extend in a efficient way the Semantic Memory of Soar with external lexical resources such as, for example, Wordnet [8].

mind when there would be conflicting interests. Another concern to be addressed is the potential necessity of keeping a collective memory to store shared knowledge about the world, social behaviors and norms developed by agents. This would be a distributed knowledge system by nature, but also would reflect high levels of redundancy, that is, agents would have a partial copy of the knowledge system (constructed as a result of learning processes) which would allow them to communicate ideas efficiently with each other. One potential issue with this approach is the representational incompatibility between agents grounded in a different cognitive architecture and, therefore, in different representational and reasoning assumptions. A possible way out, as suggested above, is that one of exploiting the potential of representational and reasoning systems *à la* DUAL-PECCS that have shown a good level of compatibility with diverse architectures. More generally, this is the kind of capability that the CMC aims to enable.

7. Problems for Macro-Cognition

As mentioned above, there is a majoritarian consensus that no additional specialized architectural modules are necessary for performing high-order capabilities of the social band. However, some specific architectural primitives may be still necessary for supporting social cognition such as visual imagery for visual-feature reasoning; pre-attentive and attentive vision sub-modules for the recognition of complex non-verbal signals of attention and emotional state, social-gaze following, face detection, etc. In the CMC, in fact, there is a real consensus over the 'minimum' number of high-level modules that are present in a cognitive architecture (e.g., perception and motor, working memory, declarative memory, and procedural memory); however, it still remains incomplete concerning other modules and mechanisms that can be useful for social cognition: emotion, direct communication, language, attention, metacognition, ethical/moral reasoning, among others [5, 36, 12, 37]. Thus, we envision that the main challenge is to identify which of these additional modules and mechanisms are completely necessary (and therefore should be included in the CMC) and what their level of involvement in social cognition is. For instance, it seems that at the lowest level, simple social interactions (such as gaze following) require at least the interplay of perception, attention and memory modules, whereas more complex social interactions (such as building rapport, negotiation, etc.) may require the interplay of additional modules such as emotion/motivations, metacognition, and language, just to name a few. So, it is necessary to define a well-structured hierarchy of a representative set of social interactions that help us identify which modules and mechanisms are strictly necessary at each level of the hierarchy and, from there, establish the architectural constraints that should be added to the CMC in order to allow social cognition to emerge.

Concerning the learning assumptions of the CMC connected the Social Band: the current version of the model [13] states that learning occurs mainly in two modules, procedural and declarative, where procedural learning involves at least reinforcement learning and procedural composition, and that more complex forms of learning involve combinations of the fixed set of simpler forms of learning. We know, however, from the Social Learning theory [4] that learning takes place in a social context and can occur purely through observation/imitation or direct instruction, even in the absence of motor reproduction or direct reinforcement. In addition to this, the field of robotics and multi-agent systems have reported [10] that learning by imitation can be supported not only by classical reinforcement learning but also by supervised learning (e.g., behavioral cloning, learning by demonstration), inverse reinforcement learning (e.g., apprenticeship learning) and transfer learning. Therefore, some questions to address at this point are: what other kind of learning, other than reinforcement learning, can be considered to model cognitively plausible social agents? What kind of architectural criteria should be taken into consideration in order to determine the 'minimum' requirements for procedural learning when modeling social cognition?

As mentioned earlier, an alternative proposal concerning the modeling of macro-cognitive phenomena has been recently pointed out by [44]. The authors applied to macro-cognition the same criticism raised by [22] as to the epistemological value of creating different models for each cognitive phenomena. In Newell's view this praxis is of limited use because it leads to a multitude of unrelated micro models. Similarly, [44] noticed that the field studying macro-cognition currently produces a vast array of ad hoc models and pointed out that this is true even when well specified cognitive architectures are used to make the models. This is because there are multiple ways to implement complex tasks in a cognitive architecture since it does not restrict the knowledge content at the rational band (see above). The adoption of Macro-cognitive architectures may offer some resolution to this problem as they are based on the proposal that the knowledge level is constrained and tends to be organized in particular ways.

In terms of relating the common model architectural principles to macro-cognition, a pragmatic approach is to specify best practices and common solutions for supporting macro-cognitive functions, such as planning, dealing with unexpected interruptions, task switching, problem solving, and dealing with large knowledge bases. This could be thought of as constituting a proto common model of macro-architecture, but one doesn't need to buy into the macro-architecture concept in order to appreciate that this sort of approach would have practical advantages for facilitating applications of the common model to real world modeling projects. A starting point for this would be to gather all the models built in common model architectures that are related to macro cognition and look for commonalities. That is, use the same methodology that was used to conceptualize the common model [13]. Likewise this same approach could be used with knowledge level languages to see if there are any commonalities.

Alternatively, an approach can be chosen in which the higher level of abstraction is treated as a more independent modeling platform with its own representations and mechanisms, but one that can be reduced to the level of the Common Model (e.g., [39]). The possible interplay between these two contrasting views is not yet entirely clear and represents, for sure, an important element of discussion and elaboration to address within this working group and between diverse CMC working groups.

8. Conclusion

We have proposed an overview of some of the main constraints and open problems that should be addressed within the CMC concerning the knowledge/rational/social levels. In our analysis we also have specified, when possible, some plausible directions to follow in order to overcome, or partially address, such problems.

Acknowledgments

This paper reflects the current thoughts of the majority of the working group on the reported issues. We also thank Phil C. Jackson and Jerald D. Kralik for their feedback. The first author acknowledges the support by a MIUR RTD-A grant from the University of Turin, Department of Computer Science. The second author is supported in part by the Center for Social Complexity, George Mason University and the Defense Thread Reduction Agency (DTRA) grant. The third author acknowledges the support of DARPA award FA8650-18-C-7826.

References

[1] John R Anderson. Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, 26(1):85–112, 2002.
[2] John R Anderson, Daniel Bothell, Michael D Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. An integrated theory of the mind. *Psychological review*, 111(4):1036, 2004.
[3] Jerry Ball, Stuart Rodgers, and Kevin Gluck. Integrating ACT-R and Cyc in a large-scale model of language comprehension for use in intelligent agents. In *AAAI workshop*, pages 19–25, 2004.
[4] Albert Bandura. Social learning theory. *General Learning Corporation*, 1971.
[5] Paul F Bello and Will Bridewell. There is no agency without attention. *AI Magazine*, 38(4):27–33, 2017.
[6] Nicholas Louis Cassimatis. *Polyscheme: a cognitive architecture for intergrating multiple representation and inference schemes*. PhD thesis, Massachusetts Institute of Technology, 2001.
[7] Antonio Chella, Marcello Frixione, and Antonio Lieto. Representational issues in the debate on the standard model of the mind. In *2017 AAAI Fall Symposium*, pages 302–307. AAAI Press, 2017.
[8] Nate Derbinsky, John E Laird, and Bryan Smith. Towards efficiently supporting large symbolic declarative memories. 1001:48109–2121, 2010.
[9] Peter Gärdenfors. *Conceptual spaces: The geometry of thought*. MIT press, 2000.
[10] Ahmed Hussein, Mohamed Medhat Gaber, Eyad Elyan, and Chrisina Jayne. Imitation learning: A survey of learning methods. *ACM Computing Surveys (CSUR)*, 50(2):21, 2017.
[11] Philip C Jackson. Toward human-level models of minds. In *2017 AAAI Fall Symposium Series*, pages 371–375, 2017.
[12] Iuliia Kotseruba and John K Tsotsos. A review of 40 years of cognitive architecture research: Core cognitive abilities and practical applications. *arXiv preprint*, 2018.
[13] J. E. Laird, C. Lebiere, and P. S. Rosenbloom. A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magazine*, 38:1–19, 2017.
[14] John Laird. *The Soar cognitive architecture*. MIT Press, 2012.
[15] Douglas B Lenat. Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11):33–38, 1995.

[16] Antonio Lieto. Heterogeneous proxytypes extended: Integrating theory-like representations and mechanisms with prototypes and exemplars. In *Biologically Inspired Cognitive Architectures 2018. BICA 2018. Advances in Intelligent Systems and Computing*, volume 848. Springer, Cham.
[17] Antonio Lieto, Antonio Chella, and Marcello Frixione. Conceptual spaces for cognitive architectures: A lingua franca for different levels of representation. *Biologically Inspired Cognitive Architectures*, 19:1–9, 2017.
[18] Antonio Lieto, Christian Lebiere, and Alessandro Oltramari. The knowledge level in cognitive architectures: Current limitations and possible developments. *Cognitive Systems Research*, 48:39–55, 2018.
[19] Antonio Lieto, Daniele P Radicioni, and Valentina Rho. Dual PECCS: a cognitive system for conceptual representation and categorization. *Journal of Experimental & Theoretical Artificial Intelligence*, 29(2):433–452, 2017.
[20] Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino, and Alessandro Oltramari. Wonderweb deliverable d18, ontology library (final). *ICT project*, 33052, 2003.
[21] George A Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.
[22] Allen Newell. You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. *Visual Information Processing: Proceedings of the Eighth Annual Carnegie Symposium on Cognition*, 1973.
[23] Allen Newell. The knowledge level. *Artificial intelligence*, 18(1):87–127, 1982.
[24] Allen Newell. *Unified theories of cognition*. Harvard University Press, 1990.
[25] Alessandro Oltramari and Christian Lebiere. Pursuing artificial general intelligence by leveraging the knowledge capabilities of ACT-R. In *Artificial General Intelligence*, pages 199–208. Springer, 2012.
[26] Mark G. Orr, Christian Lebiere, Andrea Stocco, Peter Pirolli, and William G. Kennedy. Multi-scale resolution of cognitive architectures: A paradigm for simulating minds and society. In *SBP-BRiMS*, pages 1–14, 2018.
[27] David V Pynadath, Paul S Rosenbloom, Stacy C Marsella, and Lingshan Li. Modeling two-player games in the sigma graphical cognitive architecture. In *Int. Conf. on AGI*, pages 98–108, 2013.
[28] Anand S Rao and Michael P Georgeff. Modeling rational agents within a BDI-architecture. *KR*, 91:473–484, 1991.
[29] Anand S Rao, Michael P Georgeff, et al. BDI agents: from theory to practice. In *ICMAS*, volume 95, pages 312–319, 1995.
[30] Oscar J Romero and Christian Lebiere. Cognitive modeling of behavioral experiments in network science using ACT-R architecture. In *MAS and Agent-Based Simulation*, pages 239–251, 2014.
[31] Oscar J Romero and Christian Lebiere. Simulating network behavioral dynamics by using a multi-agent approach driven by ACT-R cognitive architecture. 2015.
[32] Paul S Rosenbloom. Towards a new cognitive hourglass: Uniform implementation of cognitive architecture via factor graphs. In *9th Int. Conf. on cognitive modeling*, pages 116–121, 2009.
[33] Paul S Rosenbloom. Rethinking cognitive architecture via graphical models. *Cognitive Systems Research*, 12(2):198–209, 2011.
[34] Dario D Salvucci. Endowing a cognitive architecture with world knowledge. In *Procs. of the 36th Annual Meeting of the Cognitive Science Society*, 2014.
[35] Jonathan R Scally, Nicholas L Cassimatis, and Hiroyuki Uchida. Worlds as a unifying element of knowledge representation. *Biologically Inspired Cognitive Architectures*, 1:14–22, 2012.
[36] Brian Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, 12(1):13–24, 2002.
[37] Matthias Scheutz. The case for explicit ethical agents. *AI Magazine*, 38(4):57–64, 2017.
[38] Bernd Schmidt. *The modelling of human behaviour*, volume 132. Society for Computer Simulation International, 2000.
[39] Niels A Taatgen. Cognitive architectures: Innate or learned? In *2017 AAAI Fall Symposium Series*, pages 476–480, 2017.
[40] J Gregory Trafton, Laura M Hiatt, Anthony M Harrison, Franklin P Tamborello II, Sangeet S Khemlani, and Alan C Schultz. ACT-R/E: An embodied cognitive architecture for human-robot interaction. *Human-Robot Interaction*, 2(1):30–55, 2013.
[41] J Gregory Trafton, Alan C Schultz, Magdalena Bugajska, and Farilee Mintz. Perspective-taking with robots: experiments and models. In *Robot and Human Interactive Comm.*, pages 580–584, 2005.
[42] Lara M Triona, Amy M Masnick, and Bradley J Morris. What does it take to pass the false belief task? an ACT-R model. In *Cognitive Science Society*, volume 24, 2002.
[43] Christoph Urban. PECS: A reference model for the simulation of multi-agent systems. In *Tools and techniques for social science simulation*, pages 83–114. Springer, 2000.
[44] Robert L West and Korey MacDougall. The macro-architecture hypothesis: Modifying newell's system levels to include macro-cognition. *Biologically Inspired Cognitive Architectures*, 8:140–149, 2014.

Multi-Scale Resolution of Cognitive Architectures: *A Paradigm for Simulating Minds and Society*

Mark G. Orr[1], Christian Lebiere[2], Andrea Stocco[3], Peter Pirolli[4], Bianica Pires[1], and William G. Kennedy[5]

[1] Biocomplexity Institute of Virginia Tech
[2] Carnegie Mellon University
[3] University of Washington
[4] Institute for Human and Machine Cognition
[5] George Mason University

Abstract. [6] We put forth a thesis, the *Resolution Thesis*, that suggests that cognitive science and generative social science are interdependent and should thus be mutually informative. The thesis invokes a paradigm, the reciprocal constraints paradigm, that was designed to leverage the interdependence between the social and cognitive levels of scale for the purpose of building cognitive and social simulations with better resolution. In addition to explaining our thesis, we provide the current research context, a set of issues with the thesis and some parting thoughts to provoke discussion. We see this work as an initial step to motivate both social and cognitive sciences in a new direction, one that represents some unity of purpose and interdependence of theory and methods.

1 Introduction

The degree of overlap between cognitive science and generative social science is small despite a shared interest in human behavior and a reliance on computer simulation. The former focuses, largely, on developing computational and formal accounts of human thought, action, performance and behavior with non-trivial incorporation of neurophysiological principles when warranted. The latter approaches the question of understanding social structure and dynamics using computational and formal accounts that implement simple agents (what we call *sans cognitive*) in social contexts. We submit that the dearth of interdisciplinary work between these disciplines does not serve either well. Our central thesis, the

[6] The research is (partially) based upon work supported by the Defense Advanced Research Projects Agency (DARPA), via the Air Force Research Laboratory (AFRL). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, the AFRL or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

Resolution Thesis, is this: *the correct resolution of both cognitive and social systems depends on mutual constraints between them in the sense that the dynamics and structure of one system should inform the theoretical nature of the other*. We mean this in the context of theory development and related applications in both cognitive science and generative social science. The method implied by this thesis is what we call the reciprocal constraints paradigm–a bi-directional dependence across levels of scale w.r.t. their respective parameter specifications. [7]

Our thesis implies two claims. First, cognitive models should be able to match and predict the real world dynamics of social systems when embedded in social simulation, and, if not, the cognitive model should be questioned. Second, if an agent-based simulation is not informed by cognitive first principles, it will fail to generalize its account of the dynamics of social system to new situations.

In what follows, we will 1) flesh out the details of the reciprocal constraints paradigm, 2) provide some prior work that is directly relevant to our thesis and puts it in context of recent research, 3) address issues and their potential mitigation, and, 4) close with some brief, but potentially provocative suggestions. We deliberately exercised a narrow focus using the ACT-R cognitive architecture as our vehicle of rhetoric, partly because it reflects our expertise, and partly because this architecture is comparatively well suited for integration of both neural and social constraints. Cognitive architectures, as opposed to any cognitive model, capture what agents, in the scheme of generative social science, are supposed to do–make adaptive decisions that affect the environment.

2 The Reciprocal Constraints Paradigm

Figure 1 captures the core components of the reciprocal constraints paradigm: multiple levels of scale, multiple potentials for model types at each level, and, the constraints among levels. To understand the paradigm, it will be useful to imagine a potential implementation. Consider a modeling problem in which there is a simple social system (e.g., a multi-player repeated economic game). The cognitive model is developed, with some consideration for key neural processes, call this $\mathrm{CM}^{[1]}$, and without direct comparison to newly generated individual-level data sources (e.g., running single-subject experiments in pseudo-game like contexts). $\mathrm{CM}^{[1]}$ is then implemented in a social network graph that controls information flow (e.g., knowing past decisions of other players) and, given some other parameterizations, a simulation of the multi-player repeated game is conducted; call this $\mathrm{SM}^{[1]}$. Then, $\mathrm{SM}^{[1]}$ data is aggregated in some way isomorphic to human data in a similar experimental paradigm and an accuracy/error/confidence metric is computed, call it $\mathrm{Constraint}^{[1]}$ to map onto Figure 1. (Notice, at this point, the only direct comparison to human data was at the social system level.) $\mathrm{Constraint}^{[1]}$ would then be used–in an undefined way at this point–to change some aspect of the cognitive model, either directly within the cognitive level or,

[7] Because cognitive systems are sometimes tightly yoked to neurophysiology, we consider three levels as central to our thesis: neurophysiology, cognitive architecture, and social systems

potentially, through the neurophysiological level. Let's imagine that it makes sense to consider neurophysiological processes as the next step, a step we call Interpret$^{[1]}$ to map onto Figure 1. Now, a set of targeted neurophysiological measurements are captured by running single-subject experiments in pseudo-game like contexts which yields insight into a potential missing abstraction of neurophysiological process in the cognitive model, which we call Abstraction$^{[1]}$. The cognitive model is then refactored to incorporate Abstraction$^{[1]}$ and the process is repeated by another simulation using the next generation of the cognitive model CM$^{[2]}$. Note, this example provides only one of an infinite set of paths; the paths may be consequential to the final model and could include integration of human data at one or more points.

A fundamental part of the paradigm is the acknowledgment that scaling up from the cognitive level to the social level is different, in principle, compared to the scaling up from the neural to cognitive level. The former transition instantiates multiple isomorphic and interdependent cognitive models as a simulated system. The latter, in contrast, abstracts information processing functionalities that are assumed to be interdependent but different in nature (i.e., different functions). This is an important difference in light of what a constraint actually means.

3 Relevant Prior Work

In cognitive science, there are several relevant threads of work that address aspects that are important for the *Resolution Thesis*, e.g., on multi-agent systems [1], computational organizational theory [2], computational social psychology [3]. These efforts, however, were not directly concerned with the *Resolution Thesis*. Instead, these efforts, in the main, attempted to provide both more accurate predictions of social system level behavior and explanations that were grounded in cognitive first principles. In this section, we focus on ACT-R to illustrate efforts to either inform ACT-R from neurophysiology or use implementations of ACT-R as the agent definitions in a social simulation. These efforts, we hope, will illustrate how the state-of-the-art in infusing social simulation with cognition contrasts with the *reciprocal constraint paradigm*. Further, we offer a glimpse into how generative social science has conceptualized the integration of cognitive first principles into the behavior of agents to date.

3.1 The ACT-R Cognitive Architecture

Computational modeling aims to quantitatively capture human cognitive abilities in a principled manner. Cognitive architectures are computational instantiations of unified theories of cognition that specify the structures, representations and mechanisms of the human mind. Cognitive models of any given task can be developed using a cognitive architecture as a principled implementation platform constraining performance to the powers and limitations of human cognition. Cognitive models are not normative but represent Simons (1991) theory of bounded

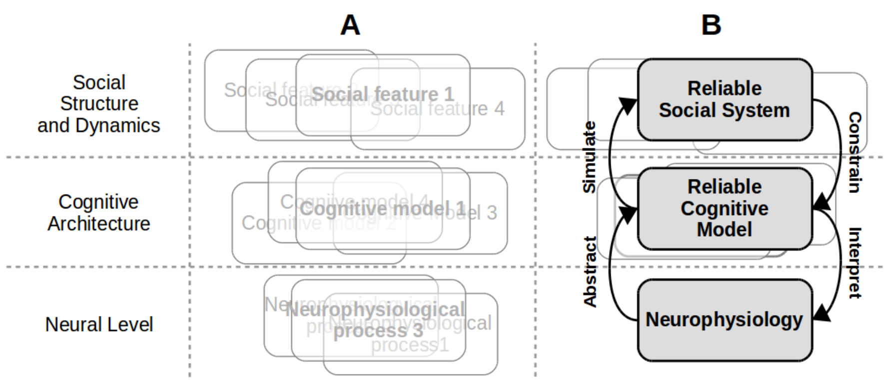


Fig. 1. ***The Architecture and Implementation of the Reciprocal Constraints Paradigm*** Each row represents a level of scale (as labeled in the left-most column). Column A is notational for the degree of variety of potential types of neural processes and cognitive models that could be constructed to capture a phenomenon and the types of features in the social space (e.g., peer-network)–i.e., it captures the feature/model space of a particular implementation. Column B shows the implementation of the reciprocal constraints paradigm; each arrow represents a kind of constraint: *Abstract*–abstraction of neural processes to cognitive processes; *Simulate*–simulating social systems in which humans behavior is defined as a cognitive architecture; *Constrain*–the feedback signal from the accuracy of the social simulation w.r.t. to empirical measurements on human systems; and, *Interpret*–refinement of the selection of neural processes that are implicated in the cognitive model. The former two constraints we call *upward constraints*; the latter are called *downward constraints*. Implementation of the paradigm will require iteration between the feature/model space and the simulation of social and cognitive models. There may be potential for automation of this paradigm once it is well developed.

rationality[4], and can also represent individual differences in knowledge and capacity such as working memory. Cognitive models can be used to generate quantitative predictions in any field of human endeavor.

ACT-R is a highly modular cognitive architecture, composed of a number of modules (e.g., working memory, procedural and declarative memory, perception and action) that operate in parallel asynchronously through capacity-limited buffer interfaces. Each module in turn consists of a number of independent mechanisms, typically including symbolic information processing structures combined with equations that represent specific phenomena and regularities (e.g., power law of practice and forgetting, reinforcement learning). Most notably, the architecture includes a number of learning mechanisms to adapt its processing to the structure of the environment. ACT-R has been applied to model human behavior across a wide range of applications (see ACT-R web site for over a thousand publications), ranging from basic experimental psychology paradigms to language, complex decision making, and rich dynamic task environments. The combination of powerful computational mechanisms and human capacity limitations (e.g., working memory, attention, etc.) provides a principled account of both human information processing capabilities as well as cognitive biases and limitations.

3.2 Neurophysiogical Constraints in ACT-R

The development of ACT-R has been guided and informed, in recent years, by the increased understanding of the computational mechanisms of the brain. For example, independent modules have been associated to specific brain regions and circuits, and this correspondence has been validated multiple times through fMRI experiments. The detailed computations of crucial ACT-R components can also be derived from the neural mechanisms they abstract. For instance, the latency to retrieve declarative information from long-term memory can be derived from the dynamics of the integrate-and-fire neural model [5], and the mechanisms for skill acquisition can be derived from reinforcement learning [5] as well as from the simulation of the large-scale effects of dopamine release in the fronto-striatal circuits [6]. In fact, the modularity of ACT-R permits to easily abstract and integrate lower-level neural principles within the architecture. While this approach does not grant the full flexibility of large-scale neural simulations, it has been repeatedly shown to be very effective in capturing features of human behavior that would otherwise have remained unexplained, while at the same time maintaining the computational parsimony of a cognitive symbolic architecture. For example, implementing the dynamics of memory retrieval permits to capture a variety of decision-making effects and paradoxes, beyond those explained by current mathematical models [7]. The modularity of ACT-R also permits to regulate the degree of fidelity of a module to its biological counterpart, without affecting the entire architecture. As an example, Stocco[8] has shown that the competition between the direct and indirect pathways of the basal ganglia can be captured by splitting production rules into opposing pairs. This procedure captures the cognitive effects of Parkinsons disease, and provides a way to model individual differences in decision-making [8] and cognitive control [9] that are due to individual differences in dopamine receptors in the two pathways. See Figure 2. This is an example of additional mechanisms that can be added to ACT-R to incorporate further biological details (i.e., the *abstract* constraint in Figure 1).

3.3 Social Simulation with ACT-R Agents

To study the dynamics of simple systems, work using ACT-R has focused on iterated two-player games, including both adversarial games (e.g., paper-rock-scissors, pitcher-batter in baseball) and social dilemmas allowing both cooperation and competition dynamics such as Prisoners Dilemma and Chicken Game [10–12]. Even in such simple systems, we have observed the emergence of complex effects such as bifurcations and stochastic resonance [13]. To scale up to more complex yet regular systems, we have modeled the emergence of group consensus and choice differentiation in networks of a few dozen nodes on tasks such as consensus voting and map coloring, respectively, and observed phenomena such as sensitivity to network rewiring parameters [14]. To study complex cognition in complex systems, we have designed and implemented an information foraging task called the Geogame that involves cooperative and competitive problem solving and have observed effects including sensitivity to network topology and tradeoffs between perceptual and memory strategies [15]. Clearly, this work represents well the *simulate* constraint in Figure 1.

A common pattern in models of social interaction using ACT-R has been to ground agent decisions in previous experiences, whether explicitly in the form of memories or implicitly by reinforcement of existing strategies, as mentioned in the previous section. We will focus here on an example using the former approach, because it has been both more common and more flexible. Models of adversarial interaction usually involve a core capability of detecting patterns in the opponent behavior and exploiting them until they disappear. For instance, playing paper rock scissors involve exploiting the human limitation in generating purely random behavior (the standard game theory solution) by detecting statistical patterns in move sequences. An expectation of an opponents next move can be generated by matching his most recent moves against previous sequences using statistical memory mechanisms. Once a pattern is being exploited, the opponent is likely to move away from it and in turn exhibit new ones, requiring a cognitive system that constantly unlearns previous patterns and learns emerging ones, rather than traditional machine learning systems that are training on a fixed set of inputs and then frozen. In that sense, social simulation is the ultimate requirement for online learning: unlike physical environments which change relatively slowly and can be mastered in a relatively static way, social interactions (especially competitive and adversarial interactions), as they involve other cognitive entities, are endlessly evolving and require constant learning and adaptivity.

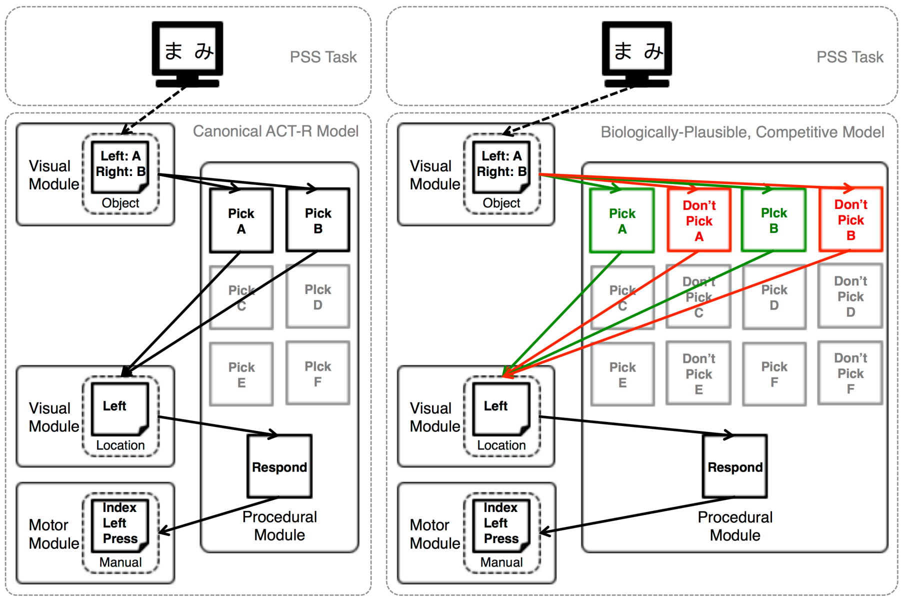


Fig. 2. ***An example (taken from [8] with permission) of how neurobiological constraints can be incorporated in a cognitive architecture.*** The two panels illustrate two alternative ways to implement a forced choice task with six possible options (A through F) in ACT-R. **(Left Panel)** A canonical ACT-R model, in which each option A...F is associated with a single, corresponding production rule (Pick A Pick F). In this model, the expected value of the different options is encoded as the expected utility of each production rule. The utility of each rule is learned through reinforcement learning in ACT-Rs procedural module, which is associated with the basal ganglia. However, the lack of biological plausibility in ACT-Rs procedural module prevents the model from capturing the results of the original study. **(Right Panel)** A biologically-plausible version of the same model, in each of the original production rules is split into two opposite actions (Pick A Pick F and Dont Pick A Dont Pick F), whose utilities are learned separately. This new version abstracts the competition between the direct and indirect pathways of the basal ganglia circuit. When equipped with this biologically-plausible version of production rules, the model can successfully reproduce the results in the neuropsychological literature, as well as capture individual differences in genetics[8] and even correctly predict new findings[9].

3.4 Comparison to the Generative Social Science Approach

Generative social simulation has historically been concerned with the simulation of interacting agents according to simple behavioral rules. We can often equate the outcome behavior of agents to a simple binary action (e.g., you either riot or dont riot) and the behavioral rules that produce this outcome to simple mathematical and logical formulations (e.g., if/else statements, threshold values). We are in debt to the many classic models that made computational social science the field it is today [16–18]. However, there has been some acknowledgment that to gain further insight into social systems, we need to decompose behavior into its underlying cognitive, emotional, and social (interactions) processes. With this, we are beginning to see a slight shift to developing models with more complex agents [19].

In this vein, an approach that has gained some traction is the use of conceptual frameworks that integrate the varied components of agent decision-making processes [20–23]. Such frameworks include BDI (Beliefs, Desires, and Intentions) and PECS (Physical conditions, Emotional state, Cognitive capabilities, and Social status) [24]. In the BDI framework, beliefs are said to be the individuals knowledge about the environment, desires contain information about the priorities and payoffs associated with the current objective, and intentions represent the chosen course of action [25]. BDI agents use a decision tree process which relies on payoff and utility maximizing functions to select goals and to determine the optimal action sequence for which to achieve those goals. The focus on optimality, however, may pose limits on its ability to model the boundedly rational agent and has been criticized for being too restrictive [25]. PECS views agents as a psychosomatic unit with cognitive capabilities residing in a social environment [26]. The PECS framework is flexible due to its ability to model a full spectrum of behaviors, from simple stimulus-response behaviors to more intricate reflective behaviors, which requires a construction of self that necessitates the agent be fully aware of its internal model. By example, Pires and Crooks [23] used the PECS framework to guide implementation of the underlying processes behind the decision to riot, applying theory from social psychology to create the agents internal model and to simulate social influence processes that heightened certain emotions and drove the agent's towards certain actions. These frameworks, while helpful for guiding implementation, are not to be considered substitutes for cognitive architectures such as ACT-R. They can, however, provide a meta-framework (sometimes called a macro-architecture) to organize knowledge and skill content in respect to a cognitive architecture (e.g., [27]).

Cognitive architectures and meta-frameworks are fundamentally complementary [28]. Cognitive architectures precisely specify the basic cognitive acts that can be used to compose complex models in a bottom up approach, but provide few constraints to guide those complex structures. Meta-frameworks provide a top down methodology to decompose complex tasks into simpler ones and structure the knowledge required, but do not include a principled grounding for that process. The combination of the two approaches can be achieved in a number of different ways. One approach is to develop integrated environments allowing modelers to flexibly leverage the two methodologies in a way that is best suited to the specific requirements of each application[29]. An alternative is to provide high-level patterns and abstractions that can be formally compiled into cognitive models in a target cognitive architecture[30].

4 Issues and Their Mitigation

4.1 Downward Constraints

Social to Cognitive This issue was laid out plain by Allen Newell about three decades ago [31] in reference to the *social band* (bands in geometric time of $> 10^4$ seconds that represent organizational behavior and other social systems). Newell, thinking in terms of the strength of a system's levels, hypothesized that social bands should be characterized as having weak strength, and therefore may not be computing much at all, in a systematic way. If Newell's surmises are correct, then constraining cognitive architectures from the social band makes little sense. Anderson's Relevance Thesis [32], put forth about a decade later, does not address the operation of social systems in terms of constraining cognitive models; his thesis is more focused on the degree to which understanding lower bands, especially the cognitive (10^{-1} to 10^1 time scale), are implicated in qualities of higher bands, e.g., educational outcomes. So, from the cognitive perspective, there might not be much signal from the social band that could serve as a useful constraint on cognitive architectures.

However, there are potential approaches towards mitigation of this problem, Newell's thesis notwithstanding. Online social communities often exhibit emergent empirical regularities. For instance, the World Wide Web exhibits many regularities including the small world organization of link structure and the distribution of the lengths of browsing paths that users exhibit. The latter has been called the Law of Surfing. Many of these regularities have been modeled at the social level using variants of statistical mechanics. The Law of Surfing [33] observes that the frequency distribution of path lengths (number of Web pages visited) is well fit by an Inverse Gaussian Distribution, that has a long positive tail. The key insight at the social level is that a Web surfer can be viewed as moving around in a kind of space analogous to the Brownian motion of a small particle on a liquid surface. In the case of the Web surfer, the movement is in the dimension of expected utility that will be received (or not) when visiting a Web page, where the expected utility from continuing on to the next page is stochastically related to the expected utility of the current page, and the Web surfer continues until a threshold expected utility is reached. This is modeled as a stochastic Wiener process. But, the Law of Surfing can also be predicted from Monte Carlo simulations with ACT-R agents [34]. In contrast to the stochastic social models, these finer-grained ACT-R agents can make predictions for specific Web tasks at specific Web sites, which can be used to predict and engineer improvements [35]. However, the emergence of the Law of Surfing from the ACT-R agent simulations is seen as constraint on the cognitive models.

In short, the social band, at least in some domains, does have structure that could constraint cognitive modeling efforts. A question that remains is to what degree will it be possible to develop general methods across the varieties of social domains for the purpose of constraining cognitive models.

Cognitive to Neurophysiology The downward Interpret arrow in Figure 1 could seem paradoxical, given that the underlying neural level is often taken as the ground truth of the entire system. Neurophysiological findings, however, are often only imperfectly understood. For instance, the existence of basal ganglia projections outside of the frontal lobe was considered impossible for a long time until recently [36]. Even when our grasp of neurophysiology is solid, cognitive architectures can be helpful in providing a functional interpretation to existing data by focusing on the computational integration of different circuits, that is, answering the question of what does this circuit do?. The most famous example in this sense is the interpretation of the activity of dopamine neurons in terms of reward prediction error signals in reinforcement learning (RL)[37]–an interpretation that borrowed from a decades-old AI theory (temporal difference learning: [38]) to solve decades of seemingly inconsistent empirical findings on the role of dopamine [39, 40]. Incidentally, this example perfectly illustrates how the Interpretation is further aided by the use of a comprehensive architecture on an agents behavior, such as that provided by RL agents. In our case, the adoption of a single architecture (such as ACT-R) to create multiple models provides the unifying framework to interpret neurophysiological data. The fact that the activity of the same neuronal process must be interpreted in the same way across multiple models of different tasks provides additional constraints to maintain the interpretation consistent.

4.2 Upward Constraints

Parsimony and Generative Social Science By uncovering some new relationship or testing some stylized hypothesis of social phenomena many classic agent-based models (e.g., [16, 18]) have demonstrated the value of modeling simple (*sans cognitive*) agents. For instance, Reynolds [41] illustrates how three simple rules of behaviors can result in the emergence of the collective behavior of a flock of birds – what looks like the highly coordinated actions of a "leader" is actually the result of three simple rules. [8] These models and many others in the computational social sciences adhere to parsimony, or keeping the model simple such that the model has just enough of right features and no more, as a main guiding principle [42]. Arguments for this approach stress the intuitive and interpretive appeal of such models [42, 43]. The purpose of the model may also dictate that the model be parsimonious (e.g., [41]). In short, parsimony in

[8] ABMs, however, can range in abstraction, from the stylized models just described to empirically-driven models; although the latter in no way implies incorporation of cognitive constraints.

respect to simple agents has served well as a strategy in generative social science. It is natural, then, to ask if cognitive modeling breaks with this notion of parsimony in modeling social systems.

We think the issue of parsimony in generative social science does not imply anything particular about the use of cognitive architectures in social simulations. Parsimony implies that model simplicity is considered in conjunction with how well a model matches empirical findings. Thus, the issue of whether to include cognitive agents, as defined in the reciprocal constraints paradigm, is largely an empirical issue. We offer that cognitive constraints may provide the right model and thus improve the degree to which a social simulation matches empirical findings. Moreover, because cognitive models inherit mechanistic constraints from cognitive architectures, they might actually end up being more parsimonious than agent-based models without such constraints.

4.3 Mere Parameter Optimization?

To deal with the challenges of scaling up cognitive models beyond the scale of tasks in the cognitive band (seconds to minutes) to tasks in the social band (weeks to months), Reitter and Lebiere [44] formulated a methodology called *accountable modeling*. That approach is not only a technical solution to scaling up the cognitive architecture but also a scientific commitment to an approach that explicitly states which aspects of the model are constrained by the architecture and which are free parameters to be estimated from data. This commitment helps determine which aspects of the social-scale simulation reflect the cognitive mechanisms and can be assumed to generalize, and which have been parameterized to reflect aspects of the situation not constrained by first principles, and thus will need to be estimated from data in new situations. Such an approach actually results in simpler, more transparent models that are explicit about their parameters rather than trying to camouflage them under a mechanistic veneer.

5 Closing Thoughts

Crossing levels of scale or analysis inevitably takes one near to deep scientific issues that echo notions pointed out 50+ years ago in Simon's "Architecture of Complexity" paper [45] (see also [46] for similar early example). Our thesis goes counter to Simon's notion of near decomposability in that it puts social structure and dynamics in the realm of convergent evidence for a cognitive theory. In this spirit, we will leave the reader with one final comment.

We see social systems as distributed and symbolic. Thus, insight into them and predictions about them should come through a distributed symbolic system–i.e., a social simulation of interacting cognitive architectures. This argument is not meant to imply that sub-symbolic processes are not part of human information processing, but only to mean that social interactions operate via symbols. For the purposes of simulating social systems, observed social structure and dynamics should be generated from the first principles of interactive artificial symbol systems.

References

1. Sun, R.: Cognition and multi-agent interaction: From cognitive modeling to social simulation. Cambridge University Press (2006)
2. Prietula, M., Carley, K., Gasser, L.: Simulating organizations: Computational models of institutions and groups. Volume 1. The MIT Press (1998)
3. Vallacher, R.R., Read, S.J., Nowak, A.: Computational Social Psychology. Routledge (2017)
4. Simon, H.A.: Bounded rationality and organizational learning. Organization science **2**(1) (1991) 125–134
5. Anderson, J.R.: How can the human mind occur in the physical universe? Oxford University Press (2007)
6. Stocco, A., Lebiere, C., Anderson, J.R.: Conditional routing of information to the cortex: A model of the basal ganglias role in cognitive coordination. Psychological Review **117**(2) (2010) 541–574
7. Gonzalez, C., Lerch, F.J., Lebiere, C.: Instance-based learning in dynamic decision making. Cognitive Science **27**(4) (2003) 591–635
8. Stocco, A.: A biologically plausible action selection system for cognitive architectures: Implications of basal ganglia anatomy for learning and decision-making models. Cognitive Science **42** (2018) 457–490
9. Stocco, A., Murray, N.L., Yamasaki, B.L., Renno, T.J., Nguyen, J., Prat, C.S.: Individual differences in the simon effect are underpinned by differences in the competitive dynamics in the basal ganglia: An experimental verification and a computational model. Cognition **164** (2017) 31–45
10. West, R.L., Lebiere, C.: Simple games as dynamic, coupled systems: Randomness and other emergent properties. Cognitive Systems Research **1**(4) (2001) 221–239
11. Lebiere, C., Gray, R., Salvucci, D., West, R.: Choice and learning under uncertainty: A case study in baseball batting. In: Proceedings of the 25th Annual Meeting of the Cognitive Science Society., Mahwah, NJ: Erlbaum (2003) 704–709
12. Lebiere, C., Wallach, D., West, R.: A memory-based account of the prisoners dilemma and other 2x2 games. In: Proceedings of International Conference on Cognitive Modeling, NL: Universal Press (2000) 185–193
13. West, R.L., Stewart, T.C., Lebiere, C., Chandrasekharan, S.: Stochastic resonance in human cognition: Act-r vs. game theory, associative neural networks, recursive neural networks, q-learning, and humans. In: Proceedings of the 27th Annual Conference of the Cognitive Science Society., Mahwah, NJ: Lawrence Erlbaum Associates (2005) 2353–2358
14. Romero, O., Lebiere, C.: Simulating network behavioral dynamics by using a multi-agent approach driven by act-r cognitive architecture. In: Proceedings of the Behavior Representation in Modeling and Simulation Conference. (2014)
15. Reitter, D., Lebiere, C.: Social cognition: Memory decay and adaptive information filtering for robust information maintenance. In: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, AAAI (2012) 242–248
16. Schelling, T.C.: Models of segregation. The American Economic Review **59**(2) (1969) 488–493
17. Axelrod, R., et al.: A model of the emergence of new political actors. Artificial societies: The computer simulation of social life (1995) 19–39
18. Epstein, J.M.: Modeling civil violence: An agent-based computational approach. Proceedings of the National Academy of Sciences **99**(suppl 3) (2002) 7243–7250

19. Epstein, J.M.: Agent_Zero: Toward neurocognitive foundations for generative social science. Princeton University Press (2014)
20. Caillou, P., Gaudou, B., Grignard, A., Truong, C.Q., Taillandier, P.: A simple-to-use bdi architecture for agent-based modeling and simulation. In: Advances in Social Simulation 2015. Springer (2017) 15–28
21. Sakellariou, I., Kefalas, P., Stamatopoulou, I.: Enhancing netlogo to simulate bdi communicating agents. In: SETN. Volume 8., Springer (2008) 263–275
22. Malleson, N., See, L., Evans, A., Heppenstall, A.: Implementing comprehensive offender behaviour in a realistic agent-based model of burglary. Simulation **88**(1) (2012) 50–71
23. Pires, B., Crooks, A.T.: Modeling the emergence of riots: A geosimulation approach. Computers, Environment and Urban Systems **61** (2017) 66–80
24. Kennedy, W.G.: Modelling human behaviour in agent-based models. In: Agent-based models of geographical systems. Springer (2012) 167–179
25. Rao, A.S., Georgeff, M.P., et al.: Bdi agents: From theory to practice. In: ICMAS. Volume 95. (1995) 312–319
26. Schmidt, B.: The modelling of human behaviour: The PECS reference models. SCS-Europe BVBA (2000)
27. West, R., Nagy, N., Karimi, F., Dudzik, K.: Detecting macro cognitive influences in micro cognition: Using micro strategies to evaluate the sgoms macro architecture as implemented in act-r. In: Proceedings of the 15th International Conference on Cognitive Modeling. (2017) 235–236
28. Lebiere, C., Best, B.J.: From microcognition to macrocognition: Architectural support for adversarial behavior. Journal of Cognitive Engineering and Decision Making **3**(2) (2009) 176–193
29. Lebiere, C., Archer, R., Best, B., Schunk, D.: Modeling pilot performance with an integrated task network and cognitive architecture approach. Human performance modeling in aviation (2008)
30. Ritter, F., Haynes, S.R., Cohen, M., Howes, A., John, B., Best, B., Lebiere, C., Jones, R.M., Crossman, J., Lewis, R.L., St. Amant, R., McBride, S.P., Urbas, L., Leuchter, S., Vera, A.: High-level behavior representation languages revisited. In: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, AAAI (2012) 242–248
31. Newell, A.: Unified theories of cognition. Harvard University Press (1990)
32. Anderson, J.R.: Spanning seven orders of magnitude: A challenge for cognitive modeling. Cognitive Science **26**(1) (2002) 85–112
33. Huberman, B.A., Pirolli, P., Pitkow, J.E., Lukose, R.M.: Strong regularities in world wide web surfing. Science **280**(5360) (1998) 95–97
34. Fu, W.T., Pirolli, P.: Snif-act: A model of user navigation on the world wide web. Human Computer Interaction **22**(4) (2007) 355–412
35. Chi, E.H., Rosien, A., Suppattanasiri, G., Williams, A., Royer, C., Chow, C., Cousins, S.: The bloodhound project: Automating discovery of web usability issues using the infoscent simulator. ACM Conference on Human Factors in Computing Systems, CHI Letters **5**(1) (2003) 505–512
36. Middleton, F.A., Strick, P.L.: The temporal lobe is a target of output from the basal ganglia. Proceedings of the National Academy of Sciences **93**(16) (1996) 8683–8687
37. Schultz, W., Dayan, P., Montague, P.R.: A neural substrate of prediction and reward. Science **275**(5306) (1997) 1593–1599
38. Sutton, R.S.: Learning to predict by the methods of temporal differences. Machine learning **3**(1) (1988) 9–44

39. Bunney, B.S., Chiodo, L.A., Grace, A.A.: Midbrain dopamine system electrophysiological functioning: a review and new hypothesis. Synapse **9**(2) (1991) 79–94
40. Schultz, W.: Getting formal with dopamine and reward. Neuron **36**(2) (2002) 241–263
41. Reynolds, C.W.: Flocks, herds and schools: A distributed behavioral model. ACM SIGGRAPH computer graphics **21**(4) (1987) 25–34
42. Miller, J.H., Page, S.E.: Complex adaptive systems: An introduction to computational models of social life. Princeton university press (2009)
43. Gigerenzer, G., Todd, P.M., ABC Research Group, t., et al.: Simple heuristics that make us smart. Oxford University Press (1999)
44. Reitter, D., Lebiere, C.: Accountable modeling in act-up, a scalable, rapid-prototyping act-r implementation. In: Proceedings of the 2010 International Conference on Cognitive Modeling. (2010)
45. Simon, H.A.: The architecture of complexity. Proceedings of the American Philosophical Society **106**(6) (1962) 467–482
46. Anderson, P.W.: More is different: Broken symmetry and the nature of the hierarchical structure of science. Science **177**(4047) (1972) 393–396

future internet

Review

Computational Social Science of Disasters: Opportunities and Challenges

Annetta Burger [1,2], Talha Oz [1,2], William G. Kennedy [1,2] and Andrew T. Crooks [1,2,3,*]

1 Department of Computational and Data Sciences, George Mason University, Fairfax, VA, 4400, USA; aburger2@gmu.edu (A.B.); toz@gmu.edu (T.O.); wkennedy@gmu.edu (W.G.K.)
2 Center for Social Complexity, George Mason University, Fairfax, VA, 4400, USA
3 Department of Geography and GeoInformation Science, George Mason University, Fairfax, VA, 4400, USA
* Correspondence: acrooks2@gmu.edu; Tel.: +1–703–993–4640

Received: 20 March 2019; Accepted: 23 April 2019; Published: 26 April 2019

Abstract: Disaster events and their economic impacts are trending, and climate projection studies suggest that the risks of disaster will continue to increase in the near future. Despite the broad and increasing social effects of these events, the empirical basis of disaster research is often weak, partially due to the natural paucity of observed data. At the same time, some of the early research regarding social responses to disasters have become outdated as social, cultural, and political norms have changed. The digital revolution, the open data trend, and the advancements in data science provide new opportunities for social science disaster research. We introduce the term computational social science of disasters (CSSD), which can be formally defined as the systematic study of the social behavioral dynamics of disasters utilizing computational methods. In this paper, we discuss and showcase the opportunities and the challenges in this new approach to disaster research. Following a brief review of the fields that relate to CSSD, namely traditional social sciences of disasters, computational social science, and crisis informatics, we examine how advances in Internet technologies offer a new lens through which to study disasters. By identifying gaps in the literature, we show how this new field could address ways to advance our understanding of the social and behavioral aspects of disasters in a digitally connected world. In doing so, our goal is to bridge the gap between data science and the social sciences of disasters in rapidly changing environments.

Keywords: disasters; computational social science; crisis informatics; disaster modeling; Web 2.0; social media; big data; volunteered geographical information; crowdsourcing

1. Introduction

The frequency of disasters is on the rise [1], and projections suggest the risk will increase in the future [2]. However, progress in the field of disaster research continues to be challenged by a multifaceted context with psychosocial, socio-demographic, socioeconomic, and sociopolitical dimensions and associated shifting definitions of what qualifies as a disaster (e.g., [3–6]). These complexities lead to a broad range of questions pertaining to the social, psychological, cultural, political, and economic impacts. What are the interacting causal factors that lead to disasters? Who is vulnerable to disasters? Which factors contribute to their vulnerabilities and to what extent? How can practitioners apply lessons from research to prevent and mitigate disasters? In addressing these questions, traditional social science methods for the collection of social data garnered from interviewing and surveying during or immediately after disasters remains a challenge. Research conclusions in the field are limited due to a paucity (or "unobservability" [7]) of data (which is discussed more throughout this paper and especially in Section 5.1) and the fact that data gathered from disaster events are heavily context dependent and extremely heterogenous [8]. Critiques of the social science side of disasters include Tierney et al.'s [9] identification of deficiencies in the knowledge base and a recommendation to find more evidence to support widely believed findings and Quarantelli's [10] work noting the disturbing deficiency of empirical disaster studies and the broad acceptance of empirical generalizations that rely on small or weak datasets. However, fresh methods and data sources are emerging from new technologies in data analysis and computational modeling and the fields of crisis informatics and computational social science (CSS). In this context, the integration of traditional social science theories, innovations in data analysis, and developments in computational modeling offer notable approaches that can address current gaps in disaster research and provide opportunities to advance the field.

In addition to the inadequacy of large, empirically validated datasets, disaster research is affected by continuously adapting environments fed by ecological, social, cultural, political, and technological changes [11] and is addressed by multiple research disciplines. In this regard, some older research findings may no longer be relevant due to changes in society, culture, technology, etc. (as is discussed in Section 2), and traditional disciplines often do not address questions that investigate these adoptions at the intersections of various social sciences. We argue that, while the research literature of disasters is often structured by discipline, understanding the interacting social processes present in disasters is subsequently challenged by disciplinary stove pipes. The purpose of this paper is to review the existing state of the art in disaster studies and relevant disciplines, identify gaps and commonalities, and discuss how computational models and new forms of data analysis can cross over and break down the traditional disciplinary barriers of the social sciences. A literature review was accomplished by leveraging existing domain literature reviews, backwards snowballing, and extensive key-word searches in Google Scholar using English words: disaster, social science, psychology, anthropology, political science, economics, computational social science, and crisis informatics. Relevant literature in the form of books, journal articles, and conference papers were selected to represent the work of three areas relevant to disaster research, as depicted in Figure 1: social science (sociology, psychology, anthropology, political science, and economics), computational social science, and crisis informatics. Representative works were further culled for exemplar questions of interest, methods used, and theories highlighted. Relevant findings were subsequently compared to identify gaps and commonalities.

As information and communication technologies (ICTs) such as the Internet of Things (IoT), smart mobile devices (including GPS and Bluetooth sensors), and advances in Web 2.0 pervade every aspect of daily life [12–14], they have also become ubiquitous in disaster events (e.g., [15,16]). Coinciding with this is the emergence of big data, innovations in data analysis that are providing us with new ways to explore disasters. Approximately a decade ago, informatics researchers (i.e., computer, information, and communication scientists) coined a term to address this aspect of disaster research—crisis informatics. Building upon Kling's [17] definition of social informatics, we define crisis informatics as the study of the design, uses, and consequences of ICTs in times of crisis. Crisis informatics in this regard approaches behavioral data largely from a technology design perspective and not necessarily for the purpose of studying the underlying social theories that explain the processes leading to observed patterns in disasters. It is primarily interested in designing systems for better disaster management.

Researchers utilizing technological tools who are interested in expanding their work beyond the area of system design can shift their attention from the field of crisis informatics to computational social science (CSS: the study of social science through computational methods). In this domain, they can leverage additional themes and theoretical tools for studying social phenomena in disasters. These include: (1) social information retrieval and data mining, (2) modeling and simulation, (3) social networks and geospatial analysis, and (4) online crowdsourcing and experimentation [18–22]. Not only can CSS provide new data sources and methodologies with the growing availability of information through advances in Internet technologies and the proliferation of the IoT and mobile devices (as is discussed in Sections 3 and 4), it has the potential to bring new theoretical and methodological insights to disaster research (discussed in Section 5). Building on CSS while leveraging what we know of crisis informatics and disaster research, we introduce the computational social science of disasters (CSSD). We define CSSD as an approach to explaining the social dynamics of disasters via computational means by adopting the relevant parts of CSS, social sciences in disaster, and crisis informatics, as depicted in Figure 1. With this approach, researchers can take advantage of the new opportunities in CSSD to advance a better understanding of social phenomena in disasters through a new set of research questions.

In the remainder of this paper, we intend to provide a comprehensive description of computational social science of disasters (Section 5). However, we first provide some background on the three scientific fields with which it overlaps (as depicted in Figure 1). In Section 2, we briefly review the domains and the approaches of each of the traditional social science disciplines to disasters. In Section 3, we describe the other encapsulating field of computational social science. Following this, in Section 4, we discuss crisis informatics and its parent field social informatics, as there have been important developments in these fields that make use of "big crisis data" [23], e.g., social media. CSSD is basically the intersection of these three fields. In the following three sections, we discuss the social sciences in disaster research, computational social science, and crisis informatics that serve as the foundations of CSSD. In Section 5, we discuss the components of CSSD and highlight some exemplar studies that capture certain elements of CSSD along with the challenges and the opportunities it brings to the study of disasters. Finally, in Section 6, we provide a summary of the paper.

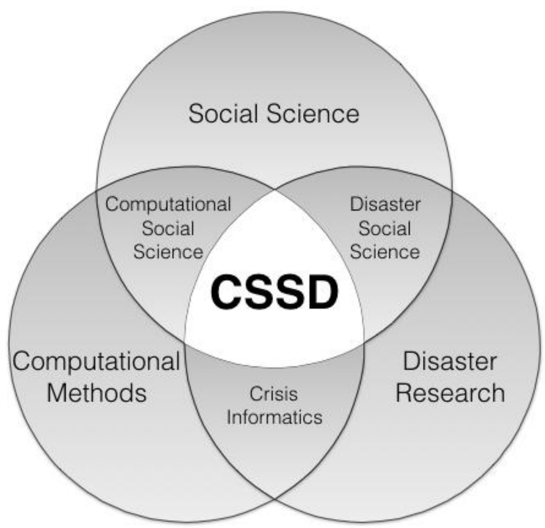


Figure 1. Relation of computational social science of disasters (CSSD) with other fields.

2. The Role of Social Science in Disaster Research

The study of disasters is part of many social science disciplines. Although sociology plays a leading role in disaster research and disaster-related policymaking, studies in this field leverage theories and methodologies from many disciplines (e.g., geography, medicine, industrial organization). Conceptually, rather than research derived from one discipline, research pertaining to disasters is popularly understood in terms of phases: preparedness, response, recovery, and mitigation. For example, in the preparedness phase, policymakers work alongside engineers and researchers to improve disaster planning and warning. Within the response phase, emergence is a core theme of disasters (and complexity science more broadly), and it has been a significant topic of research in disaster science from a variety of disciplines [24]. Disaster recovery, a long and multifaceted process, intersects with the domains of various disciplines, including psychology, economics, political science, tourism, and transportation. Finally, mitigation, which has received special attention since the 2000s, is studied by social geographers as well as environmental and sustainability scientists. How X is affected by disasters and what the impact might be if preparedness, response, recovery, and mitigation factors were varied are questions that could be asked by any discipline studying X, whatever social phenomenon X may be. The century-long history of disaster research tells us that the trans-disciplinary nature of the field has kept evolving over time [25–27].

CSSD is proposed as a subset of the study of social sciences in disasters and is discussed in Section 5. In this section, we provide some background about the foundational findings and the methods of traditional social sciences, specifically sociology (Section 2.1), psychology (Section 2.2), anthropology (Section 2.3), political science (Section 2.4), and economics (Section 2.5). Instead of doing a detailed literature review in each of these disciplines (as it would be beyond the scope of any single paper to address them), we rely on broad reviews and supplement them with additional references as needed. Our summaries of each social science discipline cover an overview, questions of interest, methods used, theory highlights, and relevant findings in the context of this review. Interested readers are referred to the papers cited for more information about the disciplines' long histories, approaches, and contributions to disaster research. It should also be noted that disasters are important topics in other disciplines such as geography, ecology, and medicine. These were not included in this paper for brevity, but a sample of key reviews is provided [28–32].

2.1. Sociology

How do individuals, groups, and societies behave in disasters and times of crisis? What are the underlying social processes? Under what conditions do behavioral patterns of social solidarity arise? How do these differ from those that lead to social conflict? What roles do gender, race, diversity, or economic inequality play throughout disaster planning, response, and recovery? These are just a small set of the questions sociologists in disaster research address [33], and a large body of related empirical work has been codified in works by Barton, Dynes, and Drabek (e.g., [8,25,33,34]). In 1994, Dynes [25] observed, "sociologists in the disaster area have had a much greater influence in the development of science and public policy than in any other [comparable] area." The dominant approach of the sociology of disasters has been event-based and integrated systems theory with the realist assumption that disaster existed at the intersection of physical agents or "hazards", such as earthquakes or tornados and vulnerable people and places [6,35].

The sociological methods of disaster research that have provided the basis of well-understood disaster theories are no different than those of any other sociological enterprise. Phillips [36] outlines four main methods: interviewing, observation, unobtrusive measures (items or traces left behind by people), and visual research through records. In a critique of disaster research methods, Quarantelli [10] recognized a reliance on retrospective and after-action interviewing, rather than systematic field observations, would lead to more reliable evidence. The context of disaster events creates unique methodological challenges, as noted by Stallings [37] and Mileti [38]. Ethical and operational considerations (1) are required to prevent physical and psychological harm to survivors and field researchers and (2) are compressed timelines that prevent adequate time to develop theory, hypothesis, and research instruments. The timetable of disaster events and research schedule is unforeseeable with a high degree of uncertainty surrounding potential subjects and behavioral events.

Two significant literature review papers on the sociology of disasters include Drabek [26], who examined the major contributions of sociology and its methodologies, and Tierney [39], who found traditional disaster research too applied and established that disasters were not distinct events but rather socially constructed by ongoing processes. Early significant findings of sociological research debunked the "disaster myths" that made up much of the cultural frames and media images of disasters, such as themes around social chaos such as panic, shock, ineffectiveness of local organizations, anti-social behavior, and low community morale (e.g., [35,40,41]). In an effort to create an inventory of sociological findings in disaster research, Drabek [8] discussed 146 themes and placed 654 major conclusions of the literature into a typology of system responses in which findings are classified into one of four disaster phases (preparedness, response, recovery, and mitigation) and six social system levels (individual, group, organizational, community, society, and international). Whereas early sociological research focused on the "event," describing disasters as a cycle of stability, disruption, and adjustment, current social constructivist approaches (e.g., [42,43]) shifted the concept of disasters towards social causation [39,44]. For example, Hurricane Andrew and the Chicago Heat Wave of 1992 were not isolated events caused by extreme weather; rather, they were socially constructed by social and economic processes that led to inequalities and created vulnerable populations.

Collective behavior, social control [45] symbolic interactionism [46], and emergent social behavior [47] have been among the popular theoretical orientations in disaster research. In his discussion of "social science research agenda for the disasters of the 21st century," Quarantelli [10] found the earlier accounts narrow and suggested five formulations relevant to disaster research: attribution theory from social psychology, satisficing theory from organizational theory, diffusion studies, network theory, and social capital. Attribution theory and satisficing theory can be applied to decision-making in the context of disasters, and diffusion studies, network theory, and social capital could help provide explanations for behavior arising from social relationships.

2.2. Psychology

Just as sociology of disasters reflects the qualitative nature of sociological studies, the disaster research conducted by psychologists is mostly quantitative, as it is the common methodology in the field of psychology. The questions are formed to understand the human mind regarding preparation for and response to disasters. What leads some people to be better prepared for disasters than others? How can disaster preparedness be encouraged? How does disaster affect the mental health of individuals and their broader community? Psychology literature on disasters can be classified into two, preparedness for risk reduction and post-disaster psychopathology. The latter can be further categorized into four topics: i) empirical predictive (predicts contributions of variables), ii) empirical epidemiological (describes incidence at population level), iii) clinical descriptive (identifies symptoms found in disaster victims), and iv) clinical intervention (describes effectiveness of different intervention approaches) [48]. Methodology in psychology is aimed at identifying and testing the underlying mechanisms of people's behavior and mental health. In disasters, these methods include a combination of screening and diagnostic reports and correlate a variety of psychosocial measures, such as insomnia, perceptions of safety, and changes in the ability to function. These are gathered through observation, interviews, and questionnaires, and they are integrated into structured experimental studies [49].

Several meta-studies reviewed this literature and highlighted the major findings. Rubonis and Bickman [48] examined the relationship between four sets of variables (the characteristics of the victim population, the characteristics of the disaster, the study methodology, and the type of post-disaster psychopathology) by reviewing 52 studies. In a similar effort, Norris [50] and Norris et al. [51] reviewed the post-disaster mental health problems and risk factors in 225 disaster samples (from 132 distinct events experienced by 85,000 individuals) quantitatively studied in the psychology literature. Rubonis and Bickman [48] found a small but positive relationship between disasters and psychopathology, and Norris [50] found post-traumatic stress disorder (PTSD) to be the most common problem occurring in post-disaster studies. Norris et al. [51] found that, among the adults they sampled, factors such as "more severe exposure, female gender, middle age, ethnic minority status, secondary stressors, prior psychiatric problems, and weak or deteriorating psychosocial resources most consistently increased the likelihood of adverse outcomes" [51], while for the youth, family factors had the greatest effect [51].

More recently, Ejeta et al. [52] identified the most common behavioral theories and models applied to disaster preparedness. Reviewing 33 articles on preparedness (including preparedness for disease outbreak, flood, and earthquake hazards), Ejeta et al. [52] found that the most common theories applied in the literature are the health belief model (HBM), the extended parallel process model (EPPM), the theory of planned behavior (TPB), and the social cognitive theories (SCT). In these studies, the main constructs of HBM (perceived susceptibility, severity, benefits, and barriers), EPPM (higher threat and higher efficacy), TPB (attitude and subjective norm), and SCT (cognitive, affective, emotional, and social influences) have been associated with disaster preparedness. However, they also noted the theories were predominately applied to natural hazards and diseases, not man-made hazards. In dealing with the effects of disasters, the review of the resources by Norris et al. [51] found that theories on coping strategies (active outreach, informed pragmatism, reconciliation), beliefs (higher self-efficacy and optimism), social support (social embeddedness, received social support, and perceived social support), and conservation of resources (including objects, conditions, personal characteristics, and energies) help explain the moderators and mediators of psychological effects of disasters. Overall, the psychological literature, while focused on preparedness and post-disaster psychopathology, has also been limited for some uses due to its lack of application to man-made disasters.

2.3. Anthropology

By approaching disaster research with holistic and comparative perspectives, anthropologists study all aspects of human life—environmental, biological, and socio-cultural—as they relate to disasters. Their work focuses on the interconnections between cultural, social, political, economic, and environmental domains to provide explanations for cultural systems in disaster. Anthropologists ask questions such as how do people and cultures understand disaster? How does culture drive socio-cultural processes and responses to disasters? How do these processes interact with the corresponding physical and technical processes? Anthropological studies cross scales from the local to global and back; they explore not only the external physical relationship between human and environment but also the internal meaning that humans produce to understand and interpret their experience. They unravel long-term processes of cultural adaptation to changing social and physical environments as revealed in archeology and history, and reveal power dynamics in the social structures of individuals and groups.

As a result of these analyses, anthropologists have uncovered complex interactions between physical, biological, and sociological systems [53,54] that involve people's adaptations to and manipulations of their physical environment and construction of sociocultural institutions, beliefs, and ethos. As part of a social process, these interactions produce disaster, the event that involves a potentially destructive natural or technological agent and a population under varying conditions of vulnerability [4,53–55]. Anthropological work has shed light on the social production of disasters and the social structures that contribute to vulnerability and risk [4]. Theories of "embodiment" have contributed to a better understanding of how culture affects individuals experiences, along with how they comprehend and cope with traumatic experiences (e.g., [56–60]). Comparative work on multiple cultures has illuminated how different societies respond and adapt to environmental changes [61] and disasters with responsive belief systems and coping strategies [62]. Longitudinal studies have shown how societies cope and adapt through multiple disasters (e.g., [63]).

Since disasters affect every feature of society as well as its relations with the environment and its individuals and communities, anthropology's holistic approach uniquely qualifies the field to study the processes of disaster and interactions that cut across domains. We can look at the findings of anthropological research temporally, studies explaining processes in pre-, (early) response-, and post-disaster phases. Anthropology has given special attention to structural conditions of pre-disaster vulnerability, such as gender inequality, global inequities, endemic poverty, racism, a history of colonial exploitation, imbalances of trade, and underdevelopment [56], and set them in the context of historical processes [64]. In responses to disasters, the themes that have been studied include changes occurring in cultural institutions (e.g., belief systems), within political organizations (i.e., power relations between individuals, the state, and international actors), and within economic systems (e.g., allocation of resources). For the post-disaster phase, a great deal of anthropological work criticizes various actors, including the relief programs, for their top-down, non-flexible strategies in which the affected populations are overlooked (which is in line with the sociological findings), or how media becomes a contested space in which actors try to control the narrative, especially in times of uncertainty [55]. As local and international communities wrestle with issues of environmental change, adaptation, and disaster mitigation, the work of anthropology provides examples of how indigenous and local knowledge can substantially contribute to solutions for community disaster risk reduction and resilience [65–67].

What differentiates anthropology from other social disciplines in disaster research is not only its emphasis on cultural comparison, but also the qualitative, contextual data gathered in the ethnographic methods, such as from interviews, longitudinal participant observations, and linguistic analysis. These contexts of disaster reveal the heterogeneity of disaster experiences in multiple realities and decision-making rationalities. With its holistic approach, the field has the potential to fill methodological and theoretical gaps between the intersecting disciplines that study disaster. In practice, its bottom-up approach balances top-down biases in emergency management and enables the incorporation of local technical knowledge, insight, skills, and needs [56,68]. Conversely, the challenge of this field is that the complex and context-rich studies can become so specific in culture and context as to limit them for general application. Anthropology has also been criticized for privileging local knowledge and problematizing the dominant modes of relief efforts [56].

2.4. Political Science

Political scientists were not present at the foundation of the modern disaster research field, and many were reluctant to study disasters because they viewed disasters primarily as engineering problems, or they maintained the widely held moral stance that there should not be a "politics of disaster" [69]. Others such as Olson [69] argued that disasters are intrinsically political events. Do disasters foster cooperation or conflict? In which condition is one or the other manifested, and why? Although these questions were asked earlier by sociologists [70], more recently, political scientists in the conflict resolution and international relations fields started to investigate it with the greater amount of data that have been collected over the recent decades. Many of the political science studies in disaster research have been quantitative in methodology, and a typical study statistically analyzes decades of data on natural disasters, the incumbents' preparedness and response, and election returns (e.g., [71,72]). Disaster research can be grouped in four subfields of political science: electoral behavior, conflict resolution, international cooperation and humanitarian aid, and political economy [73]. We discuss the first three here and review the political economy aspects of disasters in Section 2.5 under economics.

Elections are proxies for how voters judge incumbent politicians in preparedness and response to disasters, and they are an important factor in the field of electoral behavior (e.g., [74]). At times, politicians are either viewed as merely ineffective in coping with disasters or as causing the disasters. Connelley [75], in a character study of a senator, said the reason for his losing the election was his depiction of a natural disaster: "he couldn't make it rain, and now we've got him down!" Attribution of responsibility is known to be a key issue in political decision-making as Iyengar [76]. Additionally, blame—which is likely to occur in response to disasters—carries far more weight in voting behavior than that of credit [74]. Gasper and Reeves [77] found a negative relationship in the U.S. between disaster damage and the share of incumbent votes for presidents and governors. Another study in this line, which reflects upon citizen competence and government accountability, shows that "voters reward the incumbent presidential party for delivering disaster relief spending, but not for investing in disaster preparedness spending" [78].

In addition to these theories of electoral behavior, theories on conflict resolution are also tested and developed by political scientists. One study on earthquakes argued that disasters increase scarcity of resources, which subsequently provoke frustrations that lead to anger and violence [71]. Some recent studies statistically showed a link between natural rapid-onset disasters and the likelihood of conflict and rebellion (e.g., [71,72]). Nel and Righarts [72], while investigating the impact of natural disasters on civil war, found that "natural disasters significantly increase the risk of violent civil conflict both in the short and medium term" [72]. In another study looking at the root causes of conflict in climate-related disaster, Peregrine [79] used archeological evidence and found an increase in conflict only when leaders tightly controlled access to political authority, such as when using violence to secure support. Others studied more basic dynamics behind conflict behaviors. In this respect, studying the repression dynamics following rapid-onset natural disasters, researchers showed that repression is likely to increase after a disaster, but inflows of aid reduce its intensity [80].

The politics of humanitarian aid and disaster response in the international community involve both the political interests of particular governments, such as U.S. foreign disaster assistance [73], and the need to act cooperatively across traditional sovereign boundaries in international disaster assistance [81]. Political considerations may explain half of all federal disaster relief in the U.S. [82] and may determine whether a president decides to issue a disaster declaration [83]. The conflicting priorities of varying stakeholders often lead to aid policies that create subsequent disasters [84,85]. The need for international disaster cooperation and assistance can arise from civil war and failed states, such as famine in Ethiopia or Africa, or from natural disasters that cross boundaries, such as cyclones and drought. Given the challenges of human-caused climate change, it is arguable that much of today's extreme weather disasters are the result of failed political and economic systems.

2.5. Economics

The economic impact of disasters and incentives for preparedness and response are two major areas of disaster studies. How do disasters affect state and local economies? What are the economic tradeoffs between instituting policies for economic growth versus those for disaster risk mitigation? How is the overall (economic) vulnerability of a population estimated? How can the macroeconomic resilience to disasters, i.e., the ability of an economy to cope with disasters, be measured? These are a small subset of the questions that economists ask regarding disasters and their economic impacts [86]. Traditional social research methodologies exploring the economic impacts of disasters include surveys, global, state, and local measures of GDP, and market and employment reporting.

Examining basic economic indicators from a number of economic literature review papers, Kellenberg and Mobarak [87] found that natural disasters have significant impacts on short- and long-term gross domestic product (GDP), social and human capital, and the labor and real estate markets. A more recent 90-year study of U.S. disasters found that severe disasters do adversely impact economies, but milder disasters have little effect based on measures of out migration, housing prices, and poverty rates [88]. Economists looked at the impact of disasters and found varying effects on specific labor markets (see [89,90] for examples). Of interest to economists are the risk profiles of countries and which would most benefit from disaster risk management policies such as strengthening institutions and building standards, improving insurance markets, reducing corruption, and instituting more advanced warning and emergency response systems [87]. The risk insurance industry and derivative markets have been significant areas of study arising from hazards and risk reduction research (e.g., [91–94]). To find out the extent to which disasters affect the economy of a country, Albala-Bertrand [95] examined the effects of disasters on the growth rate of output of six countries by means of a quantitative macroeconomic model and found that "foreign and public disaster response may be better used to help actual victims and affected activities directly than to proceed on the rather unsound prima facie belief that the economy will be heavily affected by the disaster." Other economists have observed the increasing costs of disasters and studied how improvements to international aid for disaster victims could help protect people or improve economic outcomes [85,96,97].

While many economists have made analyses across multiple disasters, others have developed new economic measures for disasters or have been drawn to specific types of disaster or a case study of disaster, such as Hurricane Katrina. Zahran et al. [98], for example, developed quantitative measures of the mental health impacts of Katrina to explore the economics of disaster risk, social vulnerability, and mental health resilience. Yang [99] explored the impact of hurricanes on the global economy through changes in international financial flows (i.e., financial aid and migrant remittances). Hurricane Katrina has also been studied to find economic reasons for governmental failure in disasters and to measure the socioeconomic costs of disasters. For instance, Shughart [100] expanded the forms of empirical evidence used in identifying the political and economic failures that led to Hurricane Katrina: (1) maintaining existing infrastructure was cheaper than renewing the levees, (2) unlike private corporations, politician = s and bureaucrats have weak incentives, and (3) public policies such as promises of grants, loans, tax breaks, low-interest loans, and insurances had unintended consequences. In another Katrina related study, using U.S. Center for Disease Control's (CDC's) Behavioral Risk Factor Surveillance System (BRFSS) database, Zahran et al. [98] investigated the relationship between individual exposure to hurricanes and poor mental health days and evaluated the economic costs of mental health days on focal populations. Their calculations showed that natural disasters regressively punish disadvantaged population strata [98]. To address pressing problems of dwindling resources arising from environmental change, economic resilience measures are also being developed and introduced for the study of disasters with works by Xie et al. [101] and Rose [102].

2.6. Summary

Traditional social science studies of disasters have provided the foundation of our understandings of disaster, and they continue to contribute research findings and increase our knowledge base. Specifically, we show how sociological research (Section 2.1) has been primarily qualitative, exploring social organizational behavior scaling from individuals to global institutions, and temporally ordered by four main phases of disasters (i.e., preparation, response, recovery, and mitigation). We show how psychological research (Section 2.2) skews more towards quantitative data and is focused on the individuals and the theories applied to preparedness, health, planned behavior, and psychological impacts of disasters. This later topic includes coping strategies, beliefs, social support, and the uses of resources to moderate psychological effects, i.e., post-disaster psychopathology. Anthropology (Section 2.3) confirms the sociological temporal phase approach to disasters with special attention to the structural conditions resulting in vulnerabilities and organizational responses to responses to disasters. Political science (Section 2.4) examines the local, state, and international politics of disaster and comparatively focuses less on the collective preparation for or mitigation of disaster effects on populations. Finally, the economic study of disasters (Section 2.5) focuses on the economic effects of disasters, examining basic economic measures such as GDP, risk management policies, global financial flows, and financial policies. Unfortunately, progress in these sciences has been constrained by their respective disciplinary approaches and methodologies that cannot manage the quantity of events and data available for collection and study in disasters nor fully address the social and physical interactions that cross scales and boundaries. Additionally, while approaches such as case studies allow for in-depth analysis of these events, they provide limited confirmation of theory and are not generalizable to all events. In the next session, we discuss computational approaches applicable to disasters, starting with the general field of computational social science.

3. Computational Social Science

Lazer et al. [18] characterized CSS as an emerging field "that leverages the capacity to collect and analyze data at a scale that may reveal patterns of individual and group behaviors." Computational social scientists educate themselves in how to use and develop computational methods to address social science inquiries in the most effective ways, and CSS introduces new opportunities for collaboration to study the problems of social processes that cut across disciplines. We show in Section 2 how each social science field has its own sets of questions and preferred methods to address them. Paired with the foundational work of social science disaster research, the new methods in the computational social scientist's toolbox, such as computational modeling (as is discussed below), coupled with new types of datasets and corresponding analysis techniques that are now available (e.g., social media, crowdsourced data, digital data at large, machine learning algorithms, and social network analysis) make CSS a uniquely valuable field in addressing the more complex problems of disaster research. We would argue it is the integration of the variety of new datasets and computational analysis tools and modeling under the umbrella of CSS that strengthens the processes of developing and testing social theories. We discuss CSS in four main areas: automated information retrieval and open platforms (Section 3.1), social complexity and simulations (Section

3.2), social networks and geospatial analysis (Section 3.3), and online crowdsourcing and digital field experiments (Section 3.4). For a greater discussion of CSS, readers are referred to [18–21,103,104].

3.1. Information Retrieval and Open Data Systems

Advances in processing technologies have made automated information retrieval standard practice in the social sciences, and these technologies can be used to detect social, behavioral, or economic patterns. In this area, information extraction algorithms are used to collect data from disparate sources, such as census records, economic data, newspapers, and social media, and to conduct data mining and content analysis of verbal data, such as interviews, speeches, and legislative testimony [19,105]. Information retrieval was traditionally defined as "finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers) [106]". Computational social scientists engage in this activity by collecting and analyzing any digital traces that potentially address their social science inquiries (such as elections and international relations, e.g., [107–110]). There are technical challenges in this realm that include evaluation of item similarities, data scalability, and time sensitivity [111]. Salganik [112] comprehensively explores the characteristics of digital data, including its strengths (voluminous, always-on, and non-reactive) and its weaknesses (incomplete, inaccessible, non-representative, drifting, algorithmically confounded, dirty, and sensitive).

To leverage new capabilities in information retrieval, many governments and companies are adopting open data policies that allow researchers to access and study these social data. Prominent examples of such initiatives include Data.gov in the U.S. and OpenKenya [113] in the Republic of Kenya (see [114] for how such data can be used). The ever-increasing popularity of social media, enabled by Web 2.0 technology, is expanding the sources and volume of social data relevant to our daily lives through applications such as Facebook, Twitter, and Instagram, and these open sources are allowing researchers to explore a vast range of topics, including opinions during elections [115], opinions on public health [116], data on disease outbreaks [117], and studies of the connections between people and places [118,119]. The relevance and management of open-source data has become more important than ever, and they are well-positioned to support the quantitative study of disasters through the use of new computational methods, such as machine learning, natural language processing, sentiment analysis, and artificial intelligence [120].

Large-scale social data harvested from a variety of sources can be classified as part of a general category of "observational social data," and these data vary depending on researchers' interests and approaches. Their uses include identification of characteristics or patterns by quantitative and qualitative descriptions of individuals or groups, development of macro-level mathematical models of dynamics in data aggregates, identification of statistical relationships between variables and outcomes, examination of the emergent patterns on the aggregate level, calibration of parameters in computational simulations, inference of social events, and forecasting social phenomena (see [117,121–124] for a range of uses of such large-scale data).

3.2. Complexity and Simulations

CSS is primarily interested in better understanding social phenomena, and it builds on a foundation of existing social science paradigms. Two of the more salient of these paradigms are social complexity and social simulations. Social complexity is a conceptual framework for understanding the increasingly complex interactions of individuals and societies as they interact and adapt to each other and their environment [104]. A complex system is a system of subsystems (i.e., modules or components) whose intra-dependency is much stronger than inter-dependencies [125]. Complex systems can be characterized by distributional or statistical laws—in particular, power laws—and computational tools and new computer language packages, such as those in Python and R, that have made these systems tractable for analysis by a new generation of social scientists. Beyond simple description, power laws and computational tools provide new theoretical perspectives of social phenomena, including self-similarity, scaling, fractal dimensionality, emergence, self-organized criticality, meta-stability, long-range interactions, and universality [104].

In investigating complex adaptive social systems, a promising set of modeling simulation tools known as social simulations has emerged and is often called the third way of conducting social science research [126]. Social simulations and computational models not only allow for discovery of the consequences of theories in artificial societies, but by enforcing formalization in terms of coherent programs, they play a similar role in social sciences as mathematics does in physical sciences [127]. Examples of this technique applied to disasters include [128–131]. Simulation is an alternative to common static modeling approaches in social sciences. Instead of modeling the interactions among variables, the social life is modeled by interacting adaptive agents [132] in an artificial world. In this area, the generative social scientist asks, "How could the decentralized local interactions of heterogeneous autonomous agents generate the given regularity? [133]". Researchers in many social sciences use a wide variety of techniques in social simulations, including agent-based modeling, discrete event simulations, systems dynamics, microsimulations, and cellular automata (see [127,134] for reviews).

Agent-based modeling has become a dominant way of producing social simulations [135]. A distinct advantage of this technique is that it provides the ability to explicitly couple autonomous agents with geographical information when space is relevant to the simulation (e.g., [134,136]). Modeling people and their social systems is not without its challenges (such as dynamic systems, multivariate causation, and validation) [137], but agent-based models (ABMs) can operationalize the characteristics of social complexity, such as heterogeneity, autonomy, explicit space, local interactions, and bounded rationality, in controlled experiments [133]. Since agent-based modeling is well-suited to the object-oriented programming paradigm, they can be easily implemented in any object-oriented programming language, such as Java or Python. Moreover, several programming libraries and frameworks have been developed to facilitate the implementation of ABMs, such as NetLogo [138], MASON [139], GAMMA [140], and MESA [141]. In Section 5, we discuss in greater detail how simulation and these modeling techniques are also applicable to the field of CSSD.

3.3. Social Networks and Geospatial Analysis

Social network analysis and geospatial analysis are two promising computational methods for studying the structures of social and physical spaces within which humans live and interact, especially with the advent of crowdsourced information and the rise of social media [142,143]. Social network analysts see the social world as structured by a web of connected agents tied together by specific relationships [144]. Sociograms, which are graphical depictions of social networks, make social structure visible and tangible. Their representation in matrices and other visualizations on computers allow researchers to examine the properties of large networks. There are different relational types of networks, including directed (digraph), signed (valued), weighted, or multiplex, each of which reflect characteristics of social relations. The levels of analysis of networks at different levels (nodal, dyadic, triadic, n-adic, or at network-level) provide insights about the structure of the social system under study (e.g., on concepts, measures, and properties). Examples of these forms of data analysis are found in disaster studies such as [145–147], which explore the structure of organizational networks and the access and flow of information within networks. Other applications of social network analysis include human cognition and belief systems, decision-making models, organizations and meta-models, supply chains, diffusion, and international relations (for a brief review of these, see [104]).

Geography adds an important dimension to human interactions with their environments. Humans do not live in a spatial vacuum, and social reality is heavily dependent on spatial features. The gap between geography and social sciences is addressed by geographers [21,148]. Developments in geographic information systems (GIS), especially in the fields of spatial databases, positioning technologies, remote sensing, and geo-visualization, have made GIS a common tool in criminology, archaeology, public health, anthropology, economics, demography [21], and disaster research [149]. More importantly, we should note that GIS is not simply a set of technological tools; it brings "spatial thinking" to the social sciences [148] in the form of geographical information science [150]. For example, Hu et al. [151] developed a technique for grid-based tessellation of space that provides a systematic approach for prioritizing areas needing to be mapped by digital volunteers based on information value theory [152]. In this regard, geography has both benefitted from and contributed to computational social sciences [21].

3.4. Online Crowdsourcing and Field Experiments

Internet technologies have opened new frontiers in social collective action and knowledge and the gathering of scientific data in the forms of online crowdsourcing and digital field experiments (see [153–157] for reviews). Crowdsourcing can be understood as the leveraging of information technologies for individual participation in collective processes [158], such as crowdfunding, mapping applications such as Waze, and citizen science data collection efforts (e.g., Christmas Bird Count [159] and Geo-Wiki [160]). Internet platforms such as Waze, Airbnb, and Ushahidi's [161] crisis mapping applications aggregate individual action and knowledge with computational tools that enable individuals to address ongoing social problems. A significant new set of tools in the hands of computational social scientists are micro-tasking sites, such as Amazon Mechanical Turk, that provide a virtual environment for social science experiments.

Experimentation is the primary means for establishing causal relationships, and the cyber world is providing new opportunities and challenges for researchers to conduct large-scale experiments [156]. On the one hand, the "field" of the experiments, i.e., the Internet, narrows down the scope of the interventions only to those applicable in the cyber world, and it limits the ways subjects can be tracked. On the other hand, the increasing variety and prevalence of web applications in daily social life allow experiments with larger and more diverse subject pools in a shorter period of time and with greater participation. Researchers from different fields have conducted both field- and lab-like experiments in cyberspace to test the effects of controlled or natural interventions using various social computing platforms (for a recent survey of online field experiments, see [157], and for lab-like experiments, see [162]). In disaster studies, they have been used to look at crisis communication and emotions in Utz et al. [163], public behavioral responses to disaster information provided online in Liu et al. [164], and purchasing behavior post-Fukushima nuclear accident by Miyata and Wakayatsu [165]. Mao [166] provides examples of how experimental approaches in studying social computing systems can improve the design of such systems and advance our understanding of human behavior in crowd-tasking activities during crisis mapping. The online platforms available for "field" experiments include micro-tasking sites (e.g., Amazon Mechanical Turk [167]), question and answer sites (e.g., Stack Overflow [168]), collaborative encyclopedias (e.g., Wikipedia [169]), social networking sites (e.g., Facebook [170]), e-commerce sites (e.g., eBay [171]), massive open online courses (e.g., Coursera [172]), sharing economy sites (e.g., AirBnB [173]), dating sites (e.g., OkCupid [174]), massively multiplayer online games (e.g., World of Warcraft [175]), or other platforms over which experimenters can exert greater control (e.g., their own sites [176]). Technologies used for interventions in these experiments include emails with different contents [177], websites with different looks [178], bots with different strategies (server-side scripts) [169], and browser extensions with different pop-up behaviors (client-side scripts) [179].

In this section, we briefly review the computational methods developed and used to support social science inquiries with new techniques in computational models and data analysis. We also discuss how CSS provides open-source data, new theories of decision-making, social processes of aggregated behavior, complex adaptive systems, and spatial and network structure, and new experimental methods in online field experiments and social simulations. We now shift our focus to an area where computational methods intersect with disaster research and practice but do not necessarily address traditional social science research questions.

4. Crisis Informatics

The application of new computational methods to traditional fields of science has spawned numerous computational branches, such as digital anthropology, computational linguistics, and biometrics. As its name suggests, crisis informatics is a subfield of informatics, in particular, of social informatics, and can be defined as the study of the design, uses, and consequences of information and communication technologies in times of crisis [10]. As ICTs, IoT, and social media pervade every aspect of our lives, crisis informatics has increasingly become a critical tool in disaster preparedness, response, and recovery [180]. Additionally, information management problems and ineffective use of these technologies have been cited as major factors for failures in disaster management [181].

In the early days of crisis informatics research, many studies employed qualitative methods for both data collection and processing. Researchers manually monitored the ICTs, and they manually curated and classified the information. These studies could be considered mostly descriptive and used formative and interpretivist forms of inquiry (e.g., [182–184]). Recent advancements in computational and mobile technologies, the open-source culture, adoption of open data policies by companies and governments (e.g., [185]), and the popularity of social media platforms have made studies in crisis informatics both qualitative and quantitative, but the more recent studies of crisis informatics have been computational (e.g., [23,186,187]). Palen et al. [182] completed an ethnographic study of a human-induced crisis to understand what aspects of ICT were used, when they were used, and how they were used in the days following the 2011 Virginia Tech shooting event. In that study, Palen et al. [182] conducted 56 on-site, one-on-one, face-to-face interviews and manually monitored online activities of interviewees on social media sites including Facebook, Wikipedia, and Flickr. As a subset of the study, several Facebook groups as well as Wikipedia editors participated in an online, collective problem-solving task to build the list of victims before Virginia Tech officially released the names. The study found that no single online community group was able to come up with a complete list of victim names. Additionally, none of the online lists had false positives, i.e., people incorrectly listed as victims [182]. Another research effort [183] completed a qualitative longitudinal analysis of six disasters as documented by Flickr postings, which was the most popular photo sharing platform at the time of the study. Among the findings was norm development through finding group purpose or tagging nomenclatures as features of photographic contents were compared, categorized, and discussed. Panteras et al. [188] used triangulation techniques with place names paired with geo-location information in tweets and Flickr to delineate the extent of a wildfire, and Hagen et al. [189] used network analysis to identify distinct communities and influential actors from Zika-related tweets.

When we look at the definitions in the literature, we see that the focus of crisis informatics has been on the design and development of ICTs. Crisis informatics:

- "includes empirical study as well as socially and behaviorally conscious ICT development and deployment [182]",
- "strives for socially and behaviorally informed development of ICT for crisis situations [181]",
- "investigate[s] socio-technical interactions that occur during times of extreme crisis with an eye towards developing ways to support the mitigation of suffering [190]", and
- "is dedicated to finding methods for sharing the right information in a timely fashion during [significant crises] [191]."

The theoretical foundations of crisis informatics can be found in social informatics and even earlier in socio-technical systems [192]. Social informatics goes back to the 1980s, when research interests were primarily focused on the impact of computerization on the quality of work [17,193]. Social informatics itself is a subfield of socio-technical systems and is concerned with the relations between social and technical systems [194]. The field of socio-technical systems originated in 1950 from interest in optimizing the productivity of postwar industries and at a time when organizations started to be seen not only as social systems but also as technical systems [194]. We can say that crisis informatics is a study of socio-technical systems that can be used in times of disasters.

Crisis informatics researchers develop new technology capabilities as information and communication technologies advance and needs in disaster preparation and response practices emerge. ICT for disasters can be developed for use by digital volunteers (on- and off-site citizens, see Section 3.1) to allow them to crowdsource (i.e., micro-task) productively, as well as by disaster managers (formal response agencies) to provide them with contextual information to improve decision-making. This aspect of crisis informatics also enables effective coordination and collaboration between emergency responders and digital volunteers [15,192]. For example, the lack of existing road information prior to the 2010 earthquake in Haiti complicated the disaster response, but it also motivated citizen volunteers to use crowd-sourcing applications to share and update road information, as it was encountered on the ground. The large number of citizen volunteers resulted in Haiti becoming one of the best mapped road networks [129]. To fully assess and realize the potential of these technologies in times of disasters, Hughes and Tapia [192] comment that crisis informatics researchers must first understand the ways individuals and organizations "collect, organize, manage, access, share, coordinate, and disseminate information within communities during crisis situations [192]." Understanding how victims, managers, and volunteers obtain and use information constitutes a significant part of crisis informatics.

Crisis informatics as an inherently digital method is continuously incorporating innovative computational methods. Castillo's [23] book, "Big Crisis Data", focuses on methods "for processing social media messages under time-critical constraints." While Castillo [23] focused on computational methods, Meier's [15] book, "Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response", discussed crowdsourcing, the interplay between human curators (of satellite and aerial imagery, social media, and text messages), ICT, and the use of artificial intelligence in times of disasters. Imran et al. [187] also reviewed computational methods and applications for social media data retrieval and processing in the crisis informatics literature. Others (e.g., [186,195]) have discussed the history and the future of crisis informatics and provided a taxonomy of crisis analytics. The field is rapidly growing with continuing improvements to computational techniques (see [196–198] for examples). Of course, these new methods and new big data have not been immune to criticism. Spence et al. [199] addressed the challenges of social media use for collecting data related to a disaster event and the drawing of conclusive inferences from user generated content. The growing number of studies and reviews reflect an increasing interest in the application of crisis informatics to disaster events, particularly in preparedness and emergency response [200].

Interest in integrating crisis informatics with the implementation and use of decision support systems in times of crisis is reinvigorating the application of decision support system research to the disaster field. This is due not only to the growth of data availability and its near real-time nature but also to advancements in decision support systems that increasingly allow the application of knowledge management tools for tactical, operational, and strategic decision-making [201]. Such systems have a long history in urban planning and disaster management (see [202,203] for reviews). For example, systems have been built to aid decision-making during cyclones [204], floods [205], earthquakes [206], evacuations [207], disaster relief [208], and distribution [209,210] more generally. Advances in these areas of data collection and analysis linked to decision support science provide practitioners in the field of disaster and emergency management with not only basic real-time information but also actionable tactical, operational, and strategic knowledge for improved planning and response. As the field evolves, practitioners are beginning to promote necessary conversations among stakeholders to develop standards for best practice, tools, limitations, and ethics of using social media [211,212].

In crisis informatics, the emphasis is on technology, computational methodologies, and data applications rather than explanation and theory. While the field has provided a wealth of new data and analysis to the study of disasters and applications in disaster preparedness and emergency management, it has not put these advantages to use in the advancement of disaster theory. We now discuss how a new field of CSSD can close the gaps between the social sciences of disaster and the computational fields of CSS and crisis informatics.

5. Computational Social Science of Disasters

Empirical and theoretic understandings of disaster can be found at the intersection of social science, computational social science, and crisis informatics research in a combination of theories of social processes, complex adaptive systems, and the information and application of socio-technical systems, as is shown in Figure 1 and discussed in the previous sections. As a subset of computational social science, CSSD brings these domains together in the study of social and behavioral aspects of disasters and related phenomena via computational means. We can now formally define CSSD as the systematic study of the social behavioral dynamics of disasters utilizing computational methods. Computational social scientists and researchers in crisis informatics who are interested in disaster research should draw from and build upon the large body of work in sociology [8] and the other social sciences discussed in Section 2. As has been argued (e.g., [9,213]), there is a need for more evidence to support the social science findings from past quantitative studies of disasters. Through the lens of CSSD, disaster researchers can integrate new computational techniques, methodologies, and theory, which can then be used to test current understandings, develop new theories, and update policy recommendations with respect to disasters. In the remainder of this section, we sketch out how CSSD appears in practice and offer recommendations for paths forward.

The social science of disasters, CSS, and crisis informatics, as demonstrated in our review (Sections 2, 3, and 4), have no distinctive ontologies, but each has a set of preferred methodologies to address their discipline's research questions or goals. These methods are designed to bound each discipline's research questions into tractable hypotheses for testing. However, these practices also isolate the disciplines into silos that are no longer able to address and test the interdependent, nonlinear processes that cross disciplinary domains. The social science of disasters remains largely dedicated to qualitative research and is thus unable to manage the wealth of new data in ICTs and big data or to quantitatively test the complex, nonlinear social processes evident in disaster events. Computational social science and crisis informatics theories and techniques provide data and tools for explaining underlying processes and predicting the outcomes of disaster events utilizing advances in ICTs; however, we would argue their work often provides only a superficial theoretical underpinning (or is disconnected) compared to that found in the traditional social sciences of disasters. In this respect, Palen and Anderson [200] find the marginalization of social science fields within the data science community as troubling. The marginalization of the social sciences is one of the key reasons we feel the need to define CSSD and highlight its potential for advancing disaster research. Unified around a common goal to understand disasters and provide knowledge and information for improved policy decision-making, the three fields contribute unique strengths to the study of disasters. The social science of disasters provides a deep background of theory and explanation for behaviors in disasters, CSS brings theories of complexity and tools for studying complex phenomena, and crisis informatics contributes new forms of data collection and analysis.

With the integration of these fields, we can fully implement our conceptualization of CSSD. As a data-driven, theoretically informed paradigm, CSSD leverages qualitative and quantitative approaches for gathering and analyzing data and developing and testing social theory throughout the stages of disasters, as shown in Figure 2. From the social sciences, theories and conceptual models should guide data collection and analysis and computer modeling and simulation. Computational techniques in CSS and crisis informatics such as digital tracing, online crowdsourcing, and aerial imagery provide the means for gathering data {e.g., crowdsourcing, volunteered geographical information (VGI, [214]), social media, and online field experiments} using information communication technologies and smart mobile devices. These techniques also provide artificial intelligence algorithms and visualization tools in social network analysis (SNA), GIS, machine learning, and deep learning to analyze the data and retrieve evidence to develop, support, or update social theory. Data from ICTs make up the observational components of CSSD that lead to new hypotheses for online experimentation and validation of computational models such as ABMs.

Data collection feeds data analysis, social theories, and computational models, all of which together form the main elements of CSSD. In the context of CSSD, these elements operate in continuous interactions, informing each other in cycles of discovery and explanation. Social theory and models provide us with the conceptual understandings of the processes in disaster, thus they can guide data collection. Data should also inform the models and theories, because the data provide the patterns of disasters (e.g., population displacement, extent of property loss, etc.). Digital data containing various kinds of information (e.g., "big crisis data") are collected from online sources such as news reports and social media platforms and are integrated with more traditional qualitative and quantitative data. These new forms of data guide the formation of hypotheses, which are built upon the findings of social science disaster research. These hypotheses are then operationalized by identifying relevant information in the data and by finding ways to represent and integrate them into the models. The quantitative or computational models are then calibrated and run in simulations. To complete the process, the limitations, generalizability, and implications of the work are examined and inform the next cycle of data collection, theory formulation, modeling, analysis, and testing. This continuous loop of data collection and model refinement is necessary for understanding the processes and phases of disasters and their evolving nature. Although our conceptualization of CSSD has not been fully implemented in practice, we provide a few examples of the interactions among social theories, data collection and analysis, and computational modeling and simulation in Section 5.1. This is followed by Section 5.2. which outlines the challenges and opportunities arising from CSSD.

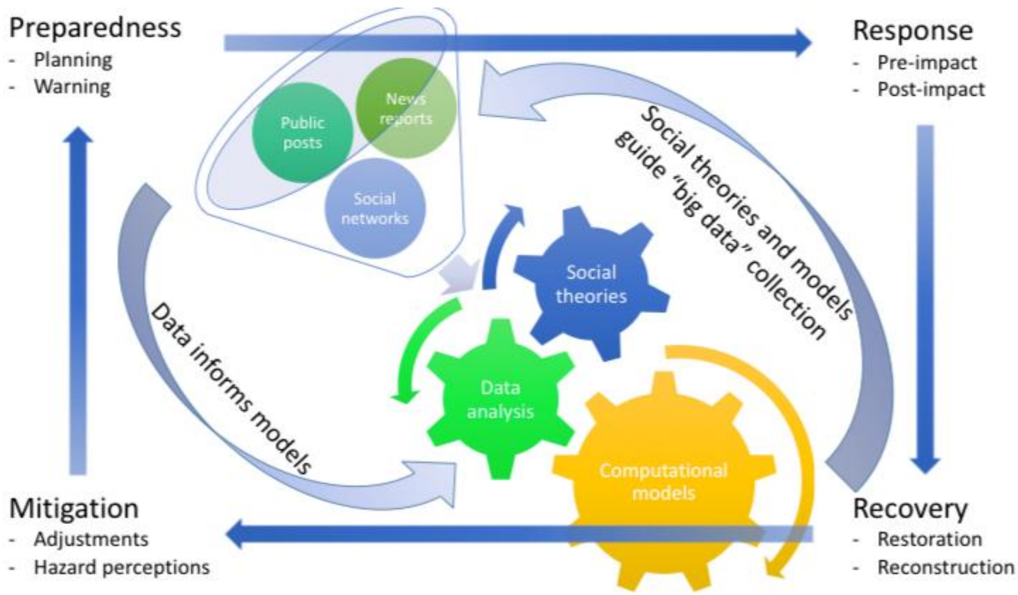


Figure 2. Interactions of data analysis, computational models, and social theory in computational social science of disasters.

5.1. Interactions Among the Components of CSSD

In this section, we discuss the interactions among data, theory, and modeling. With respect to the interactions between data informing theory, a preponderance of the CSS and crisis informatics literature only explores the application of computational techniques and new technologies (as discussed in Sections 3 and 4). However, there are limited examples of traditional scientific methodology with researchers gathering data in the interest of testing theory and producing analytic results. One such example of this disconnect is that of Olteanau et al. [215], who used crowdsourcing to label 1000 tweets from 26 different crisis situations that took place between 2012 and 2013. Their findings identified six broad categories for information communicated over Twitter during disasters (affected individuals, infrastructure and utilities, donation and volunteers, caution and advice, sympathy and emotional support, and other useful information). The observational data were used both to inform the kind of computational analysis performed on experimental datasets and to explore what types of crises elicit specific Twitter user behaviors. This work demonstrates the potential of how data could be used to test theories of human behavior in crises and disasters, such as how role theory explains individual and group behavior, but data analysis studies do not go this far [216].

Data have been used less frequently to develop computational models for social experimentation in simulation. For example, in Jumadi et al. [128], evacuation data were used to improve the social simulation of populations escaping from volcanic eruptions. The data performed multiple functions in the research—calibrating the model, verifying model dynamics, and validating the final model build. As a result of the data used in the modeling process, the researchers were able to refine the model and improve the prediction of the locations where people would evacuate to. Crooks and Wise [129] demonstrated how ICT data in the form of VGI can provide similar functions. Through data analysis of crowd contributed information, they studied the response of populations to variations in aid distribution and subsequently used social simulations to explore how rumors relating to aid availability propagated through the population.

Similar to data informing theory, the published literature showing how social science theory of disasters informs the collection of new forms of data (e.g., social media, VGI) for analysis is relatively scarce; however, we can find some examples. In the tradition of qualitative descriptive research in disaster, Lin and Margolin [217] examined inter-communal emotions and expressions tied to theories such as the social amplification of risk [218]. During the 2013 Boston bombings, they found that the people who had visited Boston or were within close proximity to it had the most predictive power for raising the level of fear, sympathy, and solidarity to Boston. Wen and Lin [219] studied the factors (geographic proximity, media exposure, social support, and gender) of distress (anxiety, sadness, and anger) after the 2015 Paris terror attacks, and compared the immediate acute responses and the ones before the attacks. Glasgow et al. [220] compared the expressions of gratitude for social support received after the 2011 Alabama tornado and the 2012 Sandy Hook school shooting (Newtown, CT), and found that, despite the Alabama victims suffering from a more severe disaster (quantitatively), they received proportionally fewer expressions of support. By examining the microblog posts (i.e., from Twitter) about the Flint water crisis, Oz and Bisgin [221] studied the attribution of responsibility and blame in a man-made disaster. Classic scientific methodologies use theories and observations to develop hypotheses for testing in further observations and experiments. Mao et al. [222] developed an online experiment to test the relationship between team size and productivity (e.g., [223]) in a realistic crisis mapping task. Not only did their work use collected data to update the existing theory on complex tasks, but it also used the existing theory to inform their data collection.

Turning to how computational models can be informed by theory, models are often built to test specific theories, and although they are not yet applied in the area of disasters, theory-based models are prevalent in the field of CSS and conflict crises. In an agent-based model of conflict in Sierra Leone, Pires and Crooks [224] tested Le Billon's [225] theory that the spatial dispersion of a resource (in this case, diamonds) leads to warlordism, secession, and mass rebellion. The theory dictated the selection of data collection from a variety of spatial data sources, including OpenStreetMap [226]. Their model subsequently provided confirmation of the theory using basic bottom-up processes operationalized in the model by enabling agents to choose whether to mine, rebel, or do nothing when varying the spatial dispersion of diamond mines and areas under government control. Traditional social science models often simplify the complex interactions of socio-economic dynamics with linear representations, e.g., increasing education and employment will improve the quality of life in a community. Although there is ample confirmation of these theories, real-world data paint a more complex picture of confounding relationships, such as global economic trade or group unrest. Exploration of these complex interactions and nonlinear relationships has been done in another agent-based model [227] that uses identity and social influence theory [228] to inform data collection and exploration of the emergence of riots. Other applications related to disasters and conflict apply the theory of planned behavior [229] to agent-based models of social-ecological systems, for example, Kniveton et al. [230] and Schwarz and Ernst [231], and these are currently being fit into a framework for mapping and comparing behavioral theories [232].

Traditionally, models were built to test a specific theory. However, with advancements in computation, models are now also used to develop theory, changing the relationship between theory and modeling [126,135]. This is specifically evident in the utilization of agent-based models. Agent-based models have been shown to be suitable for capturing the heterogeneity and the complex interactions between agents and their environments and have made progress in the development and study of theories of complexity in disaster. Perhaps the most well-studied area of agent-based models of complex human behavior in disasters is that of evacuation during fires [122,233,234], earthquakes (e.g., [235–237]), tsunamis (e.g., [238]), hurricanes (e.g., [239]), volcanic eruptions (e.g., [128]), and floods (e.g., [240]). At a finer level of spatial resolution in predicting human behavior, Helbing et al.

[241] modeled escape panic in the spirit of self-driven many-particle systems while [242] focused on more behavioral rich agents that lead to collective egress in evacuations. These and other agent-based modeling frameworks (see [243] for a review) have informed the development of theory on evacuation dynamics in emergencies and disaster. More current work builds on these efforts with further explorations of the complex dynamics in evacuation relating to second-order relationships. Wang et al. [238] simulated evacuation for a near-field tsunami in an ABM, investigating the impacts of decision-time, modes of transportation, and the availability of evacuation paths on mortality rates.

Just as models can inform theory, models can also inform data analysis. Social simulations in the form of computational models have used data from ICTs and created simulated data for analysis. Dawson et al. [240] developed an agent-based model and simulated data to analyze the risks of flooding in different scenarios and provide new insights about flood incident management. Wise's [122] agent-based model of wildfire evacuation demonstrates how, when given a set of parameters for specific scenarios, social simulations in computational models can be used to create data for analysis, predict human behavior, and provide data for policymakers. To assess the longer-term welfare impacts of urban disasters, Grinberger et al. [244,245] made several simulations spanning the three years after an earthquake. They simulated the urban dynamics (residential and non-residential capital stock and population dynamics) using both bottom-up (locational choice for workplace, residence, and daily activities) and top-down (land use and housing price) protocols and analyzed the resulting data to find that low-income groups lose both housing and embedded social support systems. Realistic population synthesis is another important aspect of social simulations of disaster response. Burger et al. [246] proposed a model to synthesize agents using public open data sources such as the U.S. Census Bureau's demographic profile dataset, business patterns dataset, and workflow (LODES) dataset. In this case, the model informs the creation of a synthetic population from census data in social simulation. This review of the interactions among components of CSSD identifies some gaps in current practice and brings us to the opportunities and challenges that need to be overcome to fully operationalize this new field of CSSD.

5.2. Challenges and Opportunities

The field of CSSD encompasses a cycle of interactions in data analysis, computational models, and social theories in the scientific process, and as discussed above, we find examples of this in the literature. In reviewing the work of these areas, we hope it has become clear to the reader that, while each element of CSSD is present, the full conceptualization of CSSD has not been achieved. Two examples of this are data not being used to inform models and few applications of disaster theory in agent-based models. More critically, there is no significant published disaster research that completes a full cycle of interactions in data analysis, computational models, and social theories. Future work needs to close these gaps to take advantage of the opportunities and address the challenges in this new field of study.

As with any emerging field, there are many challenges ranging across a wide spectrum of topical, technical, methodological, and ethical issues. Topically, we see an overreliance on case-study analysis in agent-based models, the disaster sciences, and crisis informatics research, caused in part by the context-dependent nature of disasters and the challenge of sharing data and models. Open science practices can support open exchange of research and allow for generalization to larger theory and replication with platforms sharing data (e.g., Dataverse project [247]) and models (e.g., OpenABM, the Computational Modeling in Social and Ecological Sciences (COMSES, [248]), and GitHub [249]). The nature of most new sources of data (e.g., public polls and social media) is short-term and post-event, and they contribute to understanding the processes of preparedness and response in disaster. CSSD needs to develop strategies for obtaining longitudinal sources of data for the long-term processes evident in mitigation and recovery stages of disaster. From a technical and methodological perspective, challenges in the forms of data, the collection techniques, and the machine learning algorithms all create biases in the data (as discussed in Section 3.1 and [250]). Verification and validation are also problematic. For example, deep learning algorithms suffer from a lack of human interpretability because their machine learning processes operate in a black box and do not have intermediate measures to verify whether they are performing as intended. Validation (demonstrating that results align with real-world outcomes) in this area also suffers from a number of issues, including a lack of high-quality, real-world datasets for comparison to model outputs. The complex subject matter of disasters presents a challenge with the requirement of analyzing heterogeneous variables and multiple interacting processes that prevent the isolation and evaluation of specific actors and processes.

The greatest opportunity in CSSD is the wealth of data sources now available to researchers. Big data and ICT are providing new data sources in the forms of online data collection, social media, and VGI. These sources enable the quick mapping of roads and geographic terrain of disaster event areas, individual reports of events on the ground, and more sophisticated online data collection applications and organizations, such as Ushahidi [161], Missing Maps [251], and Humanitarian OpenStreetMap Team (HOT [252]), that can now be implemented during disaster events. These data opportunities can be expanded with decision support science for improved decision-making in agent-based models such those used for wildfire training, incident command, and community outreach [253]. For instance, SimTable was used in the 2016 Sand Fire in California [254]. Not only are these platforms being used to inform policy decision-making on aid, but they also provide easier post-event data collection using the footprints of digital activities. Because the collection of disaster information can be undertaken post-event and far from the event's location, researchers can help address a major limitation in disaster research, "unobservability" [7]. ICT has also opened up a new frontier in social science experimentation through the use of Internet platforms for online field experimentation; examples include Survey Monkey [255] and Amazon Mechanical Turk [256]. Due to the inherent unpredictability of disasters' effects, crisis informatics and other disaster studies are often vulnerable to an overreliance on post-event data. Pre-event data is necessary to establish baselines of social phenomena and event causation. VGI and ICT could be leveraged to gather these data with less cost and conduct digital tracing backwards from the time of any event. Beyond ICT, there are opportunities in the use of new data analysis tools and packages widely available to data scientists that have made problems subject to multivariable causation and complex nonlinear processes both tractable and feasible on individual computer platforms. Artificial intelligence techniques for machine learning (a statistical method for describing a set of data features) and deep learning (a statistical learning that extracts features from raw data) can now be used to create knowledge and have moved beyond the domain of computer scientists into that of social scientists.

Finally, the data science community at large has yet to develop ethical standards for the collection and the handling of human subject data. Current work leverages existing standards in the social sciences, but there are risks and consequences of aggregating this information into big data. Privacy issues arise when analysis from data collected through ICT and social media (e.g., [257–259]) reveal more than what was intentionally provided, such as the identification of vulnerable individuals from the aggregated information. The level of detail available in big data increases the risk of de-identification to human subjects and requires mitigation with privacy and security controls in the use and protection of the data [112,257,260].

6. Summary and Conclusion

In this paper, we explore three research domains that contribute to the modern understanding of disasters—the social sciences of disasters, computational social science, and crisis informatics. Social science lines of inquiry contribute to our fundamental understanding of the social processes and interactions at work in disasters (Section 2). However, disciplinary structures in academic research have prevented analysis of the complex social process that cross traditional boundaries, such as scaling, long-range interactions, and tipping points. In addition, they have not been able to fully take advantage of the increasingly available sources of new data generated by the proliferation of Internet technologies and mobile devices (e.g., social media, volunteered geographical information, digital news, open data, etc.). We introduce CSS (Section 3), the exploration of social science questions through advanced computational techniques, to show how new forms of data analysis and computational models are providing a new lens with which to study the world around us. Moreover, CSS provides new theoretical underpinnings to explore the complexities and the interacting processes seen within disaster studies. We do this because our goal is to close the gap between the crisis informatics (Section 4) and the social sciences of disasters. The idea of CSSD is introduced to merge traditional social science research with advances in CSS and crisis informatics (Section 5), and we discuss the interactions among the data, theory, and modeling components of CSSD in Section 5.1, along with the opportunities and the challenges of this new avenue of disaster research in Section 5.2. CSSD provides a trans-disciplinary approach to the study and management of disasters and moves beyond simply looking at disasters from a technical or social disciplinary perspective.

Closing the gaps separating the social science of disasters, CSS, and crisis informatics, CSSD's foundation in quantitative data collection, processing, simulation, and analysis provides new knowledge at a deeper level with new forms of data and longitudinal evidence. It is important to integrate these lines of inquiry. Techniques such as computational modeling allow us to explore the patterns of and responses to the various phases of disaster, especially in the era of big data, but they do not enable us to explain the processes behind them. The social sciences of disaster contribute theory and explanation for the complex social and environmental processes involved in the construction, mitigation, response, and recovery from disasters, but they do not have the data tools to collect, compute, and analyze the immense volume and potential interactions in disaster data. Together, these lines of inquiry allow for more thorough investigations of the interactions among the development of social theory, data collection and analysis, and computational modeling of disasters. Through CSSD, we are able to leverage advantages from these domains, go beyond disciplinary boundaries, and gain a deeper understanding of the social and behavioral aspects of disasters in a digitally connected world.

Author Contributions: T.O. and A.T.C. developed the topic, and T.O. along with A.B. prepared the initial draft; A.T.C., A.B. and W.G.K. prepared the final draft for submission and all authors approved its content.

Funding: This work has been partially sponsored by the Defense Threat Reduction Agency (DTRA) grant DTRA1–16–1-0043. The Center for Social Complexity at George Mason University also supported this work.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

References

1. The United Nations Office for Disaster Risk Reduction. Disaster Statistics. Available online: https://www.unisdr.org/we/inform/disaster-statistics (accessed on 4 December 2018).
2. Mechler, R.; Bouwer, L.M. Understanding trends and projections of disaster losses and climate change: Is vulnerability the missing link? *Clim. Change* **2015**, *133*, 23–35.
3. Perry, R.W. What is a disaster? In *Handbook of Disaster Research*; Tierney, K.J., Rodríguez, H., Quarantelli, E.L., Dynes, R., Eds.; Springer: New York, NY, USA, 2007; pp. 1–15.
4. Oliver-Smith, A.; Hoffman, S.M. *The Angry Earth: Disaster in Anthropological Perspective*; Routledge: New York, NY, USA, 1999.
5. Kroll-Smith, J.S.; Couch, S.R. What is a disaster? An ecological-symbolic approach to resolving the definitional debate. *Int. J. Mass Emerg. Disasters* **1991**, *9*, 355–366.
6. Quarantelli, E.L. *What Is a Disaster? Perspectives on the Question*; Routledge: London, UK, 1988.
7. Wallace, A.F. *Human Behavior in Extreme Situations; A Study of the Literature and Suggestions for Further Research*; National Academy of Sciences: Washington, DC, USA, 1956.
8. Drabek, T.E. *Human System Responses to Disaster: An Inventory of Sociological Findings*; Springer: New York, NY, USA, 1986.
9. Tierney, K.J.; Lindell, M.K.; Perry, R.W. *Facing the Unexpected: Disaster Preparedness and Response in the United States*; Joseph Henry Press: Washington, DC, USA, 2001.
10. Quarantelli, E.L. A social science research agenda for the disasters of the 21st century: Theoretical, methodological and empirical issues and their professional implementation. In *What Is a Disaster? New Answers to Old Questions*; Quarantelli, E.L., Perry, R.W., Eds.; Routledge: Oxon, UK, 2005; pp. 325–396.

11. Cioffi-Revilla, C. Socio-ecological systems. In *Handbook of Science and Technology Convergence*; Bainbridge, W.S., Roco, M.C., Eds.; Springer: New York, NY, USA, 2016; pp. 669–689.
12. Crooks, A.T.; Schechtner, K.; Day, A.K.; Hudson-Smith, A. Creating smart buildings and cities. *IEEE Perv. Comput.* **2017**, *16*, 23–25.
13. Crooks, A.T.; Pfoser, D.; Jenkins, A.; Croitoru, A.; Stefanidis, A.; Smith, D.A.; Karagiorgou, S.; Efentakis, A.; Lamprianidis, G. Crowdsourcing urban form and function. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 720–741.
14. O'Reilly, T. Web 2.0 Compact Definition: Trying Again. Available online: http://radar.oreilly.com/2006/12/web-20-compact-definition-tryi.html (accessed on 20 February 2009).
15. Meier, P. *Digital Humanitarians: How Big Data is Changing the Face of Humanitarian Responseč*; CRC Press: Boca Raton, FL, USA, 2015.
16. Boulos, M.N.K.; Resch, B.; Crowley, D.N.; Breslin, J.G.; Sohn, G.; Burtner, R.; Pike, W.A.; Jezierski, E.; Chuang, K.Y.S. Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: Trends, OGC standards and application examples. *Int. J. Health Geogr.* **2011**, *10*, 67.
17. Kling, R. What is social informatics and why does it matter? *Inf. Soc.* **2007**, *23*, 205–220.
18. Lazer, D.; Pentland, A.; Adamic, L.; Aral, S.; Barabási, A.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; et al. Computational social science. *Science* **2009**, *323*, 721–723.
19. Cioffi-Revilla, C. Computational social science. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 259–271.
20. Watts, D. Computational social science: Exciting progress and future directions. *Bridge* **2013**, *2013*, 5–10.
21. Torrens, P.M. Geography and computational social science. *GeoJournal* **2010**, *75*, 133–148.
22. Weinberger, S. Web of war: Can computational social science help to prevent or win wars? *Nature* **2011**, *471*, 566–568.
23. Castillo, C. *Big Crisis Data: Social Media in Disasters and Time-Critical Situations*; Cambridge University Press: New York, NY, USA, 2016.
24. Drabek, T.E.; McEntire, D.A. Emergent phenomena and the sociology of disaster: Lessons, trends and opportunities from the research literature. *Disaster Prev. Manag. Int. J.* **2003**, *12*, 97–112.
25. Dynes, R.R.; Drabek, T.E. The structure of disaster research: Its policy and disciplinary implications. *Int. J. Mass Emerg. Disasters* **1994**, *12*, 5–23.
26. Drabek, T.E. Sociology, disasters and emergency management: History, contributions, and future agenda. In *Disciplines, Disasters and Emergency Management: The Convergence and Divergence of Concepts, Issues and Trends in the Research Literature*; McEntire, D.A., Ed.; Charles C Thomas Publisher: Springfield, IL, USA, 2005; pp. 61–74.
27. Quarantelli, E.L. *The Earliest Interest in Disasters and Crises, and the Early Social Science Studies of Disasters, as Seen in a Sociology of Knowledge Perspective*; Working Paper #91. University of Delaware Disaster Research Center: Newark, DE, USA, 2009.
28. Burton, I.; Kates, R.W.; White, G.F. *The Environment as Hazard*; Guilford Press: New York, NY, USA, 1993.
29. Hewitt, K. *Regions of Risk: A Geographical Introduction to Disasters*; Routledge: New York, NY, USA, 2014.
30. Gunderson, L. Ecological and human community resilience in response to natural disasters. *Ecol. Soc.* **2010**, *15*, 18.
31. Arnold, J.L. Disaster medicine in the 21st century: Future hazards, vulnerabilities, and risk. *Prehosp. Disaster Med.* **2002**, *17*, 3–11.
32. Noji, E.K. *The Public Health Consequences of Disasters*; Oxford University Press: Oxford, UK, 1996.
33. Barton, A.H. *Communities in Disaster: A Sociological Analysis of Collective Stress Situations*; Doubleday: Garden City, NY, USA, 1969.
34. Dynes, R. *Organized Behavior in Disaster*; Heath Lexington Books: Lexington, MA, USA, 1970.
35. Quarantelli, E.L.; Perry, R.W. *What is a Disaster? New Answers to Old Questions*; Routledge: Oxon, UK, 2005.
36. Phillips, B.D. *Qualitative Disaster Research*; Oxford University Press: Oxford, UK, 2014.
37. Stallings, R.A. Methods of disaster research: Unique or not? In *Methods of Disaster Research*; Stallings, R.A., Ed.; Xlibris Corporation: Bloomington IN, USA, 2003.
38. Mileti, D.S. Sociological methods and disaster research. In *Sociology of Disasters: Contributions of Sociology to Disaster Research*; Dynes, R.R., de Marchi, B., Pelanda, C., Eds.; Franco Angeli: Milan, Italy, 1987; pp. 57–69.
39. Tierney, K.J. From the margins to the mainstream? Disaster research at the crossroads. *Annu. Rev. Sociol.* **2007**, *33*, 503–525.

40. Wenger, D.E.; Dykes, J.D.; Sebok, T.D.; Neff, J.L. It's a matter of myths: An empirical examination of individual insight into disaster response. *Mass Emerg.* **1975**, *1*, 33–46.
41. Quarantelli, E.L.; Dynes, R.R. *Images of Disaster Behavior: Myths and Consequences*; Preliminary Paper #5; University of Delaware Disaster Research Center: Newark, DE, USA, 1972.
42. Hewitt, K. *Interpretations of Calamity from the Viewpoint of Human Ecology*; Allen & Unwin: Boston, MA, USA, 1983.
43. Blaikie, P.; Cannon, T.; Davis, I.; Wisner, B. *At Risk: Natural Hazards, People's Vulnerability and Disasters*; Routledge: London, UK, 1994.
44. Rodríguez, H.; Barnshaw, J. The social construction of disasters: From heat waves to worst-case scenarios. *Contemp. Sociol.* **2006**, *35*, 218–223.
45. Miller, L.M. Collective disaster responses to Katrina and Rita: Exploring therapeutic community, social capital and social control. *South. Rural Sociol.* **2007**, *22*, 45–63.
46. Norgaard, K.M.; Reed, R. Emotional impacts of environmental decline: What can native cosmologies teach sociology about emotions and environmental justice? *Theory Soc.* **2017**, *46*, 463–495.
47. Aguirre, B.E.; Wenger, D.; Vigo, G. A test of the emergent norm theory of collective behavior. *Sociol. Forum* **1998**, *13*, 301–320.
48. Rubonis, A.V.; Bickman, L. Psychological impairment in the wake of disaster: The disaster-psychopathology relationship. *Psychol. Bull.* **1991**, *109*, 384–399.
49. Norris, F.H.; Galea, S.; Friedman, M.J.; Watson, P.J. *Methods for Disaster Mental Health Research*; Guilford Press: New York, NY, USA, 2006.
50. Norris, F.H. Disaster research methods: Past progress and future directions. *J. Trauma. Stress* **2006**, *19*, 173–184.
51. Norris, F.H.; Friedman, M.J.; Watson, P.J.; Byrne, C.M.; Diaz, E.; Kaniasty, K. 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981–2001. *Psychiat. Interpers. Biol. Process.* **2002**, *65*, 207–239.
52. Ejeta, L.T.; Ardalan, A.; Paton, D. Application of behavioral theories to disaster and emergency health preparedness: A systematic review. *PLoS Curr.* **2015**, *1*, doi:10.1371/currents.dis.1331a8995ced321301466db321301400f321357829.
53. Hoffman, S.M.; Oliver-Smith, A. *Catastrophe & Culture: The Anthropology of Disaster*; School of American Research Press: Santa Fe, NM, USA, 2002.
54. Torry, W.I.; Anderson, W.A.; Bain, D.; Otway, H.J.; Baker, R.; D'Souza, F.; O'Keefe, P.; Osterling, J.P.; Turner, B.A.; Turton, D.; et al. Anthropological studies in hazardous environments: Past trends and new horizons [and comments and reply]. *Curr. Anthropol.* **1997**, *20*, 517–540.
55. Button, G. *Disaster Culture: Knowledge and Uncertainty in the Wake of Human and Environmental Catastrophe*; Routledge: New York, NY, USA, 2010.
56. Henry, D. Anthropological contributions to the study of disasters. In *Disciplines, Disasters and Emergency Management: The Convergence and Divergence of Concepts, Issues and Trends from the Research Literature*; McEntire, D., Ed.; Charles C Thomas Publisher: Springfield, IL, USA, 2005; pp 111–123.
57. Kleinman, A.; Das, V.; Lock, M.; Lock, M.M. *Social Suffering*; University of California Press: Berkeley, CA, USA, 1997.
58. Green, L. Lived lives and social suffering: Problems and concerns in medical anthropology. *Med. Anthropol. Q.* **1998**, *12*, 3–7.
59. Samuels, A. Embodied narratives of disaster: The expression of bodily experience in Aceh, Indonesia. *J. R. Anthropol. Inst.* **2016**, *22*, 809–825.
60. Sword-Daniels, V.; Eriksen, C.; Hudson-Doyle, E.E.; Alaniz, R.; Adler, C.; Schenk, T.; Vallance, S. Embodied uncertainty: Living with complexity and natural hazards. *J. Risk Res.* **2018**, *21*, 290–307.
61. Crate, S.A.; Nuttall, M. *Anthropology and Climate Change: From Encounters to Actions*; Routledge: London, UK, 2016.
62. Oliver-Smith, A. Anthropological research on hazards and disasters. *Annu. Rev. Anthropol.* **1996**, *25*, 303–328.
63. Doughty, P.L. Plan and pattern in reaction to earthquake: Peru, 1970–1998. In *The Angry Earth: Disaster in Anthropological Perspective*; Oliver-Smith, A., Hoffman, S.M., Eds.; Routledge: New York, NY, USA, 1999; pp. 234–256.
64. Barrios, R.E. What does catastrophe reveal for whom? The anthropology of crises and disasters at the onset of the anthropocene. *Annu. Rev. Anthropol.* **2017**, *46*, 151–166.

65. Oliver-Smith, A. Disaster risk reduction and applied anthropology: Disaster risk reduction. *Ann. Anthropol. Pract.* **2016**, *40*, 73–85.
66. Crate, S.A.; Nuttall, M. *Anthropology and Climate Change: From Actions to Transformations*, 2nd ed.; Routledge: New York, NY, USA, 2016.
67. Marino, E.; Lazrus, H. Migration or forced displacement? The complex choices of climate change and disaster migrants in Shishmaref, Alaska and Nanumea, Tuvalu. *Hum. Organ.* **2015**, *74*, 341–350.
68. Ride, A.; Bretherton, D. *Community Resilience in Natural Disasters*; Palgrave Macmillan: New York, NY, USA, 2011.
69. Olson, R.S. Toward a politics of disaster: Losses, values, agendas, and blame. *Crisis Manag.* **2000**, *18*, 265–287.
70. Dynes, R.R.; Quarantelli, E.L. *Community Conflict: Its Absence and Its Presence in Natural Disasters*; Preliminary Paper #18; University of Delaware Disaster Research Center: Newark, DE, USA, 1975.
71. Brancati, D. Political aftershocks: The impact of earthquakes on intrastate conflict. *J. Confl. Resolut.* **2007**, *51*, 715–743.
72. Nel, P.; Righarts, M. Natural disasters and the risk of violent civil conflict. *Int. Stud. Q.* **2008**, *52*, 159–185.
73. Drury, A.C.; Olson, R.S.; Van Belle, D.A. The politics of humanitarian aid: Us foreign disaster assistance, 1964–1995. *J. Polit.* **2005**, *67*, 454–473.
74. Lau, R.R. Two explanations for negativity effects in political behavior. *Am. J. Polit. Sci.* **1985**, *29*, 119–138.
75. Connelley, W.E. *Ingalls of Kansas: A Character Study*; Connelley, W. E.: Topeka, KS, USA, 1909.
76. Iyengar, S. *Is Anyone Responsible? How Television Frames Political Issues*; University of Chicago Press: Chicago, IL, USA,1994.
77. Gasper, J.T.; Reeves, A. Make it rain? Retrospection and the attentive electorate in the context of natural disasters. *Am. J. Polit. Sci.* **2011**, *55*, 340–355.
78. Healy, A.; Malhotra, N. Myopic voters and natural disaster policy. *Am. Polit. Sci. Rev.* **2009**, *103*, 387–406.
79. Peregrine, P.N. Reducing post-disaster conflict: A cross-cultural test of four hypotheses using archaeological data. *Environ. Hazards* **2019**, *18*, 93–110.
80. Wood, R.M.; Wright, T.M. Responding to catastrophe: Repression dynamics following rapid-onset natural disasters. *J. Confl. Resolut.* **2016**, *60*, 1446–1472.
81. Bose, B.P.C. The politics of disasters. *Ind. J. Polit. Sci.* **1994**, *55*, 119–134.
82. Garrett, T.A.; Sobel, R.S. The political economy of fema disaster payments. *Econ. Inq.* **2003**, *41*, 496–509.
83. Downton, M.W.; Pielke, R.A., Jr. Discretion without accountability: Politics, flood damage, and climate. *Nat. Hazards Rev.* **2001**, *2*, 157–166.
84. Sovacool, B.K.; Tan-Mullins, M.; Abrahamse, W. Bloated bodies and broken bricks: Power, ecology, and inequality in the political economy of natural disaster recovery. *World Dev.* **2018**, *110*, 243–255.
85. Carbonnier, G. *Humanitarian Economics: War, Disaster and the Global aid Market*; Oxford University Press: Oxford, UK, 2015.
86. Hallegatte, S. *The Indirect Cost of Natural Disasters and an Economic Definition of Macroeconomic Resilience*; Policy Research Working Paper 7357; The World Bank: Washington, DC, USA, 2015.
87. Kellenberg, D.; Mobarak, A.A. The economics of natural disasters. *Annu. Rev. Resour. Econ.* **2011**, *3*, 297–312.
88. Boustan, L.P.; Kahn, M.E.; Rhode, P.W.; Yanguas, M.L. *The Effect of Natural Disasters on Economic Activity in Us Counties: A Century of Data*; Working Paper 23410. National Bureau of Economic Research: Cambridge, MA, USA, 2017.
89. Haddad, E.A.; Teixeira, E. Economic impacts of natural disasters in megacities: The case of floods in São Paulo, Brazil. *Habit. Int.* **2015**, *45*, 106–113.
90. Kirchberger, M. Natural disasters and labor markets. *J. Dev. Econ.* **2017**, *125*, 40–58.
91. Kern, W. *The Economics of Natural and Unnatural Disasters;* WE Upjohn Institute for Employment Research: Kalamazoo, MI, USA, 2010.
92. Michel-Kerjan, E.O. Catastrophe economics: The national flood insurance program. *J. Econ. Persp.* **2010**, *24*, 165–186.
93. Mechler, R. Reviewing estimates of the economic efficiency of disaster risk management: Opportunities and limitations of using risk-based cost–benefit analysis. *Nat. Hazards* **2016**, *81*, 2121–2147.
94. Gourio, F. Disaster risk and business cycles. *Am. Econ. Rev.* **2012**, *102*, 2734–2766.
95. Albala-Bertrand, J.M. Natural disaster situations and growth: A macroeconomic model for sudden disaster impacts. *World Dev.* **1993**, *21*, 1417–1434.

96. Strömberg, D. Natural disasters, economic development, and humanitarian aid. *J. Econ. Persp.* **2007**, *21*, 199–222.
97. Yasuyuki, S.; Takasaki, Y. Natural disaster, poverty, and development: An introduction. *World Dev.* **2017**, *94*, 2–15.
98. Zahran, S.; Peek, L.; Snodgrass, J.G.; Weiler, S.; Hempel, L. Economics of disaster risk, social vulnerability, and mental health resilience. *Risk Anal. Int. J.* **2011**, *31*, 1107–1119.
99. Yang, D. Coping with disaster: The impact of hurricanes on international financial flows, 1970–2002. *B. E. J. Econ. Anal. Policy* **2008**, *8*, 1–45.
100. Shughart, W.F. Katrinanomics: The politics and economics of disaster relief. *Public Choice* **2006**, *127*, 31–53.
101. Xie, W.; Rose, A.; Li, S.; He, J.; Li, N.; Ali, T. Dynamic economic resilience and economic recovery from disasters: A quantitative assessment. *Risk Anal.* **2018**, *38*, 1306–1318.
102. Rose, A. Defining and measuring economic resilience to disasters. *Disaster Prev. Manag.* **2004**, *13*, 307–314.
103. Jungherr, A. *Analyzing Political Communication with Digital Trace Data*; Springer: Cham, Germany, 2015.
104. Cioffi-Revilla, C. *Introduction to Computational Social Science: Principles and Applications*; Springer: New York, NY, USA, 2014.
105. Crooks, A.T.; Croitoru, A.; Jenkins, A.; Mahabir, R.; Agouris, P.; Stefanidis, A. User-generated big data and urban morphology. *Built Environ.* **2016**, *42*, 396–414.
106. Kramer, A.D.; Guillory, J.E.; Hancock, J.T. Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8788–8790.
107. Tumasjan, A.; Sprenger, T.; Sandner, P.; Welpe, I. Election forecasts with Twitter: How 140 characters reflect the political landscape. *Soc. Sci. Comput. Rev.* **2011**, *29*, 402–418.
108. Bruns, A.; Burgess, J.E. #ausvotes: How Twitter covered the 2010 Australian federal election. *Commun. Polit. Cult.* **2011**, *44*, 37–56.
109. Crooks, A.T.; Masad, D.; Croitoru, A.; Cotnoir, A.; Stefanidis, A.; Radzikowski, J. International relations: State-driven and citizen-driven networks. *Soc. Sci. Comput. Rev.* **2014**, *32*, 205–220.
110. Schuchard, R.; Crooks, A.T.; Stefanidis, A.; Croitoru, A. Bots in nets: Empirical comparative analysis of bot evidence in social networks. In Proceedings of the 7th International Conference on Complex Networks and Their Applications, Cambridge, UK, 11–13 December 2018; Aiello, L.M., Cherifi, C., Cherifi, H., Lambiotte, R., Lió, P., Rocha, L.M., Eds.; Volume 2, pp. 424–436.
111. Manning, C.D.; Raghavan, P.; Schütze, H. *Introduction to Information Retrieval*; Cambridge University Press: Cambridge, UK, 2008.
112. Salganik, M.J. *Bit by Bit: Social Research in the Digital Age*; Princeton University Press: Princeton, NJ, USA, 2017.
113. Republic of Kenya. Kenya Opendata. Available online: https://www.opendata.go.ke (accessed on 14 June 2018).
114. Mahabir, R.; Stefanidis, A.; Croitoru, A.; Crooks, A.T.; Agouris, P. Authoritative and volunteered geographical information in a developing country: A comparative case study of road datasets in Nairobi, Kenya. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 24.
115. Maynard, D.; Gossen, G.; Funk, A.; Fisichella, M. Should i care about your opinion? Detection of opinion interestingness and dynamics in social media. *Fut. Internet* **2014**, *6*, 457–481.
116. Vraga, E.; Stefanidis, A.; Lamprianidis, G.; Croitoru, A.; Crooks, A.T.; Delamater, P.L.; Pfoser, D.; Radzikowski, J.; Jacobsen, K.H. Cancer and social media: A comparison of traffic about breast cancer, prostate cancer, and other reproductive cancers on Twitter and Instagram. *J. Health Commun.* **2018**, *3*, 181–189.
117. Stefanidis, A.; Vraga, E.; Lamprianidis, G.; Radzikowski, J.; Delamater, P.L.; Jacobsen, K.H.; Pfoser, D.; Croitoru, A.; Crooks, A.T. Zika in Twitter: Temporal variations of locations, actors, and concepts. *JMIR Public Health Surveill.* **2017**, *3*, e22.
118. Mearns, G.; Simmonds, R.; Richardson, R.; Turner, M.; Watson, P.; Missier, P. Tweet my street: A cross-disciplinary collaboration for the analysis of local Twitter data. *Fut. Internet* **2014**, *6*, 378–396.
119. Jenkins, A.; Croitoru, A.; Crooks, A.T.; Stefanidis, A. Crowdsourcing a collective sense of place. *PLoS ONE* **2016**, *11*, e0152932.
120. Shah, D.V.; Cappella, J.N.; Neuman, W.R. Big data, digital media, and computational social science: Possibilities and perils. *Ann. Am. Acad. Polit. Soc. Sci.* **2015**, *659*, 6–13.
121. Wang, F.; Cui, P.; Sun, G.; Chua, T.S.; Yang, S. Guest editorial: Special issue on information retrieval for social media. *Inf. Retr.* **2012**, *15*, 179–182.

122. Wise, S. Using Social Media Content to Inform Agent-Based Models for Humanitarian Crisis Response. Ph.D. Thesis, George Mason University, Fairfax, VA, USA, 2014.
123. Birkin, M.; Harland, K.; Malleson, N.; Cross, P.; Clarke, M. An examination of personal mobility patterns in space and time using Twitter. *Int. J. Agric. Environ. Inf. Syst.* **2014**, *5*, 55–72.
124. Zhao, W.X.; Jiang, J.; Weng, J.; He, J.; Lim, E.P.; Yan, H.; Li, X. Comparing Twitter and traditional media using topic models. In *Advances in Information Retrieval*; Clough, P., Foley, C., Gurrin, C., Jones, G.J.F., Kraaij, W., Lee, H., Mudoch, V., Eds.; Springer: Berlin, Germany, 2011; pp 338–349.
125. Simon, H.A. *The Sciences of the Artificial*, 3rd ed.; MIT Press: Cambridge, MA, USA, 1996.
126. Axelrod, R. Advancing the art of simulation in the social sciences. In *Simulating Social Phenomena*; Conte, R., Hegselmann, R., Terno, P., Eds.; Springer: Berlin, Germany, 1997; pp 21–40.
127. Gilbert, N.; Troitzsch, K.G. *Simulation for the Social Scientist*, 2nd ed.; Open University Press: Milton Keynes, UK, 2005.
128. Jumadi, J.; Heppenstall, A.; Malleson, N.; Carver, S.; Quincey, D.; Manville, V. Modelling individual evacuation decisions during natural disasters: A case study of volcanic crisis in Merapi, Indonesia. *Geosciences* **2018**, *8*, 196.
129. Crooks, A.T.; Wise, S. GIS and agent-based models for humanitarian assistance. *Comput. Environ. Urban Syst.* **2013**, *41*, 100–111.
130. Lewis, B.; Swarup, S.; Bisset, K.; Eubank, S.; Marathe, M.; Barrett, C. A simulation environment for the dynamic evaluation of disaster preparedness policies and interventions. *J. Public Health Manag. Pract.* **2013**, *19*, S42–S48.
131. Adam, C.; Gaudou, B. Modelling human behaviours in disasters from interviews: Application to Melbourne bushfires. *J. Artif. Soc. Soc. Simul.* **2017**, *20*, 12.
132. Macy, M.W.; Willer, R. From factors to factors: Computational sociology and agent-based modeling. *Ann. Rev. Sociol.* **2002**, *28*, 143–166.
133. Epstein, J.M. *Generative Social Science*; Princeton University Press: Princeton, NJ, USA, 2007.
134. Crooks, A.T.; Malleson, N.; Manley, E.; Heppenstall, A.J. *Agent-Based Modelling and Geographical Information Systems: A Practical Primer*; Sage: London, UK, 2019.
135. Crooks, A.T.; Castle, C.J.E.; Batty, M. Key challenges in agent-based modelling for geo-spatial simulation. *Comput. Environ. Urban Syst.* **2008**, *32*, 417–430.
136. Benenson, I.; Torrens, P.M. *Geosimulation: Automata-Based Modelling of Urban Phenomena*; John Wiley & Sons: London, UK, 2004.
137. Davis, P.K.; O'Mahony, A.; Gulden, T.R.; Osoba, O.A.; Sieck, K. *Priority Challenges for Social and Behavioral Research and Its Modeling*; RAND Corporation: Santa Monica, CA, USA, 2017.
138. Wilensky, U. *Netlogo. Http://ccl.Northwestern.Edu/netlogo*; Northwestern University, Evanston, IL, USA, 1999.
139. Luke, S.; Cioffi-Revilla, C.; Panait, L.; Sullivan, K.; Balan, G. Mason: A multi-agent simulation environment. *Simulation* **2005**, *81*, 517–527.
140. Taillandier, P.; Vo, D.; Amouroux, E.; Drogoul, A. Gama: A simulation platform that integrates geographical information data, agent-based modeling and multi-scale control. In *Principles and Practice of Multi-Agent Systems: Lecture Notes in Computer Science Volume 7057*; Desai, N., Liu, A., Winikoff, M., Eds.; Springer: New York, NY, USA, 2012; pp. 242–258.
141. Masad, D.; Kazil, J. Mesa: An Agent-Based Modeling Framework. In Proceedings of the 14th Python in Science Conference, Austin, TX, USA, 6–12 July 2015; Huff, K., Bergstra, J., Eds.; pp. 53–60.
142. Croitoru, A.; Crooks, A.T.; Radzikowski, J.; Stefanidis, A.; Vatsavai, R.R.; Wayant, N. Geoinformatics and social media: A new big data challenge. In *Big Data Techniques and Technologies in Geoinformatics*; Karimi, H.A., Ed.; CRC Press: Boca Raton, FL, USA, 2014; pp 207–232.
143. Croitoru, A.; Crooks, A.T.; Radzikowski, J.; Stefanidis, A. Geovisualization of social media. In *The International Encyclopedia of Geography: People, the Earth, Environment, and Technology*; Richardson, D., Castree, N., Goodchild, M.F., Kobayashi, A.L., Liu, W., Marston, R., Eds.; Wiley Blackwell: Hoboken, NJ, USA, 2017.
144. Wasserman, S.; Faust, K. *Social Network Analysis: Methods and Applications*; Cambridge University Press: New York, NY, USA, 1994.
145. Schweinberger, M.; Petrescu-Prahova, M.; Vu, D.Q. Disaster response on September 11, 2001 through the lens of statistical network analysis. *Soc. Netw.* **2014**, *37*, 42–55.
146. Comfort, L.K.; Ko, K.; Zagorecki, A. Coordination in rapidly evolving disaster response systems: The role of information. *Am. Behav. Sci.* **2004**, *48*, 295–313.

147. Abbasi, A.; Kapucu, N. Structural dynamics of organizations during the evolution of interorganizational networks in disaster response. *J. Homel. Secur. Emerg. Manag.* **2012**, *9*, 23.
148. Goodchild, M.F.; Janelle, D.G. Toward critical spatial thinking in the social sciences and humanities. *GeoJournal* **2010**, *75*, 3–13.
149. Goodchild, M.F. GIS and disasters: Planning for catastrophe. *Environ. Urban Syst.* **2006**, *30*, 227–229.
150. Longley, P.A.; Goodchild, M.F.; Maguire, D.J.; Rhind, D.W. *Geographical Information Systems and Science*; 3rd ed.; John Wiley & Sons: New York, NY, USA, 2010.
151. Hu, Y.; Janowicz, K.; Couclelis, H. Prioritizing disaster mapping tasks for online volunteers based on information value theory. *Geogr. Anal.* **2017**, *49*, 175–198.
152. Howard, R.A. Information value theory. *IEEE Trans. Syst. Sci. Cybern.* **1966**, *2*, 22–26.
153. Roche, S.; Propeck-Zimmermann, E.; Mericskay, B. Geoweb and crisis management: Issues and perspectives of volunteered geographic information. *GeoJournal* **2013**, *78*, 21–40.
154. Hudson-Smith, A.; Batty, M.; Crooks, A.T.; R., M. Mapping tools for the masses: Web 2.0 and crowdsourcing. *Soc. Sci. Comput. Rev.* **2009**, *27*, 524–538.
155. Haklay, M. Citizen science and volunteered geographic information: Overview and typology of participation. In *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (vgi) in Theory and Practice*; Sui, D., Elwood, S., F., G.M., Eds.; Springer: New York, NY, USA, 2013; pp. 105–124.
156. Kittur, A.; Chi, E.H.; Suh, B. Crowdsourcing User Studies with Mechanical Turk. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; pp. 453–456.
157. Chen, Y.; Konstan, J. Online field experiments: A selective survey of methods. *J. Econ. Sci. Assoc.* **2015**, *1*, 29–42.
158. Howe, J. *Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business*; Crown Business: New York, NY, USA, 2008.
159. National Audubon Society. Christmas Bird Count. Available online: https://www.audubon.org/conservation/science/christmas-bird-count (accessed on 24 December 2018).
160. Perger, C.; LeDrew, E.; See, L.; Fritz, S. Geography geo-wiki in the classroom: Using crowdsourcing to enhance geographical teaching. *Fut. Internet* **2014**, *6*, 597–611.
161. Ushahidi. Ushahidi: Open Source Software for Information Collection, Visualization and Interactive Mapping. Available online: http://www.ushahidi.com/ (accessed on 26 November 2018).
162. Peer, E.; Samat, S.; Brandimarte, L.; Acquisti, A. Beyond the turk: An empirical comparison of alternative platforms for crowdsourcing online behavioral research. *Soc. Sci. Res. Netw.* **2015**, doi:10.2139/ssrn.2594183.
163. Utz, S.; Schultz, F.; Glocka, S. Crisis communication online: How medium, crisis type and emotions affected public reactions in the fukushima daiichi nuclear disaster. *Public Relat. Rev.* **2013**, *39*, 40–46.
164. Liu, B.F.; Fraustino, J.D.; Jin, Y. How disaster information form, source, type, and prior disaster exposure affect public outcomes: Jumping on the social media bandwagon? *J. Appl. Commun. Res.* **2015**, *43*, 44–65.
165. Miyata, T.; Wakamatsu, H. Irrational reputational damage on wakame seaweed in Sanriku district after the Fukushima nuclear disaster: Revealed preference by auction experiment. *Fish. Sci.* **2015**, *81*, 995–1002.
166. Mao, Q. Experimental Studies of Human Behavior in Social Computing Systems; Ph.D. Thesis, Harvard University, Boston, MA, USA, 2015.
167. Berinsky, A.J.; Huber, G.A.; Lenz, G.S. Evaluating online labor markets for experimental research: Amazon.Com's mechanical turk. *Polit. Anal.* **2012**, *20*, 351–368.
168. Oktay, H.; Taylor, B.J.; Jensen, D.D. Causal Discovery in Social Media Using Quasi-Experimental Designs. In Proceedings of the First Workshop on Social Media Analytics, Washington DC, USA, 25–28 July 2010; Melville, P., Leskovec, J., Provost, F., Eds.; pp. 1–9.
169. Cosley, D.; Frankowski, D.; Terveen, L.; Riedl, J. In Suggestbot: Using Intelligent task Routing to help people find work in Wikipedia. In Proceedings of the 12th International Conference on Intelligent User Interfaces, Honolulu, HI, USA, 28–31 January 2007; pp. 32–41.
170. Bond, R.M.; Fariss, C.J.; Jones, J.J.; Kramer, A.D.; Marlow, C.; Settle, J.E.; Fowler, J.H. A 61-million-person experiment in social influence and political mobilization. *Nature* **2012**, *489*, 295–298.
171. Ayres, I.; Banaji, M.; Jolls, C. Race effects on ebay. *RAND J. Econ.* **2015**, *46*, 891–917.
172. Anderson, A.; Huttenlocher, D.; Kleinberg, J.; Leskovec, J. Engaging with massive online courses. In Proceedings of the 23rd International Conference on World Wide Web, Seoul, Korea, 7–11 April 2014; pp. 687–698.

173. Fradkin, A.; Grewal, E.; Holtz, D.; Pearson, M. Bias and reciprocity in online reviews: Evidence from field experiments on airbnb. In Proceedings of the Sixteenth ACM Conference on Economics and Computation, Portland, OR, USA, 15–19 June 2015; pp. 641–641.
174. Rudder, C. *Dataclysm: Love, Sex, Race, and Identity—What our Online Lives Tell us About our Offline Selves*; Crown Publishing: New York, NY, USA, 2014.
175. Richter, S.; Lechner, U. Transactive memory systems and shared situation awareness: A world of warcraft experiment. In Proceedings of the International Conference on Organizational Learning, Knowledge and Capabilities, Amsterdam, The Netherlands, 26–28 April 2009.
176. Salganik, M.J.; Dodds, P.S.; Watts, D.J. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* **2006**, *311*, 854–856.
177. Zhang, J.; Brackbill, D.; Yang, S.; Becker, J.; Herbert, N.; Centola, D. Support or competition? How online social networks increase physical activity: A randomized controlled trial. *Prev. Med. Rep.* **2016**, *4*, 453–458.
178. Bakshy, E.; Eckles, D.; Yan, R.; Rosenn, I. Social influence in social advertising: Evidence from field experiments. In Proceedings of the 13th ACM Conference on Electronic Commerce, Valencia, Spain, 4–8 June 2012; pp. 146–161.
179. Munson, S.A.; Lee, S.Y.; Resnick, P. Encouraging reading of diverse political viewpoints with a browser widget. In Proceedings of the International AAAI Conference on Weblogs and Social Media, Boston, MA, USA, 8–11 July 2013; pp. 419–428.
180. Soden, R.; Palen, L. Informating crisis: Expanding critical perspectives in crisis informatics. *Proc. ACM Hum. Comput. Interact.* **2018**, *2*, 162.
181. Hagar, C. Crisis informatics: Introduction. *Bull. Am. Soc. Inf. Sci. Technol.* **2010**, *36*, 10–12.
182. Palen, L.; Vieweg, S.; Sutton, J.; Liu, S.B.; Hughes, A.L. Crisis informatics: Studying crisis in a networked world. In Proceedings of the Third International Conference on E-Social Science, Ann Arbor, MI, USA, 7–9 October 2007.
183. Liu, S.B.; Palen, L.; Sutton, J.; Hughes, A.L.; Vieweg, S. In search of the bigger picture: The emergent role of on-line photo sharing in times of disaster. In Proceedings of the Information Systems for Crisis Response and Management Conference, Washington, DC, USA, 4–7 May 2008.
184. Vieweg, S.; Hughes, A.L.; Starbird, K.; Palen, L. Microblogging during two natural hazards events: What Twitter may contribute to situational awareness. In Proceedings of the 28th International Conference on Human Factors in Computing Systems, Atlanta, GE, USA, 10–15 April 2010; pp. 1079–1088.
185. Roy, J. Open data and open governance in canada: A critical examination of new opportunities and old tensions. *Fut. Internet* **2014**, *6*, 414–432.
186. Qadir, J.; Ali, A.; ur Rasool, R.; Zwitter, A.; Sathiaseelan, A.; Crowcroft, J. Crisis analytics: Big data-driven crisis response. *J. Int. Humanit. Action* **2016**, *1*, 12.
187. Imran, M.; Castillo, C.; Diaz, F.; Vieweg, S. Processing social media messages in mass emergency: A survey. *ACM Comput. Surv.* **2015**, *47*, 67.
188. Panteras, G.; Wise, S.; Lu, X.; Croitoru, A.; Crooks, A.T.; Stefanidis, A. Triangulating social multimedia content for event localization using Flickr and Twitter. *Trans. GIS* **2015**, *19*, 694–715.
189. Hagen, L.; Keller, T.; Neely, S.; DePaula, N.; Robert-Cooperman, C. Crisis communications in the age of social media: A network analysis of zika-related tweets. *Soc. Sci. Comput. Rev.* **2018**, *36*, 523–541.
190. Hourcade, J.P.; Nathan, L.P. Human computation and conflict. In *Handbook of Human Computation*; Michelucci, P., Ed.; Springer: New York, NY, USA, 2013; pp. 993–1009.
191. Summarization, P.S.U.f.D. Predicting salient updates for disaster summarization. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Beijing, China, 26–31 July 2015; Association for Computational Linguistics; pp. 1608–1617.
192. Hughes, A.L.; Tapia, A.H. Social media in crisis: When professional responders meet digital volunteers. *J. Homel. Secur. Emerg. Manag.* **2015**, *12*, 679–706.
193. Fichman, P.; Rosenbaum, H. *Social Informatics: Past, Present and Future*; Cambridge Scholars Publishing: Newcastle upon Tyne, UK, 2014.
194. Trist, E. The evolution of socio-technical systems: A conceptual framework and an action research program. In *Perspectives in Organization Design and Behavior*; Van de Ven, A.H., Joyce, W.F., Eds.; John Wiley: London, UK, 1981; pp. 19–75.

195. Reuter, C.; Hughes, A.L.; Kaufhold, M.A. Social media in crisis management: An evaluation and analysis of crisis informatics research. *Int. J. Hum. Comput. Inter.* **2018**, *34*, 280–294.
196. Imran, M.; Castillo, C.; Diaz, F.; Vieweg, S. Processing social media messages in mass emergency: Survey summary. In Proceedings of the Companion Proceedings of the The Web Conference, Lyon, France, 23–27 April 2018; ACM: New York, NY, USA, 2018; pp. 507–511.
197. Andreotta, M.; Nugroho, R.; Hurlstone, M.J.; Boschetti, F.; Farrell, S.; Walker, I.; Paris, C. Analyzing social media data: A mixed-methods framework combining computational and qualitative text analysis. *Behav. Res. Methods* **2019**, doi:10.3758/s13428–13019–01202–13428.
198. Aragón, P.; Volkovich, Y.; Laniado, D.; Kaltenbrunner, A. When a Movement Becomes a Party: Computational Assessment of New Forms of Political Organization in Social Media. In Proceedings of the Tenth International AAAI Conference on Web and Social Media, Cologne, Germany, 17–20 May 2016; AAAI Press: Palo Alto, CA, USA, 2015; pp. 12–21.
199. Spence, P.R.; Lachlan, K.A.; Rainear, A.M. Social media and crisis research: Data collection and directions. *Comput. Hum. Behav.* **2016**, *54*, 667–672.
200. Palen, L.; Anderson, K.M. Crisis informatics—New data for extraordinary times. *Science* **2016**, *353*, 224–225.
201. Arnott, D.; Pervan, G. A critical analysis of decision support systems research revisited: The rise of design science. In *Enacting Research Methods in Information Systems*; Willcocks, L.P., Sauer, C., Lacity, M.C., Eds.; Palgrave Macmillan: Cham, Germany, 2016; Volume 3, pp 43–103.
202. Brail, R.K. *Planning Support Systems for Cities and Regions*; Lincoln Institute of Land Policy: Cambridge, MA, USA, 2008.
203. Magiswary, D.; Raman, M.; Kaliannan, M. Knowledge management systems in support of disasters management: A two decade review. *Technol. Forecast. Soc. Change* **2013**, *80*, 1834–1853.
204. Zerger, A.; Smith, D.I. Impediments to using GIS for real-time disaster decision support. *Comput. Environ. Urban Syst.* **2003**, *27*, 123–141.
205. Shim, K.-C.; Fontane, D.G.; Labadie, J.W. Spatial decision support system for integrated river basin flood control. *J. Water Resour. Plan. Manag.* **2002**, *128*, 190–201.
206. Fiedrich, F.; Gehbauer, F.; Rickers, U. Optimized resource allocation for emergency response after earthquake disasters. *Saf. Sci.* **2000**, *35*, 41–57.
207. de Silva, F.N.; Eglese, R.W. Integrating simulation modelling and gis: Spatial decision support systems for evacuation planning. *J. Oper. Res. Soc.* **2000**, *51*, 423–430.
208. Zhang, D.; Zhou, L.; Nunamaker Jr, J.F. A knowledge management framework for the support of decision making in humanitarian assistance/disaster relief. *Knowl. Inf. Syst.* **2002**, *4*, 370–378.
209. Fikar, C.; Gronalt, M.; Hirsch, P. A decision support system for coordinated disaster relief distribution. *Expert Syst. Appl.* **2016**, *57*, 104–116.
210. Hadiguna, R.A.; Kamil, I.; Delati, A.; Reed, R. Implementing a web-based decision support system for disaster logistics: A case study of an evacuation location assessment for Indonesia. *Int. J. Disaster Risk Reduct.* **2014**, *9*, 38–47.
211. Palen, L.; Hughes, A.L. Social media in disaster communication. In *Handbook of Disaster Research*; 2nd ed.; Rodríguez, H., Donner, W., Trainor, J.E., Eds.; Springer: Cham, Germany, 2018; pp 497–518.
212. Dharma, D.; Soden, R.; LaLone, N. Crisis informatics for everyday analysts: A design fiction approach to social media best practices. In Proceedings of the 2018 ACM Conference on Supporting Groupwork, Sanibel Island, FL, USA, 7–10 January 2018; ACM: New York, NY, USA, 2018; pp. 230–243.
213. Quarantelli, E.L.; Lagadec, P.; Boin, A. A heuristic approach to future disasters and crises: New, old, and in-between types. In *Handbook of Disaster Research*; Tierney, K.J., Rodríguez, H., Quarantelli, E.L., Dynes, R., Eds.; Springer: New York, NY, USA, 2007; pp 16–41.
214. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221.
215. Olteanu, A.; Vieweg, S.; Castillo, C. What to expect when the unexpected happens: Social media communications across crises. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Vancouver, BC, Canada, 14–18 March 2015; ACM: New York, NY, USA, 2015; pp. 994–1009.
216. Parsons, T. *The Social System*; The Free Press/Macmillan: New York, NY, USA, 1964.
217. Lin, Y.R.; Margolin, D. The ripple of fear, sympathy and solidarity during the boston bombings. *EPJ Data Sci.* **2014**, *3*, 31.

218. Kasperson, R.E.; Renn, O.; Slovic, P.; Brown, H.S.; Emel, J.; Goble, R.; Kasperson, J.X.; Ratick, S. The social amplification of risk: A conceptual framework. *Risk An.* **1988**, *8*, 177–187.
219. Wen, X.; Lin, Y.R. Sensing distress following a terrorist event. In Proceedings of the 2016 International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction and Behavior Representation in Modeling and Simulation, Washington, DC, USA, 28 June–1 July 2016; Xu, K.S., Reitter, D., Lee, D., Osgood, N., Eds.; pp. 377–388.
220. Glasgow, K.; Vitak, J.; Tausczik, Y.; Fink, C. "With your help…We begin to heal": Social media expressions of gratitude in the aftermath of disaster. In Proceedings of the 2016 International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction and Behavior Representation in Modeling and Simulation, Washington, DC, USA, 28 June–1 July 2016; Xu, K.S., Reitter, D., Lee, D., Osgood, N., Eds.; pp. 226–236.
221. Oz, T.; Bisgin, H. Attribution of responsibility and blame regarding a man-made disaster: #flintwatercrisis. In Proceedings of the 4th International Workshop on Social Web for Disaster Management, Indianapolis, IN, USA, 24–28 October 2016.
222. Mao, A.; Mason, W.; Suri, S.; Watts, D.J. An experimental study of team size and performance on a complex task. *PLoS ONE* **2016**, *11*, e0153048.
223. Steiner, I.D. *Group Process and Productivity*; Academic Press: New York, NY, USA, 1972.
224. Pires, B.; Crooks, A.T. The geography of conflict diamonds: The case of Sierra Leone. In Proceedings of the 2016 International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction and Behavior Representation in Modeling and Simulation, Washington, DC, USA, 28 June–1 July 2016; Xu, K.S., Reitter, D., Lee, D., Osgood, N., Eds.; pp. 335–345.
225. Le Billon, P. The political ecology of war: Natural resources and armed conflicts. *Polit. Geogr.* **2001**, *20*, 561–584.
226. OpenStreetMap. Openstreetmap. Available online: https://www.openstreetmap.org/ (accessed on 19 March 2019).
227. Pires, B.; Crooks, A.T. Modeling the emergence of riots: A geosimulation approach. *Comput. Environ. Urban Syst.* **2017**, *61*, 66–80.
228. Stets, J.E.; Burke, P.J. Identity theory and social identity theory. *Soc. Psychol. Q.* **2000**, *63*, 224–237.
229. Ajzen, I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **1991**, *50*, 179–211.
230. Kniveton, D.; Smith, C.; Wood, S. Agent-based model simulations of future changes in migration flows for Burkina Faso. *Glob. Environ. Change* **2011**, *21*, S34–S40.
231. Schwarz, N.; Ernst, A. Agent-based modeling of the diffusion of environmental innovations—An empirical approach. *Technol. Forecast. Soc. Change* **2009**, *76*, 497–511.
232. Schlüter, M.; Baeza, A.; Dressler, G.; Frank, K.; Groeneveld, J.; Jager, W.; Janssen, M.A.; McAllister, R.R.; Müller, B.; Orach, K.; et al. A framework for mapping and comparing behavioural theories in models of social-ecological systems. *Ecol. Econ.* **2017**, *131*, 21–35.
233. Shi, J.; Ren, A.; Chen, C. Agent-based evacuation model of large public buildings under fire conditions. *Autom. Constr.* **2009**, *18*, 338–347.
234. Korhonen, T.; Hostikka, S.; Heliövaara, S.; Ehtamo, H. Fds+ Evac: An Agent Based Fire Evacuation Model. In *Pedestrian and Evacuation Dynamics 2008*; Klingsch, W.F.W., Rogsch, C., Schadschneider, A., Schreckenberg, M., Eds.; Springer: Wuppertal, Germany, 2010; pp. 109–120.
235. Mas, E.; Suppasri, A.; Imamura, F.; Koshimura, S. Agent-based simulation of the 2011 great east japan earthquake/tsunami evacuation: An integrated model of tsunami inundation and evacuation. *J. Nat. Disaster Sci.* **2012**, *34*, 41–57.
236. D'Orazio, M.; Spalazzi, L.; Quagliarini, E.; Bernardini, G. Agent-based model for earthquake pedestrians' evacuation in urban outdoor scenarios: Behavioural patterns definition and evacuation paths choice. *Saf. Sci.* **2014**, *62*, 450–465.
237. Torrens, P.M. High resolution space-time processes for agents at the built–human interface of urban earthquakes. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 964–986.
238. Wang, H.; Mostafizi, A.; Cramer, L.A.; Cox, D.; Park, H. An agent-based model of a multimodal near-field tsunami evacuation: Decision-making and life safety. *Transp. Res. Part C Emerg. Technol.* **2016**, *64*, 86–100.
239. Chen, X.; Meaker, J.W.; Zhan, F.B. Agent-based modeling and analysis of hurricane evacuation procedures for the florida keys. *Nat. Hazards* **2006**, *38*, 321–338.

240. Dawson, R.J.; Peppe, R.; Wang, M. An agent-based model for risk-based flood incident management. *Nat. Hazards* **2011**, *59*, 167–189.
241. Helbing, D.; Farkas, I.; Vicsek, T. Simulating dynamical features of escape panic. *Nature* **2000**, *407*, 487–490.
242. Torrens, P. A computational sandbox with human automata for exploring perceived egress safety in urban damage scenarios. *Int. J. Digit. Earth* **2017**, *11*, 369–396.
243. Ren, C.; Yang, C.; Jin, S. Agent-Based Modeling and Simulation on Emergency Evacuation. In Proceedings of the First International Conference on Complex Sciences, Shanghai, China, 23–25 February 2009; Zhou, J., Ed.; pp. 1451–1461.
244. Grinberger, A.Y.; Lichter, M.; Felsenstein, D. Simulating urban resilience: Disasters, dynamics and (synthetic) data. In *Planning Support Systems and Smart Cities*; Geertman, S., Ferreira, J., Goodspeed, R., Stillwell, J., Eds.; Springer: Cham, Germany, 2015; pp. 99–119.
245. Grinberger, A.Y.; Felsenstein, D. Dynamic agent based simulation of welfare effects of urban disasters. *Comput. Environ. Urban Syst.* **2016**, *59*, 129–141.
246. Burger, A.; Oz, T.; Crooks, A.T.; Kennedy, W.G. Generation of Realistic Mega-City Populations and Social Networks for Agent-Based Modeling. In Proceedings of the Computational Social Science Society of Americas Conference, Santa Fe, NM, 19–22 October 2017.
247. Dataverse Project. Open Source Research Data Repository Software. Available online: https://dataverse.org/ (accessed on 11 March 2019).
248. COMSES. Network for Computational Modeling in Social and Ecological Sciences. Available online: https://www.comses.net/ (accessed on 11 March 2019).
249. ABMGIS. Agent-Based Modelling and Geographical Information Systems. Available online: https://github.com/abmgis/abmgis (accessed on 11 March 2019).
250. Hargittai, E. Is bigger always better? Potential biases of big data derived from social network sites. *Ann. Am. Acad. Polit. Soc. Sci.* **2015**, *659*, 63–76.
251. Missing Maps. Putting the World's Vulnerable People on the Map. Available online: https://www.missingmaps.org/ (accessed on 11 March 2019).
252. Humanitarian OpenStreetMap Team. Available online: https://www.hotosm.org/ (accessed on 11 March 2019).
253. Guerin, S.; Carrera, F. Sand on fire: An interactive tangible 3d platform for the modeling and management of wildfires. *WIT Trans. Ecol. Environ.* **2010**, *137*, 57–68.
254. SimTable. Available online: https://www.simtable.com/ (accessed on 12 March 2018).
255. Survey Monkey. Available online: https://www.surveymonkey.com/ (accessed on 11 March 2019).
256. Amazon Mechanical Turk. Available online: https://www.mturk.com/ (accessed on 11 March 2019).
257. Ali, S.; Islam, N.; Rauf, A.; Din, I.; Guizani, M.; Rodrigues, J. Privacy and security issues in online social networks. *Fut. Internet* **2018**, *10*, 114.
258. Golbeck, J.; Mauriello, M.L. User perception of facebook app data access: A comparison of methods and privacy concerns. *Fut. Internet* **2016**, *8*, 9.
259. Friedland, G.; Sommer, R. Cybercasing the Joint: On the Privacy Implications of Geotagging. In Proceedings of the Fifth USENIX Workshop on Hot Topics in Security (HotSec 10), Washington, DC, USA, 10 August 2010.
260. Elwood, S.; Leszczynski, A. Privacy, reconsidered: New representations, data practices, and the geoweb. *Geoforum* **2011**, *42*, 6–15.

ABSTRACT

DISASTER THROUGH THE LENS OF COMPLEX ADAPTIVE SYSTEMS: EXPLORING EMERGENT GROUPS UTILIZING AGENT BASED MODELING AND SOCIAL NETWORKS

Annetta G. Burger

George Mason University, 2020

Dissertation Director: Dr. Andrew Crooks

Disasters have become more frequent and intense in the last decades and are a significant challenge to the health and well-being of local communities and regions. As a potential solution to this problem attention has been drawn to community resilience and the building of social networks that support or hinder local response and recovery. Research on disasters and community resilience has shown how the ability to leverage social capital through a community's social networks is fundamental to the ability of individuals and communities to respond to disaster events, but there is little understanding of how the evolution of social networks can impact disaster response and recovery. A computational framework and agent-based model of disasters can provide a virtual laboratory for testing social network effects and uncover their role, function and underlying mechanisms in community resilience. Agent-based models are suited to test bottom-up dynamics and the interactions of variables that lead to the nonlinear relationships in disasters. To what extent can an agent-based model characterize the social networks that emerge in response to a no-warning disaster event such as a Nuclear Weapon of Mass Destruction impacting Manhattan Island? To explore this question this research reviews theories of disaster, primarily from sociological and anthropological research, and builds a conceptual model of disasters from which to develop an agent-based model. The agent-based model represents social networks relevant in both the normal commuting patterns of New York City and the emergent social networks responding to a Nuclear Weapon of Mass Destruction impacting Manhattan Island. Network representations of social groups along with physical representations of the community shows how individuals adapt and respond to the disaster in the initial response. Integrating agent-based models with social network analysis provides new spaces for scientific inquiry into disasters, the dynamics of social networks in resilient communities, and those areas of complexity most often explored today with qualitative methodologies.

ABSTRACT

COLLECTIVE STRESS IN THE DIGITAL AGE

Talha Oz, Ph.D.

George Mason University, 2020

Dissertation Director: Dr. Andrew Crooks

Collective stress occurs when communities are faced with unfavorable circumstances in which they fear losing (or lose) conditions of life they are accustomed to. Such stressors are plenty in current times, ranging from complications of pandemics, technological and man-made disasters, toxic job environments, economic crises, government oppressions, terror attacks, and political breakdowns are all acting as catalysts for collective stress. Since their impacts can be devastating and multi-faceted, a better understanding of social behaviors before (e.g. in preparation), during (e.g. in first responses), and after (e.g. in recovery) them is needed to prevent and alleviate their effects. While going through all these phases, in our current digital age we actively use information and communication technologies (ICT) both at work (e.g. calendar and e-mail) and outside work (e.g. social media). I argue and demonstrate in this dissertation that by examining ICT data with computational techniques, we can understand, model, and theorize about collective stress related social behaviors in ways that were not possible before, and thus handle their

complications more effectively. In particular, I show how computational social science (CSS) can solve the common scientific problems of behavioral unobservability (temporal and spatial), limited study extent (sample size and longevity), and informant subjectivity (biases in self-report based measurements). Thus, this dissertation makes a paradigmatic contribution to the field of collective stress research. Regarding more specific theoretical, methodological, and empirical contributions, this dissertation contains three studies each of which is the first computational social scientific study in its own domain. Collective stress researchers have made calls particularly for a need of an objective work stressor measurement strategy, for more empirical studies on blame attribution (in collective stress situations) and also on adoption of teleworking out of necessity; and this dissertation responds to each of them with a separate essay. The study on work stressors is a solution-oriented research that guides People Analytics practitioners in achieving better employee experience by showing how to measure stressors in organizations using commonplace workplace ICT. The other two studies serve to the advancement of theories by forming and testing hypotheses using the data collected and analyzed during the 2016 Flint Water Crisis and COVID-19 pandemic, from social media and workplace ICT (calendar and workplace messaging apps), respectively. Thus, by building novel research designs (such as retrospective cohort analysis), implementing new computational and quantitative methods (such as combining data from multiple sources, conducting large scale social network analysis, and sentiment analysis), and exploiting newly available data sources (social media and work ICT), this dissertation shows how computational social science can increase our understanding of collective stress in the digital age.

Abstract

GEO-TEXTUAL DATA ANALYTICS: EXPLORING PLACES AND THEIR CONNECTIONS

Xiaoyi Yuan, PhD

George Mason University, 2020

Dissertation Director: Andrew Crooks

Place is defined by physical, social, and economic activities and processes. Understanding the complexity of socially constructed places is a fundamental question in geography, sociology, and many other social sciences. Meanwhile, the growing amount of user volunteered geographic information (VGI) leads us to study place through a new perspective. For instance, Flickr users report local activities in various geographic locations that capture individualistic experiences and impressions of the locations. Many previous studies utilizing non-textual VGI have focused primarily on analyzing geographical footprints of places, which separated place from its meaning. This dissertation argues that the textual part of VGI provides us with unprecedented opportunities for deriving patterns of place meanings on an individual level. More specifically, three research questions are pursued in this dissertation. First, how to quantify placeness (i.e., place identities) that has been traditionally studied via theoretical and qualitative methods? Second, as place being innately interconnected, how can we assess connections between places in networks so that we can apply network science to analyze complex connections between places? Third, as geo-textual data can also reveal social events, how to trace critical events across places using geo-textual data? In order to answer these research questions, this dissertation leverages advances in

machine learning, natural language processing and network analysis techniques on geo-textual data. By doing so this dissertation is able to build foundations for geo-textual data analytics and thus providing a new lens to study places and the connections between them from the bottom up. Overall, this dissertation showcases an interdisciplinary effort in computational social science research that combines computational textual data analytics and social scientific theories including human geography and sociology.