BOSTON COLLEGE
School of Social Work

EXAMINING THE CONFLICT-RELATED PREDICTORS AND SPATIOTEMPORAL DYNAMICS OF ATTACKS AGAINST HUMANITARIAN ASSISTANCE: AN ALGORITHMIC MODELING APPROACH

A dissertation
by

CARA M. ANTONACCIO

Submitted in partial fulfillment
of the requirements for a degree of
Doctor of Philosophy

SEPTEMBER 2022

Examining the Conflict-Related Predictors and Spatiotemporal Dynamics of Attacks Against Humanitarian Assistance: An Algorithmic Modeling Approach

A dissertation

by

CARA M. ANTONACCIO

Dissertation Chair: Dr. Christopher P. Salas-Wright

Abstract

Around the world, attacks against humanitarian aid workers pose a pervasive and intransigent threat to health and human rights, but evidence about the complex factors that predict perpetrators' behavior as well as attack outcomes remains quite limited. While previous studies have addressed several aspects of local and global trends of attacks against humanitarian assistance, more evidence is also needed to understand the dynamics of recurrent incidents, small-scale attacks, as well as patterns of events across time and space; and how the observed trends are driven by conflict-related and contextual factors. In this dissertation, we investigate the predictors and spatiotemporal of attacks against humanitarian assistance from 1997 to 2022 using publicly-available datasets from the Aid Worker Security Database (AWSD, Humanitarian Outcomes, 2022) and the Armed Conflict Dataset version 22.0 (ACD; Glaeditsch et al., 2002; Davies et al., 2022).

Keywords: prevention, prediction, exploratory spatial analysis, armed conflict, health and human rights, humanitarian assistance

DEDICATION

For my mom, Katie

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I. Introduction

In December 2015, a gunship operated by the United States Air Force conducted a series of five aerial bombing runs on a hospital in Kunduz, Afghanistan which destroyed the emergency ward and killed fourteen *Médecins Sans Frontières* (MSF) staff members, 24 patients, and four caregivers (Benton and Atshan, 2016). This attack is but one example among many—around the world, attacks against aid workers and denials of humanitarian access pose growing threats to health and human rights. Since monitoring began in 1997, more than 2,186 aid workers have been killed, 1,665 kidnapped, and 2,308 wounded in attacks against humanitarian assistance (Humanitarian Outcomes, 2022). The impact of attacks on affected communities is not limited to acute casualties—they also compromise population access to critical health and social protection services. In affected communities, the loss of humanitarian personnel and destruction of critical infrastructure increases population risks of death and disability due to elevated incidence rates of preventable conditions like malnutrition, malaria, and diarrheal disease (Briody et al., 2018).

In addition to acute morbidity and mortality, attacks against humanitarian assistance also damage public trust in health and protection services. In some humanitarian emergencies, distrust of aid workers due to attacks and denials of access increase the risk of subsequent attacks in the short-term, and undermine health and social protection systems in the long-term (Haar et al., 2021). Although most conflict parties acknowledge the protected status of civilians and humanitarian actors as a defining principle of International Humanitarian Law (IHL), the extent to which recognition of these principles translates into respect for protected populations varies substantially. In some conflict settings, qualitative evidence points to armed forces and groups specifically targeting humanitarian personnel and infrastructure to achieve tactical and strategic objectives, in defiance of international legal protections. In other contexts, humanitarian personnel and infrastructure are impacted not via premeditated targeting but rather as a “collateral” effect of armed conflict. Despite evidence from across contexts which underscores the devastating impact of attacks against humanitarian assistance, more research is needed to examine the role, nature, and motivations of the armed forces and groups working in proximity to aid organizations, dynamics of conflict and how, if at all, these characteristics predict attacks against humanitarian assistance. More evidence is also needed to understand who is affected by humanitarian violence, who perpetrates attacks and how, where humanitarian workers are, and what they are doing when attacks occur (Hoelscher et al., 2017).

Purpose, Aims, and Research Questions

This dissertation contributes to previous research by examining how conflict-related and contextual factors predict attacks against humanitarian assistance, using a sample of 170 countries from 1997 to 2021; exploring the distribution of attacks against humanitarian assistance from 1997 to 2022 across time and space, using a density-based spatiotemporal clustering algorithm (ST-DBSCAN); and investigating the local dynamics of attacks against aid workers in emerging clusters, which represent areas where the risks facing humanitarian operations may be growing. Three studies investigate the following primary research questions: 1) how, if at all, do armed conflict- and context-related factors predict attacks against humanitarian assistance; 2) how, if at all, do attacks against humanitarian assistance form patterns over time and geographical space; and 3) where, if anywhere, are new clusters of risk for attacks against humanitarian assistance emerging, globally?

Significance

Pursuant to effective, evidence-based protection and prevention strategies, it is important to understand how conflict and context predict the risk of attacks against humanitarian assistance, and to detect and demonstrate how attacks form patterns across time and space. Greater knowledge in these areas can be used strategically—to plan, allocate scarce resources, and establish evidence-based policies and protocols to keep humanitarian aid workers safe in high-risk areas. With better knowledge of why attacks occur in specific conflicts and humanitarian contexts, it will be easier to build evidence-based interventions to protect aid workers and modify attackers' behavior. In addition, results from the three studies may be continuously updated and made publicly available via web-based mapping and decision support tools in order to better protect aid workers and prevent unnecessary casualties due to attacks against humanitarian assistance.

Terminology and Conceptual Issues

The importance of conceptual clarity

Across the literature on attacks against humanitarian assistance, many key terms like ‘attack’ and ‘humanitarian’ are defined and operationalized inconsistently, which affects the clarity, interpretation, and comparability of findings across studies (Mahieu et al., 2020). Pursuant to clarity and consistency with the extant literature on attacks against humanitarian assistance, this section will review the key terms and conceptual issues related to the humanitarian space, attacks, and armed conflict.

Key terms and conceptual issues

The humanitarian space. The humanitarian space refers to the people (i.e., personnel, aid recipients, and patients) and things (i.e., facilities, transport and equipment) which are involved in providing assistance to populations in the wake of conflict, natural disasters, and other humanitarian emergencies. Outcomes of attacks affecting people include: wounding, kidnapping, killing, harassing, threatening, robbing, and intimidating personnel, patients, or those trying to access humanitarian assistance (SHCC, 2020). Examples of attacks against *things* include forced entry, encircling, looting, shooting into, bombing, shelling, or other forceful interference with facilities, equipment, and transports involved in humanitarian operations (including assaults on, theft of, and interference with ambulances and other medical vehicles) (Terry, 2013). In addition, improper use of the Red Cross and Red Crescent emblems and the misuse of facilities and signs intended to designate protected humanitarian entities are additional forms of violence recognized by the International Committee of the Red Cross and Red Crescent Societies (ICRC) and Customary International Humanitarian Law (IHL) as critical violations of humanitarian protections (Hague Regulations, 1907).

What constitutes an "attack"? Legal protections for humanitarian assistance were first enshrined in the Geneva Conventions of 1864 and 1949 (ICRC, 1949). The Fourth Geneva Convention explicitly condemns attacks against humanitarian assistance, stating that "civilian hospitals organized to give care to the wounded and sick, the infirm and maternity cases, may in no circumstances be the object of attack, but shall at all times be respected and protected by the Parties to the conflict" (Article 18). However, while the Geneva Conventions condemn attacks and explicitly describe the parties to be protected, they do not explicitly define what constitutes an attack. Moreover, the terms "attack" and "violence", are often used interchangeably in the literature and by researchers, advocates, legal analysts, and policymakers. Some argue that the word "attack" is more powerful, as it evokes the legal protections of the Geneva Conventions. Interestingly, though, the WHO and Safeguarding Health Care Coalition (SHCC) actually avoid using the word “attack”, suggesting a more appropriate term is “incident” (Mahieu et al., 2020). The latter, they argue, includes interferences like administrative access barriers, denials, protests, or demonstrations that impede healthcare provision and recognizes, at least tacitly, that indiscriminate violence and crossfire are often the source of the threat.

These gray areas are not empirically unimportant; for example, although attacks against humanitarian assistance include both violent and non-violent incidents, it is common in the literature to refer to attacks as ‘violence’, which can limit the scope of inquiry to physical assault and may neglect to capture threats, interferences, and other non-violent forms of attack. Examples of violent attacks include bombing, robbery, hijacking, shooting/gunfire, searching of facilities, fire and arson, abducting humanitarian workers, assault, torture, execution, violent demonstrations, and sexual violence. Examples of non-violent attacks include forced closure of facilities, military use and takeover, cyberattacks, denial or delay of access to populations in need of assistance, forcing aid workers to act against their ethical code, administrative harassment, obstruction, psychological violence or threat of violence (SHCC, 2020, p. 22).

Criteria for documenting incidents. While there is generally agreement about what is and is not an attack, a consensus has not been reached about precisely how to document incidents. Two sets of inclusion criteria, broadly, are used to define which incidents are measured as attacks. The first set of inclusion criteria documents only deliberate and intentional attacks against clearly-identifiable humanitarian operations. This approach is useful because by limiting the scope of inquiry to identifiable entities, there is greater confidence that the perpetrator was cognizant of the humanitarian status of the target (Mahieu et al., 2020). The SHCC for example, uses a discrete monitoring system that only reports direct attacks against the health care sector. However, this approach may neglect to capture some key forms of attack (e.g., psychological violence; blockages; harassment, intimidation, and threats of violence; one-sided violence; attacks on other sectors which affect health). Similarly, restrictive criteria for documenting attacks are also limited in conflict settings like Syria—where attacks are extremely common despite the fact that ambulances are camouflaged, humanitarian aid workers maintain a low profile, and facilities are built underground to discourage deliberate targeting (Elamein et al., 2017).

The second set of criteria casts a broader net to include all attacks that affect civilians, without considering intentionality or humanitarian identifiability. For example, the World Health Organization (WHO) defines an attack as any act of obstruction or threat of violence (verbal or physical) that interferes with the availability, access, or delivery of humanitarian assistance during an emergency (Yousuf et al., 2021; Mahieu et al., 2020). The WHO definition captures incidents based on impact, with criteria that do not distinguish between violence against the health sector and an attack perpetrated against another aid sector that affects health. Broader criteria for documenting attacks such as the definition used by the WHO offers a "health in all policy" lens for documenting attacks which captures attacks on roads, markets, education, and related non-medical arms of the humanitarian system. In the three dissertation studies, we use the stronger term "attack" versus incident or events. Attacks are sampled under the Aid Worker Security Database using the broad criteria set forth by the WHO (i.e., humanitarian aid not only includes health centers and clinical services but also other forms of assistance, infrastructure, and emergency services).

Conflict-related and contextual risk factors for attacks on aid workers

Conflict presence, intensity, type, and actors. In this section, we will describe the key theoretical risk factors that we will explore in chapter two and how they are hypothesized to predict attacks against humanitarian assistance. We assess the importance of armed conflict as a predictor of attacks in a given country-year, based on existing theory suggesting that the strategic use of violence may influence where and how aid workers are targeted (Hoelscher et al., 2017). Similarly, we examine the importance of the type of dispute, which is also thought to influence how armed groups target humanitarian assistance, as parties to conflict may be more suspicious of humanitarians in conflicts where insurgents are seeking control over the government and in other cases territorial control (Hoelscher et al., 2017). The structure, organization, and leadership of armed forces and groups in some settings is likewise thought to incentivize attacks on civilians or aid workers. Evidence from some conflict settings finds that armed forces and groups intentionally target humanitarian personnel and infrastructure to achieve tactical and strategic objectives, in defiance of international laws and humanitarian principles (Boutton et al., 2018).

Evidence also suggests that violence against civilians and humanitarian aid operations happens together, although this relationship is disputed by some (Hoelscher et al., 2017; Narang et al., 2017). The logic behind concurrent aid worker attacks and one-sided violence is that when civilians account for a large proportion of conflict-related casualties and combatants use violence indiscriminately against non-combatants, combatants may be more likely to attack neutral parties involved in humanitarian assistance—including hospitals and clinics, which are frequently subjected to damage during the implementation of critical services to alleviate direct humanitarian needs (Madhiwalla & Roy, 2009; Hoelscher et al., 2017).

Contextual risk factors. Along with dispute-level factors, existing evidence suggests that the political, social, and economic contexts may also play a role in attacks against humanitarian assistance. This includes macroeconomic outcomes like per-capita gross national income (PCGNI) and national inequality. In terms of the peacekeeping environment, operational coordination between humanitarian organizations and UN peacekeeping organizations (PKOs) has also been proposed as potentially important for humanitarian security, as certain conflict actors are opposed to aid operations' real and perceived political, development, and humanitarian agendas (Boutton, 2018). Although PKOs are designed to create space for secure humanitarian operations in conflict-affected areas, it is thought that because of their international and/or 'Western' associations, humanitarian workers are vulnerable to violence despite the presence of PKOs (Hoelscher et al., 2017). This is despite the fact that the majority of humanitarian aid workers are nationals in the country they work in. Taken together, this range of conflict-related and contextual factors has been proposed as theoretical predictors of attacks against humanitarian assistance, and will be included as hypothetical predictors of attacks against humanitarian assistance in our cross-national time series analysis.

Theoretical Framework

Parametric and non-parametric predictive modeling and theory testing

Chapter 2 of the dissertation compares parametric and nonparametric statistical learning approaches in terms of their accuracy in modeling attacks against humanitarian assistance using on conflict-related and contextual predictors. Contemporary applications of statistical learning techniques are found in the literature on health and medicine, business, infrastructure and engineering systems, cyber-security, disaster planning, and counterterrorism (Zhao, 2021). This theoretical approach is used to light the complex nature (i.e., non-linear, interactive) of relationships among phenomena and offers a methodological framework for researchers to build models with better face validity (see Figure 1). Accurate predictive models also facilitate a straightforward evaluation of competing theories, making this analytical approach the gold standard of theory testing and establishing baseline measures (Boateng et al., 2018). In the following sections, we review the two statistical learning paradigms that guide our predictive analysis along with the respective strengths and limitations of each.

Figure 1. Predictive relationship of interest in statistical learning

Parametric predictive models. Common parametric modeling techniques include linear regression (e.g., ordinary least squares [OLS], logit, probit); lasso, elastic net, and penalized regression models; and principal components analysis. Parametric models assume that a relationship exists between a predictor, X, and an outcome, Y, and model the relationship between variables according to the function $(X'X) - 1(X'Y)$ where $Y = \beta_0 + \beta_1 X_1 + \varepsilon$. Parametric modeling starts by assuming $N \sim$ i.i.d. Normal, Poisson, Weibull; and that inputs, X_i, influence the outcome, Y, according to a known data generation process that can be evaluated using in-sample goodness-of-fit tests and examination of model residuals (see Figure 2).

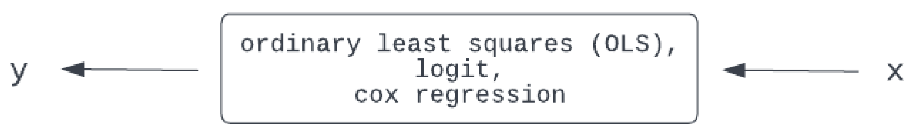

Figure 2. Theoretical predictive relationship in parametric modeling

Linear (logistic) regression. Logistic regression is the most common form of parametric predictive analysis and is the form we compare to nonparametric predictive techniques in the study. Logistic regression, models the function $f(N)$ as a linear combination of independent variables to separate zeros from ones in the data. We model the data generation process in logistic regression as a function of the proportion of positive cases (i.e., an attack was observed) in the data (i.e., the mean of the dependent variable). Logistic regression attempts to model the predictive relationship $f(N)$ by creating a hyperplane through the data that does the best job of effectively placing 0s on one side of the line and 1s on the other. Statistical significance in parametric predictive models is interpreted as the probability of observing a relationship between X and Y at least as extreme as that observed, assuming that the null hypothesis of no association represents the true state of reality. Parametric models fit only a fraction of the points in the data, and consider just one possible relationship – which reduces the amount of data needed to estimate a mode. In logistic regression, $\rho = \frac{1}{1+e^{-x_i \beta}}$, where $-x_i\beta = \mu_i$. When interpreting model results, X affects $Y \Rightarrow \rho(x) < 0.05$.

Parametric modeling approaches like logistic regression set restrictions on the shape of the relationship $f(N)$, and the resulting statistical models do a poor job of accurately predicting Y if the true relationship between X and Y is nonlinear (Breiman, 2001a). If a simple binary classification is desired, a linear regression model may be suitable for a researcher's needs—however, it is not often that data fit cleanly into such simple orderly spaces. Moreover, statistical significance is a poor measure of accuracy in many prediction cases (Poldrack et al., 2020). The frequent incorrect use of null hypothesis significance testing (i.e., formulating hypotheses as Bayesians, but testing them as frequentists), "garbage can" models, and multicollinearity are also major concerns in parametric predictive modeling approaches (Ward et al., 2010). The limitations of linear regression models result from the parametric assumptions used to fit the data. The basic assumptions of predictive regression models (i.e., perfectly measured predictors, error term correlated with response, linear relationship between X and Y) are often misunderstood and misused, and in the above cases, alternative methods can produce more accurate predictions for outcomes of interest (Shmeuli, 2010).

Nonparametric predictive models. Compared to parametric approaches like linear (logistic) regression, the nonparametric algorithmic models used in this study assume that nature, N, is complex and unknown—perhaps even unknowable (see Figure 3). In nonparametric modeling approaches, ρ is no longer important.

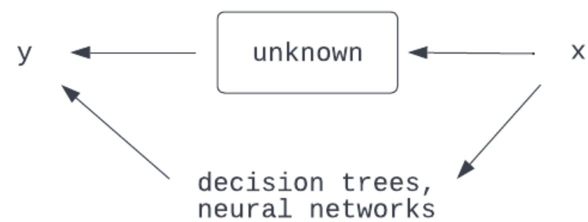

Figure 3. Theoretical prediction relationship in nonparametric modeling

Flexibility estimating $f(N)$. Nonparametric modeling techniques that allow flexibility estimating $f(N)$ can strengthen the accuracy and reliability of model predictions, which has implications for producing more accurate and reliable evidence on which to base policy advice (Shmueli, 2010). Whereas logistic regression models the data space according to a joint distribution between the outcome, Y, and predictor variables, X_i, nonparametric models seek to simply to find a function, f, that does a good job of using X to predict Y without assuming that the true relationship follows a distribution that is normal, negative binomial, or beta (Breiman, 2001a). Instead of assuming *a priori* a relationship exists in the data and testing whether or not it is true (in contrast to parametric models like linear regression), nonparametric modeling uses multiple methods to estimate the relationship $f(N)$ directly from the data, with model validation resting on out-of-sample predictive accuracy. Some nonparametric statistical learning methods, like ridge regression and the lasso are based on the linear model but use techniques like regularization to enhance the accuracy of model predictions. Others methods, like deep neural networks are total black boxes where it is almost impossible to determine the relationship between the predictors and outcome measure.

Tree-based algorithmic models. Decision trees are a nonparametric modeling technique that use recursive partitioning of the data space by some value (i.e., a constant), to make the data space more homogenous on Y with each additional partition (Kamiński et al., 2018). Decision tree models stratify or segment a predictor space into a number of simple regions to achieve to build a predictive model (see Figure 4). Tree-based models are nonparametric in that decisions regarding cut-points in the data do not rely on the distribution of the variables in the model. The algorithm starts by inducing a single split on the data space, creating two daughter nodes at the end of each partition branch. The set of splitting rules used to segment the data can be summarized in a decision tree.

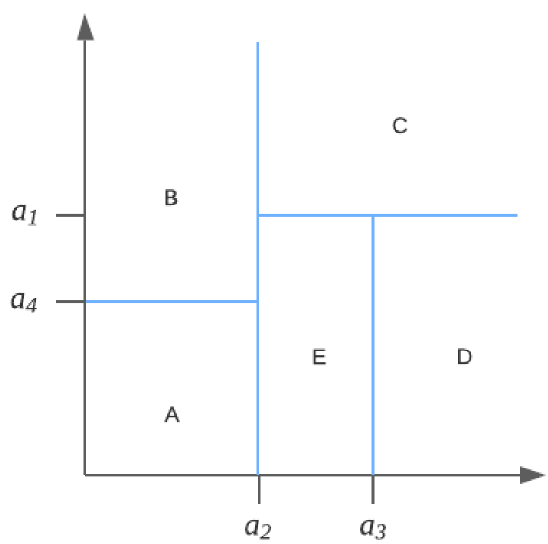

Figure 4. Segmentation of a data space in tree-based models

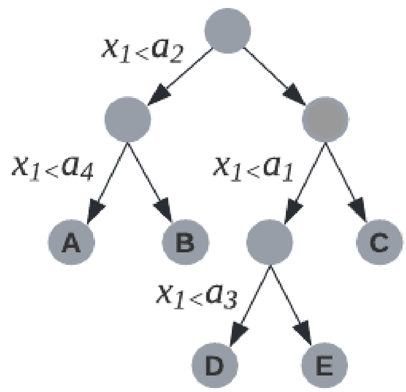

Figure 5. Decision tree data splitting procedure

Decision tree-based models are both simple and easily interpretable. Imagine the whole data set exists at the top of a tree, and the algorithm selects the variable to partition the data such that the X variable cut-point optimizes different values of Y in the newly generated regions, R. Each split is like a yes/no question – for example, is the value of the predictor variable less than a certain threshold? If yes, then right, if no, then left (see Figure 4 and Figure 5). Additional partitions are made in the same way, producing more branches and nodes, continuing until a stopping criterion (set by researcher) is met. This process is repeated iteratively until some stopping criterion (i.e., depth or error rate) is achieved. The resulting structure shows a nonlinear and interactive tree structure, which is interpreted as X_1 & X_2 & X_3 ... leads to $Y = 0$ or $Y = 1$. As depicted in Figure 6, random forests are an ensemble learning method that produces multiple decision trees, which are then combined to arrive at a consensus, or a single prediction (Breiman, 2001b).

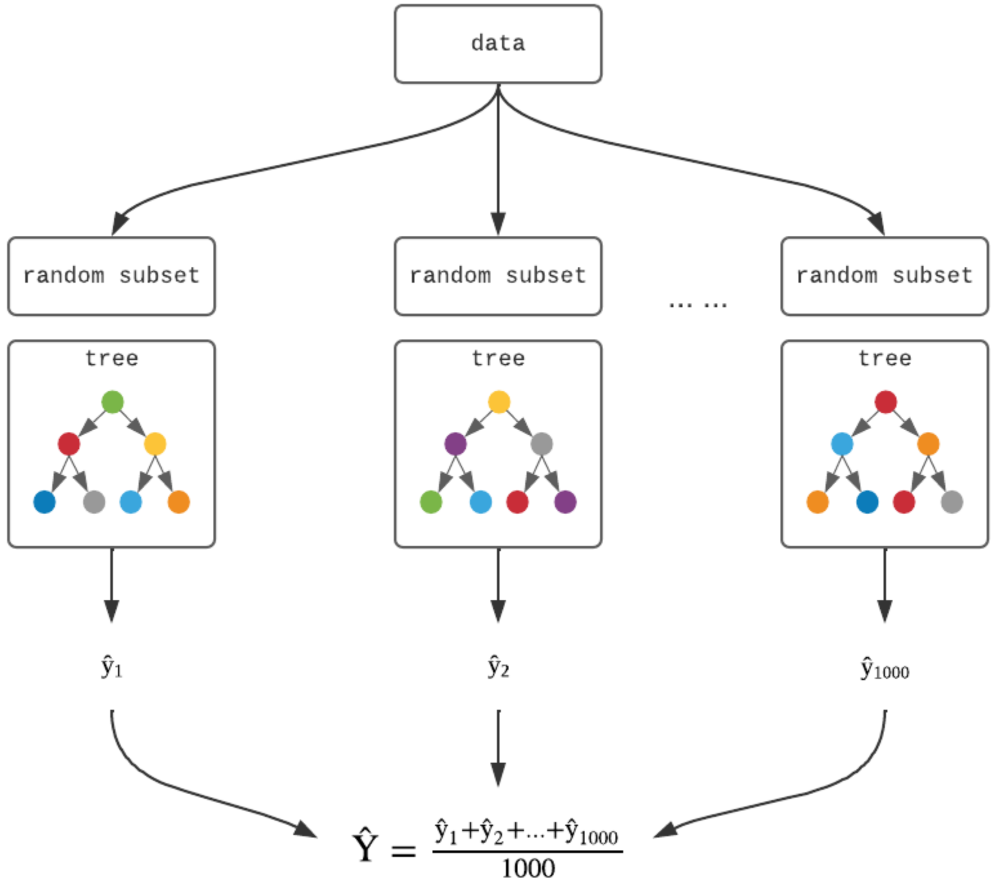

Figure 6. Random forest algorithmic modeling procedure

Tobler's First Law of Geography

A major concern in regression-based statistical learning models is how to address spatial autocorrelation, or the tendency for outcomes to "cluster" in particular areas. The analyses in chapters 3 and 4 explore spatial autocorrelation of attacks directly, through the theoretical lens of Tobler's First Law of Geography, network theory, and diffusion theory. According to Tobler's Law, "everything is connected to everything else, but near things are more related than distant things" (p. 234). Applied in the context of the spatiotemporal point process of attacks against humanitarian assistance, Tobler's Law would suggest that two attacks that occur closer together are more likely to be similar that those occurring in different clusters. The spatiotemporal analyses in chapters 3 and 4 examine how, if at all, Tobler's Law is validated in the case of attacks against humanitarian assistance, through the applied lens of network theory and diffusion theory.

Network theory. Network theory examines how objects or, in the case of this dissertation, attacks create a network or pattern across dimensions (i.e., space and time). Diffusion theory considers how, over time, phenomena spread from their origin (core) point. Meade and Emch (2010) classify spatiotemporal diffusion patterns into three types, including hierarchy, contagion, and relocation, each with its own epidemiological meaning and distinct evolution mechanism. Point networks, their diffusion dynamics and relationships are depictured in figures 7 and 8, where A:1 represents the first attack at a certain point on the Earth's surface, and subsequent attacks over time occur around it ("neighboring"). Time is a key parameter we add emphasize through this theory, where attacks that occur closer together in time and space are more similar to or even related to one another. In this dissertation, this theory is operationalized where two or more attacks are considered a spatiotemporal cluster if they occur within adjacent 365-day periods and 1-degree longitude/latitude (approximately 100 kilometers; see Figure 7).

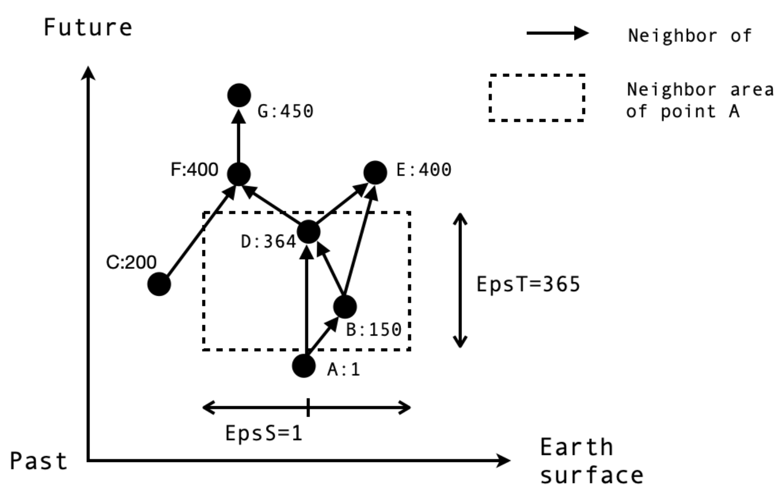

Figure 7. Context of a spatiotemporal relationship

Diffusion theory. According to Midlarsky (1980), on a global scale, violence occurs in a process of risk contagion where the likelihood of an attack increases if neighboring locations experienced violence recently. We examine this risk contagion process through the perspective of diffusion theory, which is described in detail, below. Cluster growth, movement, and change are referred to as diffusion, which is an important theoretical process in chapter 4, where a spatiotemporal density-based clustering algorithm is used to detect emerging clusters. In this dissertation, emerging clusters are defined as new networks of attacks against humanitarian assistance that have emerged since July 2021.

The full series of attacks that emerge in a spatiotemporal network is referred to as a spatiotemporal cluster, and clusters can grow and interact over time, as depicted in figure 8. Numbers behind points indicate the appearance time of the event. Line weights represent relationships among points in the cluster ($EpsT$=Maximum time distance between consecutive attacks to consider as the same cluster; $EpsS$=Maximum spatial distance to consider as cluster). Cluster change is understood as a process determined by actor decision-making and actions that play out across time and space, which are modeled as a diffusion process through a spatiotemporal network. Clusters can grow and reduce in density across time and space.

Spatiotemporal relationships and diffusion

According to Kuo et al. (2021), there are four types of spatiotemporal relationships in the diffusion process framework (see Figure 8). Panel A depicts direct spatiotemporal reachability: given the two points p and q, if p is a neighbor of q and q is a core point, then p is directly spatiotemporally reachable from q (Ester et al. 1996). Panel B depicts spatiotemporal reachability, where given a point p and a core point q, p is spatiotemporally reachable from q if there is a chain-like series of attacks $k_1, k_2, \ldots, k_n$, where $k_1 = q$ and $k_n = p$, such that k_{i+1} is directly spatiotemporally reachable from k_i, for $1 \leq i < n$, $k \epsilon D_t$. Panel C depicts spatiotemporal connectedness, where given two points, p and q, if a point k exists and both p and q are spatiotemporally reachable from k, then p and q are spatiotemporally connected. Panel D demonstrates indirect spatiotemporal connectedness, where given two events p and q, if a core point k exists such that p and q are spatiotemporally connected to k simultaneously, then p and q are indirectly spatiotemporally connected. Figure 9 depicts spatiotemporal cluster evolution types, which demonstrate how specific points are characterized by both single patterns as well as interaction patterns. Results of the analyses in chapters 3 and 4 investigate both point interactions as well as cluster diffusion patterns.

The three-paper dissertation that follows uses a predictive analysis approach in concert with exploratory spatiotemporal cluster analyses to examine the dynamics of attacks against humanitarian assistance from 1997. The first study (chapter 2) examines the conflict-related and contextual predictors of attacks against humanitarian assistance. The second study (chapter 3) explores the spatiotemporal clustering of attacks between 1997 and 2021, and examines the dynamics of perpetration in major spatiotemporal hotspots. The third and final study (chapter 4) detects and illustrates the distribution of new, emerging clusters of attacks against humanitarian assistance between July 2021 and July 2022, and explores the dynamics of perpetration and cluster diffusion in emerging clusters.

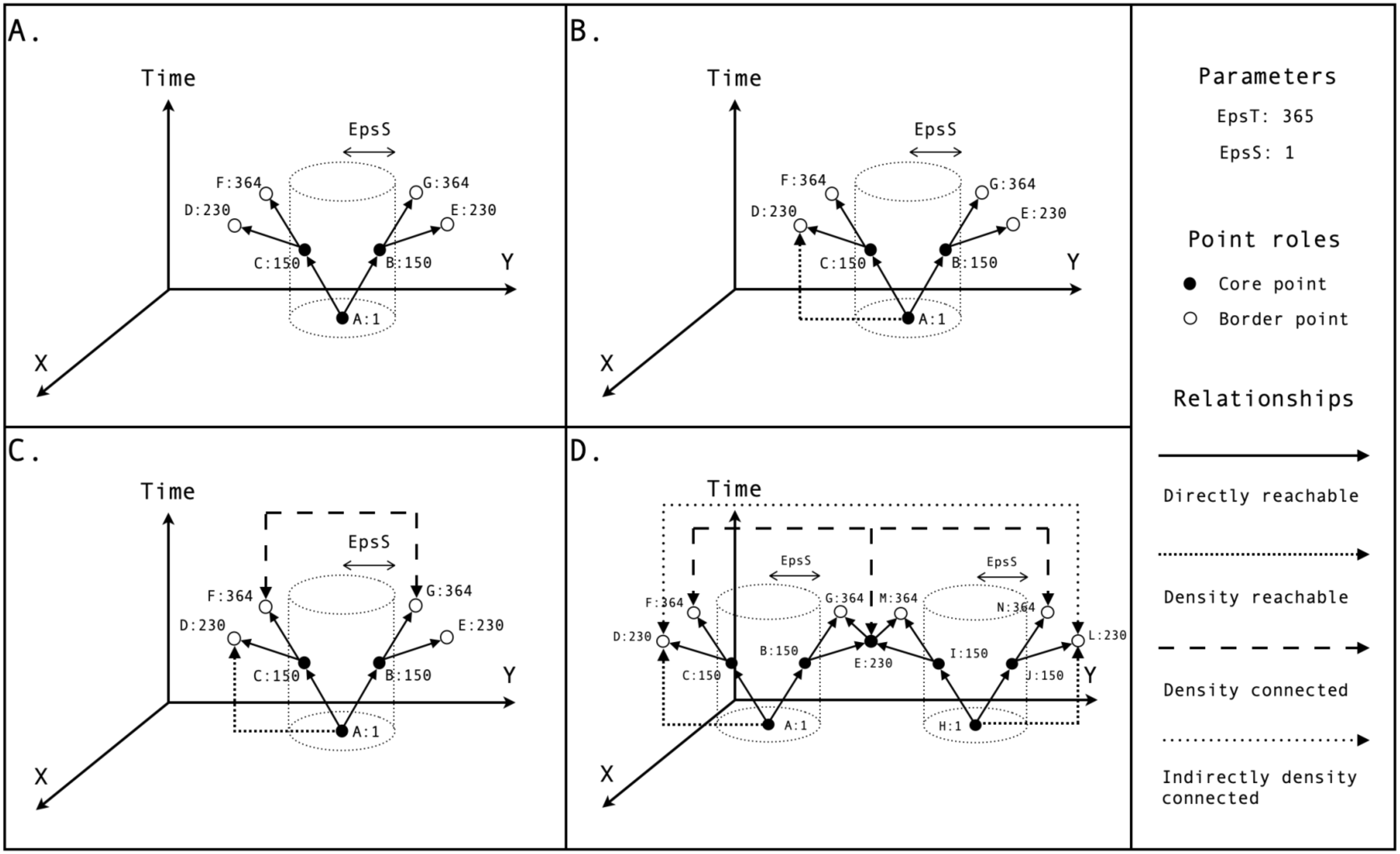

Figure 8. Spatiotemporal diffusion process, including point roles and relationships in a network of attacks

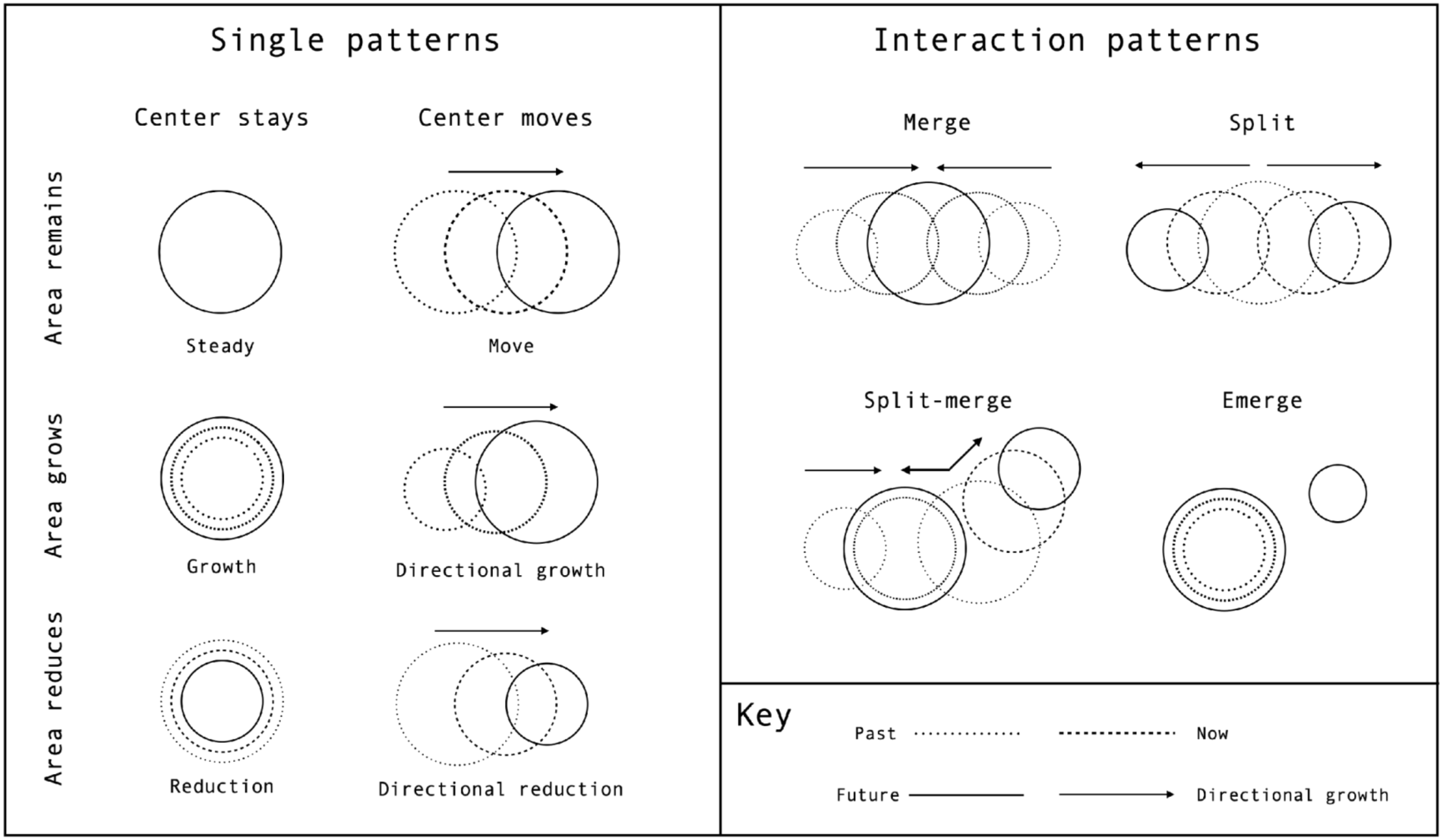

Figure 9. Spatiotemporal cluster evolution types based on diffusion theory, depicting point patterns, density, and interaction

Chapter II. Examining the Conflict-Related and Contextual Predictors of Attacks Against Humanitarian Assistance, Globally, from 1997 to 2021

Abstract

The need for humanitarian assistance is increasing, globally, due to worsening conditions of conflict, climate disasters, and other emergencies. In this study, we use data from the Aid Worker Security Database (Humanitarian Outcomes, 2022) and the Armed Conflict Dataset (to examine the key conflict-related and contextual factors that best predict attacks against humanitarian assistance. Our analytical approach compares the common parametric regression modeling techniques with nonparametric methods like gradient boosting and random forest model ensembles in terms of their relative accuracy in predicting the occurrence of attacks against humanitarian assistance during a given country-year. Our results indicate that the most important predictors of attacks against humanitarian assistance include conflict presence, conflict intensity, and one-sided violence against civilians. The findings presented in this study have important implications for the design of targeted, evidence-based interventions to protect humanitarian operations and prevent unnecessary casualties in, and in the wake of, attacks against aid workers.

In 2022, more than 274 million people – 1 in 29 people, globally, need humanitarian assistance and protection due to conditions of armed conflict, natural disasters, and other crises. Concerningly, however, attacks against aid workers have become a defining feature of humanitarian emergencies in recent decades. In 2021, alone, 140 aid workers were killed and since 1997, more than 1,665 aid workers have been kidnapped, 2,186 killed, 2,308 wounded in attacks on humanitarian assistance. South Sudan, for example, has proven throughout its short post-independence era to be a profoundly complex and insecure operational environment for humanitarian operations (Crombé & Kuper, 2019). Nickerson (2015) describes one attack on a South Sudanese hospital in 2014: "[A]rmed fighters entered the Malakal Teaching Hospital run by Médecins Sans Frontières, robbed patients and their families of cash and mobile phones, and shot those who had nothing to give – killing 14 patients who were lying in their hospital beds" (Nickerson, 2015). In Iraq, a very different context from South Sudan, qualitative interviews with health care providers from 2014 to 2017 found that during the Islamic State (IS) occupation of Mosul, they had been subjected to "terrifying working environments, the strict separation between the sexes, restricted movement, and continuous monitoring by the Al-Hesba morality police" (Michlig, 2019).

These horrific incidents are just two examples of the intransigent problem of attacks against humanitarian operations, both by state forces and non-state actors: since 1997, more than 1,665 people have been kidnapped, 2,186 killed, 2,308 wounded in attacks on humanitarian assistance, and the impact of attacks against humanitarian assistance extends far beyond acute destruction and casualties (Humanitarian Outcomes, 2022). In the wake of an attack against humanitarian assistance, access to civilians in need of critical health and social protection services is often disrupted or discontinued, and in affected communities, the loss of personnel and destruction of critical health and social protection infrastructure is associated with elevated morbidity and mortality, due to preventable conditions like malnutrition, malaria, and diarrheal diseases (Rubenstein et al. 2013).

Evidence from across global conflict settings emphasizes that attacks against humanitarian assistance foment distrust of health and social protection services in the short-term and impact the ability of these systems to function and provide services to populations in need long after an acute emergency ends (Haar et al., 2021). Despite clear evidence of the insidious impacts of attacks against humanitarian assistance, little is known about the conflict-related and contextual factors that predict such events. In this study, we compare the performance of parametric (logistic regression and support vector machines [SVM]) and nonparametric models (i.e., random forest and gradient boosting) in terms of accuracy in predicting country-years where an attack against humanitarian assistance was observed from 1997 to 2021.

Literature Review

Early human rights investigations to present

Some of the first published evidence on attacks against humanitarian assistance includes human rights investigations in El Salvador and Nicaragua, which documented "a pattern of harassment, torture, murder, and disappearance of doctors, dentists, nurses, other health workers, and medical school faculty members, and assaults on health care institutions" (Eisenberg et al., 1983; Geiger et al., 1989; Garfield et al.1987). In 2010, Rubenstein and colleagues published the first global call to action for the protection of humanitarian assistance during armed conflict in the first global qualitative review of the evidence on attacks. Marton (2011) conducted a thorough human rights investigation to document attacks against aid workers occurred during Israel's attack on the Gaza Strip from late 2008 into early 2009. From the early 2010s on, a new wave of human rights investigations and legal analysis about attacks on humanitarian assistance has taken hold, catalyzed in part by the October 2015 aerial bombardment of the Médecins Sans Frontières hospital in Kunduz, Afghanistan by the United States military. The attack on the Kunduz trauma center brought new, global media attention to humanitarian insecurity (Bouchet-Saulnier et al., 2018).

Epistemic streams in the literature

Two broad epistemic streams have emerged in the literature on attacks against humanitarian assistance, and each has made important contributions. The first stream, represented in large part by policymakers and legal analysts, is primarily concerned with legal advocacy and documentation of attacks against humanitarian assistance (Briody et al., 2018; Elamein et al., 2017; Fardousi et al., 2019). To this group of scholars, the purpose of documentation is not to understand intent, but whether attacks are precipitated by parties' failure to comply with rules and protections established to safeguard humanitarian assistance. They invoke a useful analogy from criminal law in the common practice to infer intent from consequence (i.e., although a shooting may not be intended to kill, if the victim dies, a murder charge can still be brought) (Terry, 2013). Scholars in this stream contend that it is difficult and not critically relevant to understand the motivations and intent behind an attack. They argue that even in the absence of clear evidence of intent to harm, attacks against humanitarian assistance subvert the obligations of parties to armed conflict: to distinguish between military and civilian objects, engage in proportional attacks, and take precautions to protect civilians (Mahieu et al., 2020).

A more recent epistemic stream of primarily qualitative research takes a speculative approach, comprising researchers pursuing possible explanations or "causes" of the problem of attacks against humanitarian assistance. This stream of literature points to a range of potential social, political, and conflict-related predictors of attacks, but often without producing corresponding empirical evidence (Haar et al., 2021). Numerous research studies have been done to capture the qualitative experiences of aid workers affected by attacks (e.g., Footer et al., 2018; Tammi, 2021; Singh et al., 2022), but few studies have addressed the perpetrators responsible for attacks, including the context of incidents and whether humanitarian assistance was the intended target of attack. Only three studies have focused on identifying perpetrators of attacks or considering the motives of perpetrators (i.e., Hoelscher et al. 2017; Elamein et al. 2018; and Narang et al., 2017). With the exception of an instrument development study and one research study by Hoelscher et al. (2017) who examined the factors that predict attacks using inferential statistical analysis, very few studies have used quantitative methods to understand attacks against humanitarian assistance.

Evidence from existing research describes attacks on people (e.g., clinicians, project managers, administrators, security guards, ambulance drivers, and translators), obstructions of access to populations in need(e.g., stopping ambulances at checkpoints), discrimination and coercion (e.g., pressure to provide assistance to one individual/group over another), as well as attacks on and misuse of facilities and property (e.g., vandalism, theft, and ambulance theft by armed actors) (Haar et al., 2021). Qualitative interviews with health care workers in Syria underscore the myriad ways conflict causes disruptions in their work, training and resources; barriers and hazards to providing services; mental health risks; and threats to their safety (Abdelrahman et al., 2021). Numerous studies also describe the local impacts of attacks against humanitarian assistance using interview techniques and population-level health indicators. In addition to human rights investigations and research, the evidence has grown to include policy and legal analyses investigating the impacts of attacks in ongoing conflicts in Syria, South Sudan, and others (Elamein et al., 2017; Crombé & Kuper, 2019).

Only one quantitative study, to our knowledge, has examined conflict-related and contextual predictors of attacks against humanitarian assistance. Hoelscher et al. (2017) found evidence of a relationship between conflict presence and intensity, but did not find evidence that aid workers are at increased risk in countries with one-sided violence, stating that the effect is "essentially zero" (14). This, they argue, is consistent with findings from Narang et al. (2017), who suggest that there are different motivations for attacking aid workers versus civilians. Hoelscher et al. (2017) highlight that the absence of one-sided violence as a significant predictor is promising because civilians are particularly vulnerable and dependent on humanitarian assistance in settings with one-sided violence. The authors also found no evidence of an effect of NATO presence as a factor predicting attack risk; however, their analyses found that the presence of peacekeeping operations is associated with elevated risk for attacks against humanitarian assistance.

Methods

Sample

This study compares the aforementioned statistical learning models in terms of their ability to accurately predict attacks against humanitarian assistance during a given country-year using time-series data from a cross-national sample of 170 countries over 25 years (1997-2021) (N=4250 country-years). The primary source of information about attack outcomes is the Aid Worker Security Database (AWSD), an event dataset with details about 3,380 attacks against humanitarian assistance. Attacks include incidents affecting non-governmental organizations (NGOs), the International Movement of the Red Cross/Red Crescent (ICRC), UN agencies belonging to the Inter-Agency Standing Committee on Humanitarian Affairs (FAO, OCHA, UNDP, UNFPA, UNHCR, UNICEF, UN-Habitat, WFP and WHO) as well as IOM, UNRWA, UNMAS and when applicable, the World Bank and locally contracted staff (e.g., drivers, security guards, and others) (Humanitarian Outcomes, 2022).

Information about attacks was sourced, coded, and manually entered by human operators who review the incidents to ensure that they meet the criteria for inclusion. Attacks in the AWSD date to 1997, with more complete information available from 2000-onwards. Historical events are continually updated as new information becomes available, and when a coder cannot determine whether the incident meets the parameters, the incident is referred to two other members of the database team for review and assessment. Attacks affecting contracted workers and vendors of the humanitarian organizations are included if they occur in the course of their work supporting the humanitarian mission. The AWSD does not include violence against UN peacekeepers, human rights workers, election monitors, or purely political, religious, or advocacy organizations (Humanitarian Outcomes, 2022).

Measurement

Theoretical predictors of attacks against humanitarian assistance are presented (see Table 1). All variables are measured at the country-year level. Conflict-related predictors from the UCDP/PRIO Armed Conflict Dataset version 22.0 include total battle-related deaths, state violence against civilians (total casualties), and the type of armed conflict in country-years with conflict since 1997 (Gleditsch et al. 2002; Davies et al., 2022). Contextual predictors of attacks against humanitarian assistance include per capita gross national income (PCGNI), national inequality in terms of the Gini Index, national homicide rate as a measure of generalized insecurity and violence, political risk using the Polity5 score (Marshall and Gurr, 2021), and characteristics of the peacekeeping environment, including PKO presence and size.

In terms of armed conflict and the parties involved, the structure, organization, and leadership of armed forces and groups are thought to motivate attacks against civilians or humanitarian workers in specific contexts. According to evidence from some conflicts, armed forces and actors purposefully attack humanitarian personnel and infrastructure in order to achieve tactical and strategic objectives, in violation of international law and humanitarian principles (Boutton et al., 2018). The strategic use of violence may impact where and how assistance workers are targeted. In addition to armed conflict, another factor that is thought to predict or coincide with attacks against humanitarian assistance is one-sided violence (i.e., violence perpetrated against civilians). Some scholars argue that violence against civilians and humanitarian operations often happens together, and that when civilians account for a large proportion of conflict-related casualties, combatants use violence indiscriminately against non-combatants and may be less hesitant to attack neutral parties (Madhiwalla & Roy, 2009; Hoelscher et al., 2017).

In this study, we examine the following research questions: 1) To what extent are conflict-related dynamics important in predicting attacks against humanitarian assistance; and 2) how, if at all, do contextual factors predict attacks against humanitarian assistance? Hypotheses are stated in non-directional terms due to the fact that the objective of nonparametric modeling (i.e., random forest) is not a one-tailed, directional hypothesis test about the presence of a linear relationship between each predictor and the outcome, but instead to disentangle the relative *importance* of each predictor relative to the *other* variables in the dataset. In this way, the nonparametric models presented in this study do not rely on traditional directional hypotheses, but instead allow nonlinear relationships to exist between predictors and the outcome across the probability range of the outcome. This differs from standard regression techniques, which impose and test the significance of a directional, linear relationship between the outcome of interest and each individual predictor.

Hypothesis 1a. The presence and intensity of conflict will be important predictors of attacks.

Hypothesis 1b. Attacks are more likely to occur in country-years where more one-sided violence is observed.

Hypothesis 1c. Attacks are more likely to occur during country-years where civil war and internationalized civil war is observed.

Hypothesis 1d. The types of armed actors that are involved in conflict are important factors in predicting attacks against humanitarian assistance.

Hypothesis 2a. Gross national income per capita will be an important predictor of attacks against humanitarian assistance.

Hypothesis 2b. National inequality will be important for predicting attacks against humanitarian assistance.

Hypothesis 2c. Political risk will be an important predictor of attacks against humanitarian assistance in a given country-year.

Hypothesis 2d. Peacekeeping activities and budget will be important predictors of humanitarian security in countries where PKOs are present.

Table 1
Operationalization of theoretical predictors of attacks against humanitarian assistance

Predictor	Construct	Operationalization
Conflict-related	Conflict intensity	Battle-related deaths (total BRDs)
	Type of armed conflict	Inter-state
		Extra-state (colonial; conflict with colony)
		Extra-state (imperial; state vs. nonstate)
		Civil war (for central control)
		Civil war (over local issues)
		Regional internal
		Intercommunal
		Non-state war (in non-state territory)
		Non-state war (across state borders)
Contextual factors	Income	Per capita gross national income (PCGNI, per country-year)
	National inequality	Gini Inequality Index (0 low to 1 high)
	One-sided violence	State violence against civilians (total casualties)
	Generalized violence and insecurity	National homicide rate
	Political risk	ICRG polity scores (0 low to 12 very high), absolute and squared
	Peacekeeping activities	PKO presence (1 yes 0 no)
		Peacekeeping size (budget, USD$)

Results

Comparing the performance of parametric and nonparametric models

We compare the performance of parametric (logistic regression and support vector machines [SVM]) and nonparametric models (i.e., random forest and gradient boosting) in terms of accuracy in predicting country-years where an attack against humanitarian assistance was observed from 1997 to 2021. For each statistical learning algorithm, we conducted ten initial model hyperparameter tuning iterations, 30 featuring engineering iterations and 25 final model tuning iterations. Table 2 presents a model performance comparison for the parametric and non-parametric models. The random forest and gradient boosting algorithms were the most accurate—both predicted whether an attack was observed at a rate of 91 percent accuracy in the test sample (N=425). SVM and logistic regression were less accurate and predicted attack outcomes correctly at a rate of 84 and 82 percent, respectively.

Table 2
Parametric and non-parametric predictive model performance

Model	Accuracy	Precision	Recall	F_1	Average Precision (APS)
Random Forest	0.91	0.92	0.40	0.55	0.75
Gradient Boosting	0.91	0.80	0.48	0.60	0.72
SVM	0.84	0.58	0.81	0.58	0.71
Logistic Regression	0.82	0.41	0.79	0.54	0.69

Of the models that we compared, random forest was the most precise, with attacks predicted in 92 percent of the country-years where an attack was observed. In contrast, the gradient boosting, SVM, and logistic regression models were less precise. In the SVM model, attacks were predicted in 80 percent of observed country-years, compared to 58 percent in the SVM model and only 41 percent in logistic regression. The SVM model had the highest sensitivity at 81 percent, closely followed by logistic regression at 79 percent. Gradient boosting and the random forest model were less sensitive, each predicting attacks in 48 and 40 percent of the country-years where an attack was observed, respectively.

The F_1 score is the harmonic mean of precision and recall. The gradient boosting model demonstrated the highest F_1 score, followed by SVM, random forest, and logistic regression. The random forest model had the highest APS, followed by gradient boosting, SVM, and logistic regression. Table 3 presents the random forest model confusion matrix, with the frequencies and proportions of correct and incorrect predictions in the test sample. Values in the confusion matrix correspond to the reported metrics for model accuracy, precision, and recall. As mentioned previously, the random forest model accurately predicted 92% of observed attacks against humanitarian assistance; however, the model predicted that attacks occurred at a higher rate than was actually observed (high false positive rate).

Table 3
Random forest confusion matrix

Observed	Predicted		
	No	Yes	% Correct
No	23	35	39.7
Yes	2	365	99.5
% Correct	92.0	91.3	91.3

The Receiver Operating Characteristic (ROC) curve plots the true positive rate (TPR, the proportion of positive outcomes that are correctly predicted – also known as sensitivity, recall, or probability of detection) on the y axis against the false positive rate (FPR, the proportion of country-years that are falsely predicted to have an attack – also known as fall-out, probability of false alarm, or false discovery rate) on the x axis. The ROC curve represents the threshold for observing an attack across the probability range, and the area under the ROC curve (ROC-AUC) is a measure of how well the model can distinguish between country-years with and without attacks against humanitarian assistance (see Figure 10). Our previous conclusion that the random forest model was the best-performing of the algorithms tested is likewise reflected in the results of the ROC-AUC, which is largest for the random forest model.

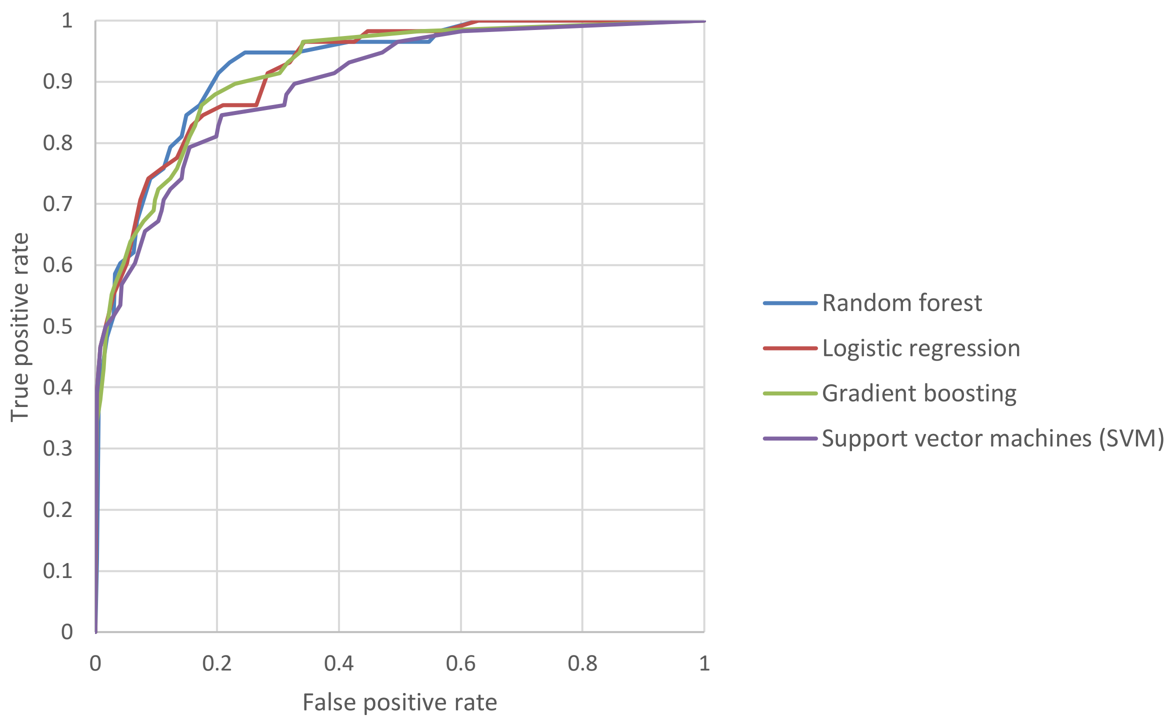

Figure 10. Receiver Operating Characteristic (ROC) curve for all models

The precision versus recall curve demonstrates the proportion of country-years where an attack was observed to those where an attack was predicted (also known as precision) on the y axis against the proportion of positive outcomes (i.e., country-years where an attack is observed) that are correctly predicted (also known as recall, sensitivity, probability of detection or true positive rate [TPR]) on the x axis (see Figure 11). This metric is preferred over the ROC curve for evaluating attack predictions because of the imbalanced nature of the outcome in this study (i.e., more country-years did not have an attack than did). A perfect model is represented as a point with coordinates of (1,1); a skilled model is represented by a curve that bends towards a coordinate of (1,1); and a no-skill classifier appears on the plot as a horizontal line with a precision value proportional to the number of positive cases in the dataset (in a balanced dataset, 0.5). As demonstrated by the precision-recall curve in Figure 11, the model with the largest area under the curve (AUC) is the random forest model. The random forest model demonstrated the greatest skill in terms of both the ROC-AUC and the AUC of the precision-recall curve.

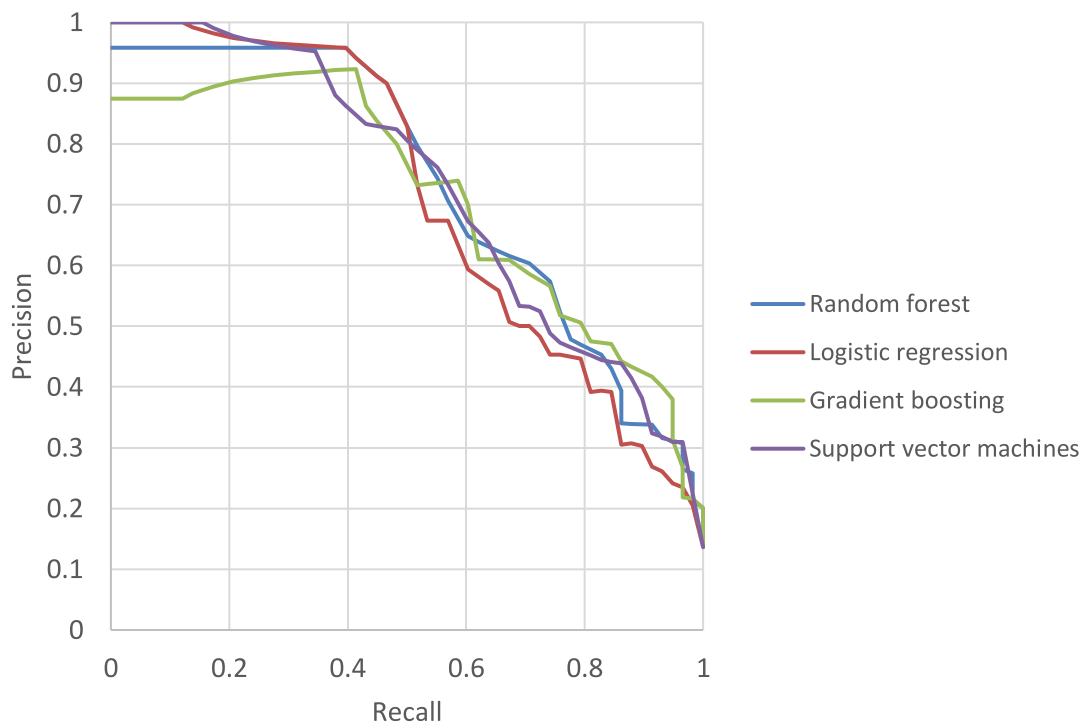

Figure 11. Precision-recall curve for all models

Relative importance of attack predictors

Table 4 presents the importance of the hypothesized predictors of attacks against humanitarian assistance, highlighting the relative contribution of all these variables in predicting whether an attack was observed in a given country-year. The most important variables in predicting attacks during a given country-year were the presence of armed conflict in the form of a governance dispute and the intensity of armed conflict (battle-related deaths). Other important predictors of attacks include the magnitude of violence against civilians (one-sided violence), per-capita gross national income (PCGNI), presence and size of peacekeeping operations, and level of political risk (Polity5). Internationalized intra-state conflict and national inequality were less important predictors of attacks. In contrast, the presence of inter-state conflict, disputes over territory and generalized violence were not important predictors of whether an attack was observed in a given country-year, each with zero percent importance in the best-performing model.

Table 4
Variable importance in predicting attacks

Variable	Importance (%)
Governance dispute	100
Total battle-related deaths	47
One-sided violence	18
PCGNI	15
Population size	14
PKO presence	11
Polity5	9
Internationalized intra-state conflict	8
PKO size	4
GINI index	3

Discussion

The present study provides a number of important insights about the predictors of attacks against humanitarian assistance in a cross-national sample of 170 countries from 1997 to 2021. First, it underscores how the presence and intensity of armed conflict are important factors in terms of understanding the probability of observing an attack against humanitarian assistance during a given country-year. This is consistent with findings from the existing literature and inferential analysis by Hoelscher et al. (2017), who point to conflict presence and intensity as key factors involved in attacks. This makes sense, too, as aid workers are simply more likely to be present and providing assistance in country-years where armed conflict is present. The relative importance of conflict intensity is also logical, since humanitarian aid workers are increasingly encouraged to operate and provide assistance as close to the frontlines of conflict as possible—where conflict is more intense and aid workers are operating on the frontlines, there are more chances that they will experience intentional and/or unintentional attacks as a function of proximity to the violence.

Second, the findings of this study allow us to consider how different types of conflict have differential predictive links to attacks. That is, the nature of the conflict, itself—especially in terms of armed insurrections over national governance—is linked to the safety of humanitarian workers. In addition, in contrast to previous findings by Hoelscher et al. (2017), our analyses found that attacks against humanitarian assistance are indeed predicted by one-sided violence. This finding is in contrast to Narang et al. (2017), who argued that there are different motivations for attacks against civilians and attacks against humanitarian groups. The present study also found, in contrast to the existing theoretical narrative, that contextual factors (e.g., national inequality, population size, and political risk) were relatively less important as predictors of attacks. This finding contrasts with the extant theoretical narrative about contextual drivers of attacks, which places emphasis on political, social, and macroeconomic factors in tandem with armed conflict. However, this is perhaps not surprising, as the settings most affected by attacks against humanitarian assistance (e.g., Afghanistan, Syria, Somalia, and South Sudan) are among the most intense and protracted conflicts in the world.

Third, this study finds clear evidence that nonparametric modeling approaches may be better suited for the task of predicting attacks against humanitarian assistance than parametric regression modeling. The present study found that the random forest model has a better ability to interrogate the complex dynamics driving attacks against humanitarian assistance compared to the typical parametric modeling approaches found in the extant research. The finding that nonparametric analysis techniques (i.e., the random forest algorithm) demonstrate better performance at accurately classifying attack outcomes has both methodological and practical implications, in that these models may be used to inform real-world decision-making by protection stakeholders who supervise and carry out humanitarian operations.

Limitations

Despite the strengths of the methods and insights provided, this study is not without limitations. Most importantly, attacks against humanitarian assistance often go unreported, and there may be disproportionate underreporting of attacks in some of the most remote and risky operational contexts. Thus, our data may be missing important information about some of the riskiest operational contexts for humanitarian assistance. Similarly, what data are available remain incomplete for many cases, which may limit the extent of what we are able to learn from our analyses. These challenges are directly related to the nature of the dataset, which is not a formal surveillance system for attacks, but rather a publicly available database of events that are documented and verified by a team of humans.

Conclusion

This is the first study, to our knowledge, to use a nonparametric predictive modeling approach to understand the conflict-related and contextual predictors of attacks against humanitarian assistance, globally. Findings indicate that the presence of an armed conflict—especially armed conflict in the form governance dispute—and conflict intensity are critical factors in understanding the likelihood of attacks against humanitarian assistance. Somewhat surprisingly, contextual factors, such as sociopolitical and macroeconomic indicators, were relatively less important in terms of predicting attacks. We also see that nonparametric modeling approaches are especially useful in comparison to traditional regression modeling, which has both scientific and, critically, real world relevance. In light of the growing global need for humanitarian assistance due to armed conflict, climate crises, and other disasters, it is increasingly important to critically evaluate the evidence on the dynamics of conflict and the specific contexts in which aid workers are at greatest risk.

Chapter III. Characterizing the Dynamics of Spatiotemporal Hotspots of Attacks Against Humanitarian Assistance from 1997 to 2021: An Algorithmic Modeling Approach

Abstract

Despite the large body of research and published reporting on the impact of attacks against humanitarian assistance and experiences of aid workers who have been affected, few studies have been done to explore 1) how, if at all, attacks form patterns across space and time—and 2) how, if at all, spatiotemporal patterns relate to attack outcomes and perpetrator characteristics. Based on a sample of geo-referenced attack data from the Aid Worker Security Database, this study uses a spatiotemporal clustering algorithm to interrogate the distribution and patterns of attacks against humanitarian assistance, globally, from 1997 to 2021. Results from the study indicate a non-random distribution of attacks against humanitarian assistance, globally, during the 25-year observation period. During this period, four clusters of attacks with 100 or more reported incidents were detected (i.e., Afghanistan, Syria, Somalia, and South Sudan). Smaller clusters were also identified in the clustering algorithm. In addition, the algorithm detected 603 attacks that did not occur in a defined spatiotemporal cluster, which we refer to as "spatiotemporally isolated". The findings from this study broadly indicate that attacks against humanitarian assistance spatiotemporal hotspots form cyclical patterns in time and space, in the form of localized bursts of attacks. Cluster results and diffusion dynamics can be used much like weather patterns, which are used to forecast and understand future patterns. This approach can support evidence-based protection and prevention activities in key risk areas.

A common claim in the literature is that the risks facing humanitarian aid workers are growing, and some scholars suggest that this is because global conflict is becoming more "brutal" (Fast et al., 2010). Although the targeting of protected humanitarian actors during war has been documented throughout history and across contexts, a dearth of rigorous research makes it difficult to verify claims about if and where attack rates are growing over time. A plausible explanation for the perceived increase in overall attack risk may be improvements in data collection and the development of systematic reporting tools, and greater attention to attacks in mass media, which have strengthened the data on the scope and impact of attacks (Hoelscher et al., 2017). Counter to the notion that conflict is becoming more intense or deadly, battle-related deaths have steadily *decreased* globally from 120,000 battle-related deaths per year to 50,000 per year since 2014, whereas attacks against humanitarian assistance have increased over the same period (see Figure 12).

Thoroughly understanding the spatial and temporal patterns of attacks is vital to unpacking the etiology of perpetrator behavior. Although previous research has addressed certain aspects of local and global trends of attacks against humanitarian assistance, the geographical focus of research remains an enduring research gap. That is, while numerous studies highlight qualitative evidence from specific countries and conflicts where attacks against humanitarian assistance are common, more evidence is needed to understand how attacks against humanitarian assistance form patterns across time and space; how, if at all, these patterns relate to conflict-related and contextual factors; and the dynamics of emerging clusters, recurrent incidents, and small-scale attacks. This study aims to detect and visualize diffusion patterns of attacks against humanitarian assistance, first by exploring attack clustering, broadly, and then by examining the evolution of spatiotemporal hotspots.

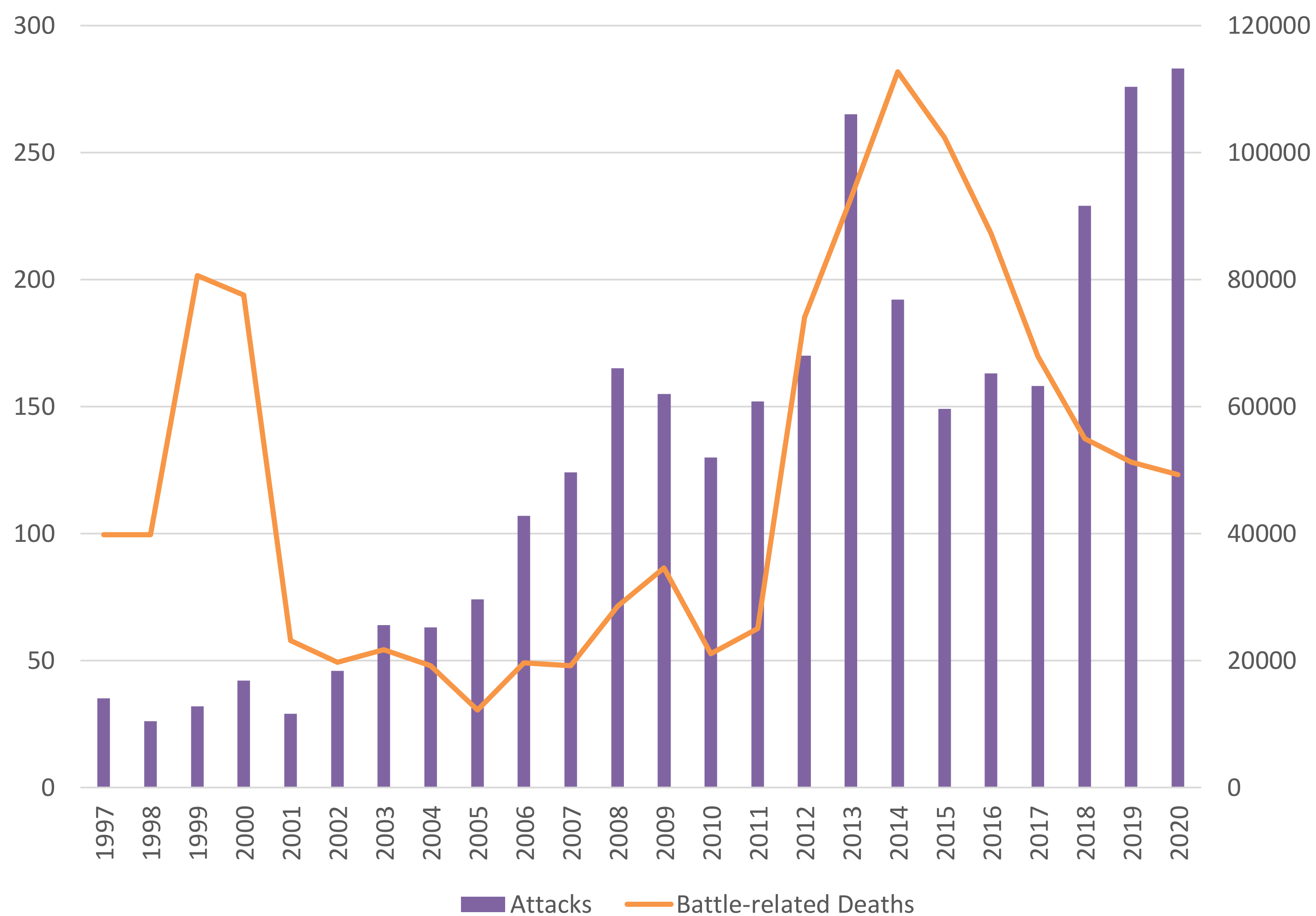

Figure 12. Trend of attacks against humanitarian assistance and battle-related deaths, globally, from 1997 to 2020

Theoretical Framework

Although little quantitative research has been done to examine the spatial and temporal dynamics of attacks against humanitarian assistance, a vast literature across a diverse range of disciplines offers useful theoretical foundation to investigate the distribution and patterns of attacks against humanitarian assistance. This study draws upon existing theory used to characterize the spatiotemporal processes of crime, terrorism, and infectious disease epidemics. Crime, terrorism, disease outbreaks are similar to attacks against humanitarian assistance in that they are point processes that involve stakeholder decision-making, bringing the affected parties together in particular locations in space and time (Felson and Eckert, 2018; Wortley & Mazerolle, 2008; Chainey, 2008). Research on terrorism and armed conflict indicates that actors make an expected cost-benefit analysis; and make bounded, rational decisions regarding their attacks. Intuitively, armed actors often add geography and time into their decision calculus to make rational choices in attacks.

Tobler's First Law of Geography

The theoretical basis of chapters 3 and 4 is Tobler's First Law of Geography (Tobler, 1970). According to Tobler's Law, "everything is connected to everything else, but near things are more related than distant things" (p. 234). When applied to the spatiotemporal analysis of attacks on humanitarian assistance, this implies that two attacks that occur near together are more likely to be similar than those that occur more distant from each other. Through the applied lenses of network theory and diffusion theory, the spatiotemporal analysis investigates how, if at all, attacks against humanitarian assistance form patterns over time and geographical space. Owing to advancements in geographic information systems and statistical modeling technologies, density-based spatiotemporal clustering techniques are becoming more widely used in public health, epidemiological research, and other social and natural sciences. For example, evidence on a global scale demonstrates how conflict events occur in a process of diffusion or risk contagion, where the likelihood of an event increases if a neighboring location experienced an event recently.

Diffusion theory

Cliff et al. (1981) defined expansion diffusion in the context of social, environmental, and health problems as a clustering process that begins at a fixed center (point A) but enlarges its area gradually over time. The authors focused on evolution in terms of the change in a cluster's center and area over time (Cliff et al., 1981). To pick out the cases that should be included at each time step (day), $EpsT$ is the variable that we use:

$$D_t = \{p|t - EpsT \leq T(p\} \leq t\},$$

where D_t is the data set at day t, and $T(p)$ is the appearance day of event p.

Expanding upon Cliff's arguments related to spatiotemporal diffusion theory, Barreto et al. (2008) considered how the geographic area a problem ranges could reflect its severity, concluding that the geographical characteristics of a problem can be used to profile its evolution in detail. Similarly, Lian et al. (2005) examined, in the context of West Nile Virus spread, how a cluster can split into two or more parts – which implies that in addition to single patterns and point relationships, cluster interaction patterns also exist. Spatiotemporal point patterns and relationships are presented in figures 13 and 14. In figures 13 and 14, numbers behind points indicate the appearance time of the point. Maximum spatial distance for two adjacent attacks in a network is defined in the variable $EpsS$. In figure 13, the dashed line around points B and D indicates the spatiotemporal neighbor area of point A.

In quantitative terms, the definition of spatiotemporal neighbors ($STNB_{i,t}$) is:

$$\Delta T(i,j) = T(j) - T(i)$$

$$STNB_{i,t} = j \in D_t | dist(i,j) \leq EpsS \ \cap \ \Delta T(i,j) \leq EpsT \cap i \neq j\}$$

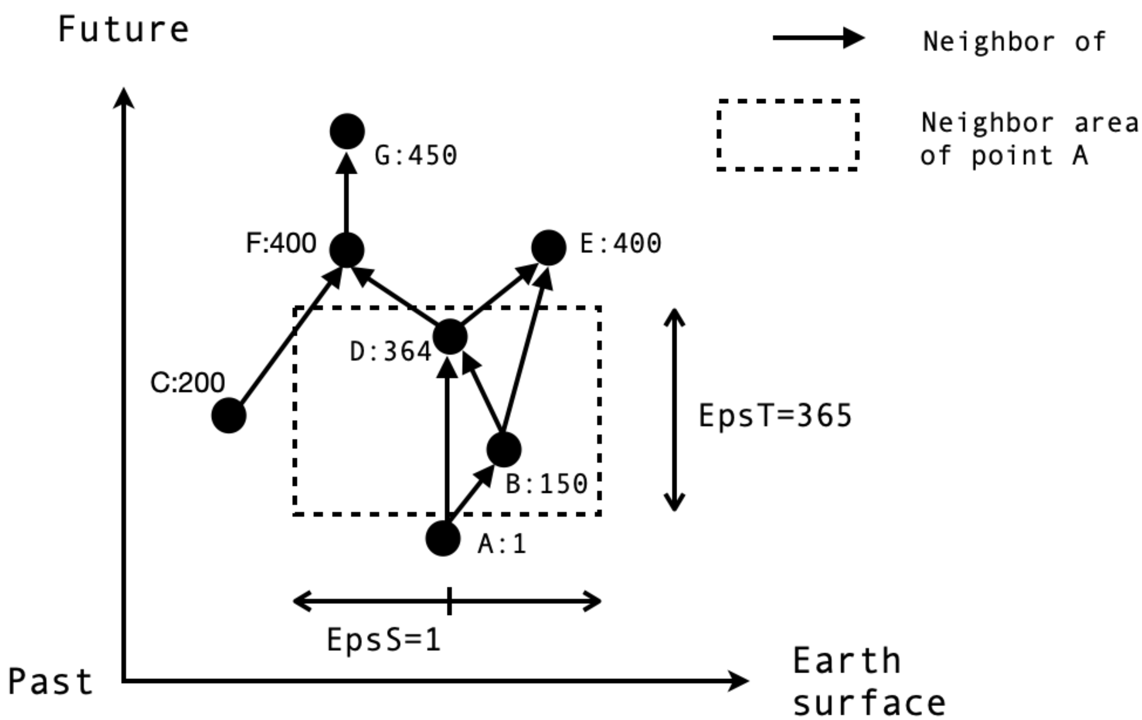

Figure 13. Context of a cluster in terms of spatiotemporal neighbor relationships

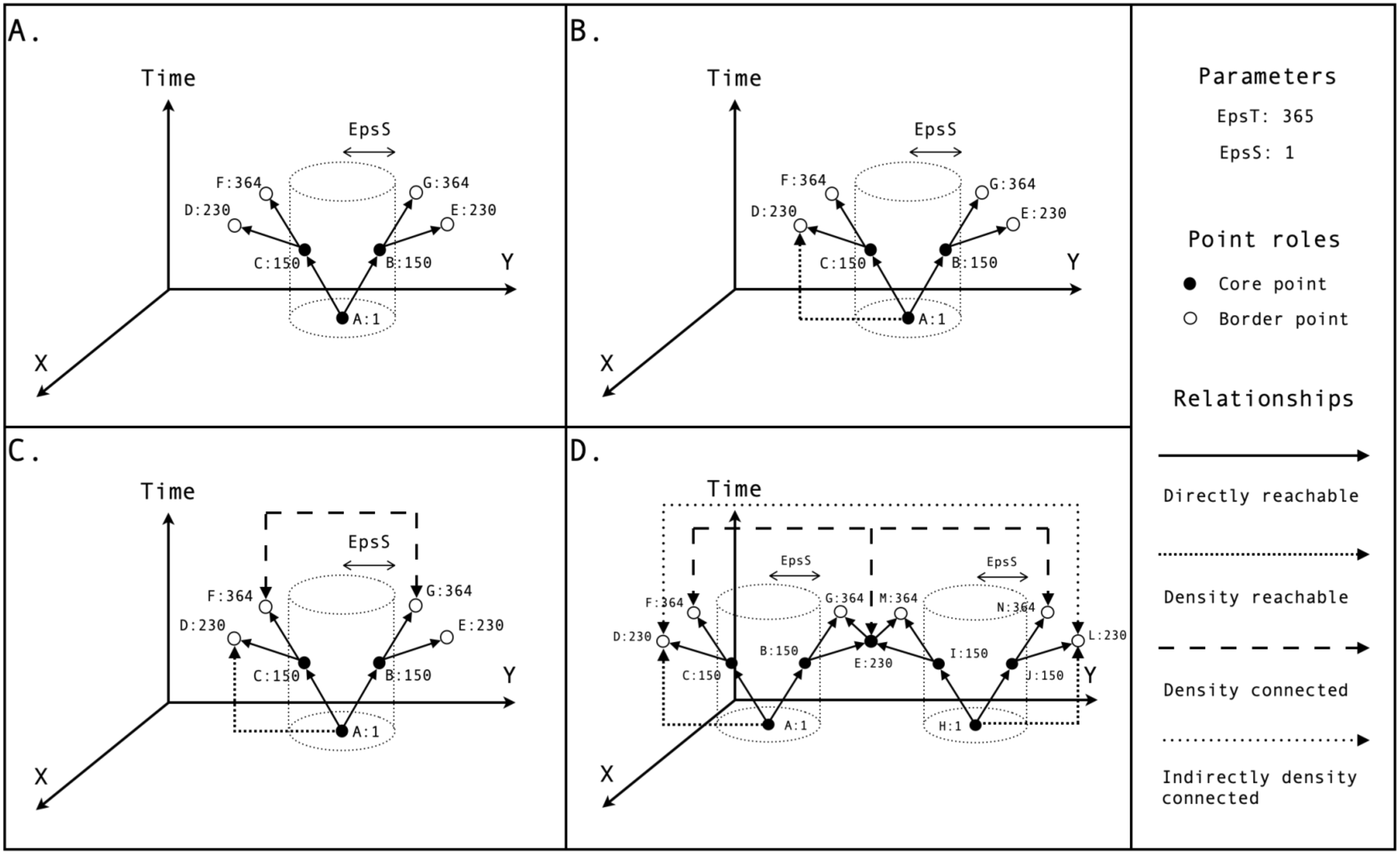

Figure 14. Diffusion process, depicting emergence and interactions in a spatiotemporal network of attacks

Methods

Cognizant of the gaps in the extant literature, the aim of the present study is to 1) understand how attacks against humanitarian assistance form patterns, relate, and diffuse across time and space, and 2) explore the conflict-related and contextual factors involved in attacks in the identified spatiotemporal hotspots.

Design and sample

The present exploratory spatial analysis uses a spatiotemporal density-based clustering algorithm for applications with noise (ST-DBSCAN). The Aid Worker Security Database (AWSD) is a publicly available dataset which includes 3,378 geo-referenced attacks against humanitarian assistance from 1997 to 2022. Spatial data elements in the study include a map with basic geographic features for overlaying the attack data, attack locations in the form of geographical information system (GIS) coordinates, attributes, and time variables. Spatial data are aggregated using a polygon grid of thematic units based on internationally-recognized geopolitical boundaries. To prepare data for the spatiotemporal analyses, we prepared the attack and spatial data sets and identified appropriate maps for the indicated years to layer under attack clusters.

Analytical approach

In this section, we describe all of the steps taken in the spatiotemporal cluster analysis. The complete and annotated density-based clustering algorithm that we implemented using R is presented in Appendix A. The following space and time parameters were used to define clusters and density of attacks against humanitarian assistance: the maximum spatial distance of adjacent attacks in the same cluster ($EpsS$, defined as 1 degree longitude and latitude, or approximately 100 kilometers), the maximum distance for time between events in the same cluster ($EpsT$, defined as 365 days); and the minimum number of attacks to consider as a cluster ($MinPts$, defined as 2 attacks). Using these parameters, $TC = \{TC_1, TC_2, \dots, TC_k\}$ is the set of clusters at time step t and $TC_t = \{C_1, C_2, \dots, C_k\}$ and $D = \{p_1, p_2, \dots, p_n\}$ is the set of spatial points (i.e., attack locations).

We began by building an edge-weighted directed graph (G) with vertices N and empty edges. Then, we created an empty dynamic data set (dd), along with a KD-tree for all points (N) in D. In the KD-tree, we identified all neighborhood pairs whose spatial difference between the two points is equal to or smaller than $EpsS$. Among the spatial adjacent neighborhood pairs, we selected the pairs whose time lag was equal to or smaller than $EpsT$. The earlier point in each pair was marked as the home point and the later point as the neighbor point, and these pairs were inserted into G as the in-flow edges of their neighbor points. At each time step t, old points were removed from dd if they appeared before $1 - EpsT$. From G, taking the out-flow edges whose home point was one of the old points, we added the new points that appeared at time t into dd. For each new point, we identified the in-flow edges and the core point of each of these edges.

We inserted the edges into G as the outflow edges of their origin points, respectively. For each point (p) in dd, we obtained the out-flow edges of p in G, and if the number of these edges was equal to or large than $MinPts$, p was marked as the cluster's core. When p had not been visited in the current time step t, for each core point (p) in dd a depth-first search was used to find all points that were spatiotemporally reachable by p along the out-flow direction. These points were then assigned these points to the same cluster as p and marked as visited. Using a depth-first search, we found one cluster each time and then started the search from an unvisited core point until there was no core point in and dd. When a core point was visited more than once, this implied that it could be reached by two or more clusters, so attacks belonging to these clusters were assigned to the same cluster. We test the null hypothesis that the spatiotemporal distribution of attacks is random, that zero spatiotemporal clusters are observed in the data.

Results

In this section, we present the results of algorithmic cluster analysis which systematically identified spatiotemporal clusters of attacks against humanitarian assistance as well as interaction patterns and diffusion of risk. The ST-DBSCAN algorithm detected 273 distinct spatiotemporal clusters of attacks against humanitarian assistance from 1997 to 2022. As shown in Figure 16, the algorithm identified 4 clusters with more than 100 attacks in the following hotspots: Afghanistan and Pakistan from 2001 to 2021 (N=413 attacks); South Sudan from 2011 to 2021 (N=380); Syria from 2011 to 2021 (N=296 attacks); and Somalia from 2007 to 2021 (N=105 attacks). In addition to the four clusters with more than 100 attacks, the analysis identified 8 clusters with 31 to 100 attacks; 9 clusters with 21 to 30 attacks; 14 clusters with 11 to 20 attacks; and 237 clusters with 2 to 10 attacks.

Although the majority of the attacks occurred in a defined cluster of at least two events, the analysis also detected 603 attacks that did not occur as part of a defined pattern ("spatiotemporally isolated"). These attacks are isolated such that no other attacks were observed within 365-days in a distance of 100 kilometers. Detailed maps are presented to demonstrate the attack clusters. Maps indicate the location and density of attacks (indicated by marker size), as well as the duration of the cluster in terms of days (indicated by the color of the marker, with lighter markers representing older events and dark representing events that were observed more recently). A discussion of the four largest clusters detected in the spatiotemporal clustering algorithm, each with 100+ attacks against humanitarian assistance, follows.

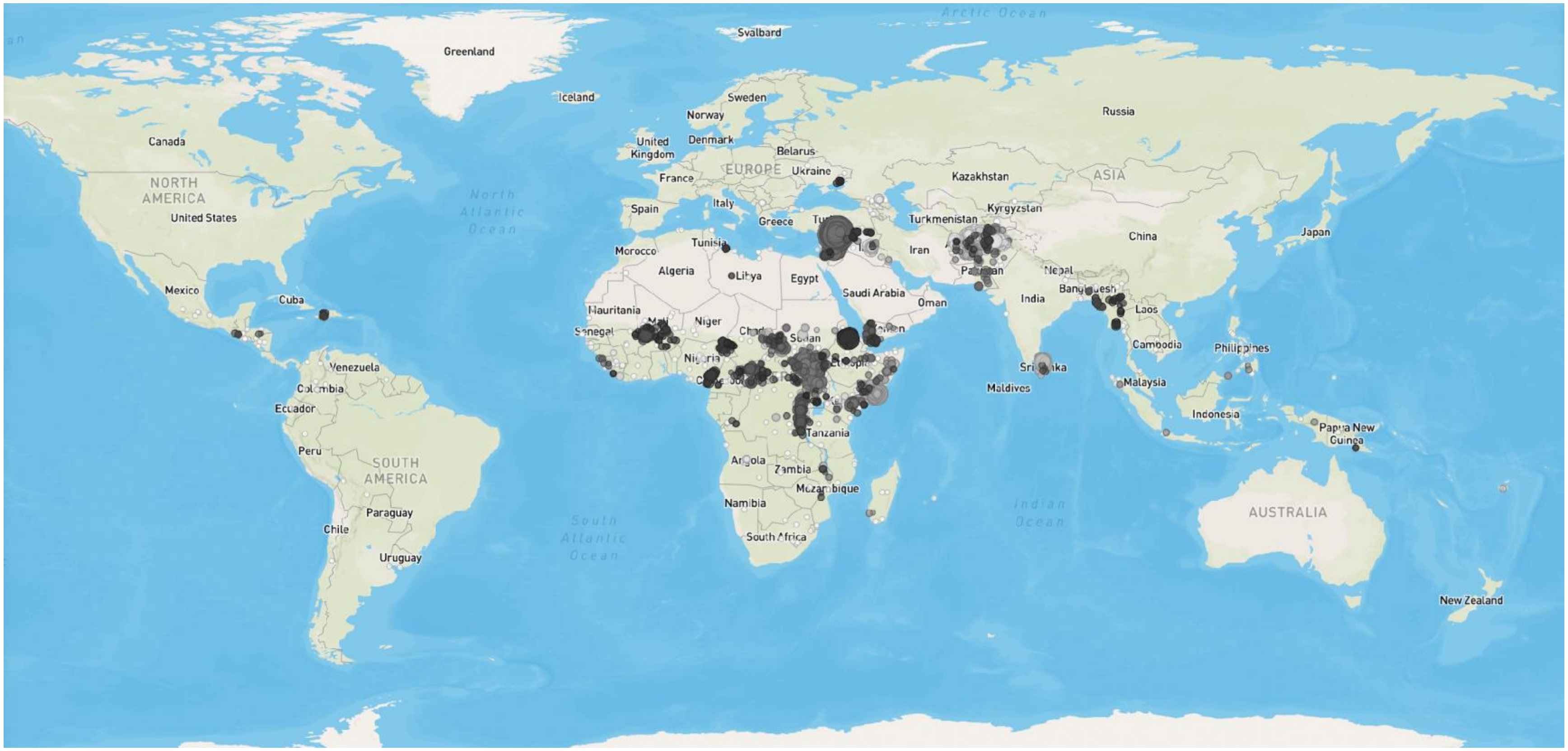

Figure 16. Spatiotemporal clusters of attacks against humanitarian assistance from 1997 to 2022 (N=273)

Spatiotemporal hotspots with more than 100 attacks

Afghanistan and Pakistan from 2001 to 2021 (N=413). The largest spatiotemporal cluster of attacks against humanitarian assistance that the algorithm detected was in Afghanistan and Western Pakistan from 2001 to 2021. The distribution and diffusion dynamics of attacks over time and space are depicted in Figures 17 and 18, respectively. In Figure 18, lighter events indicate attacks that occurred in the more distant past. The map indicates that attacks began in the Waziristan region along the border between Afghanistan and Pakistan in December 2001, then begin to grow and move North and West across Afghanistan while at the same time expanding into Pakistan. After the late 2010s, attacks in this cluster became more densely concentrated around Kabul in the Eastern part of Afghanistan and are observed at a lower rate in Western Afghanistan and Pakistan. The observed spatiotemporal patterns form localized bursts of elevated attack risk over time. To view the diffusion animation (i.e., to watch the spatial patterns play out over time), click on each image.

Figure 17. Diffusion dynamics of attacks through hotspot in Afghanistan and Pakistan from 2001 to 2021 (N=418)

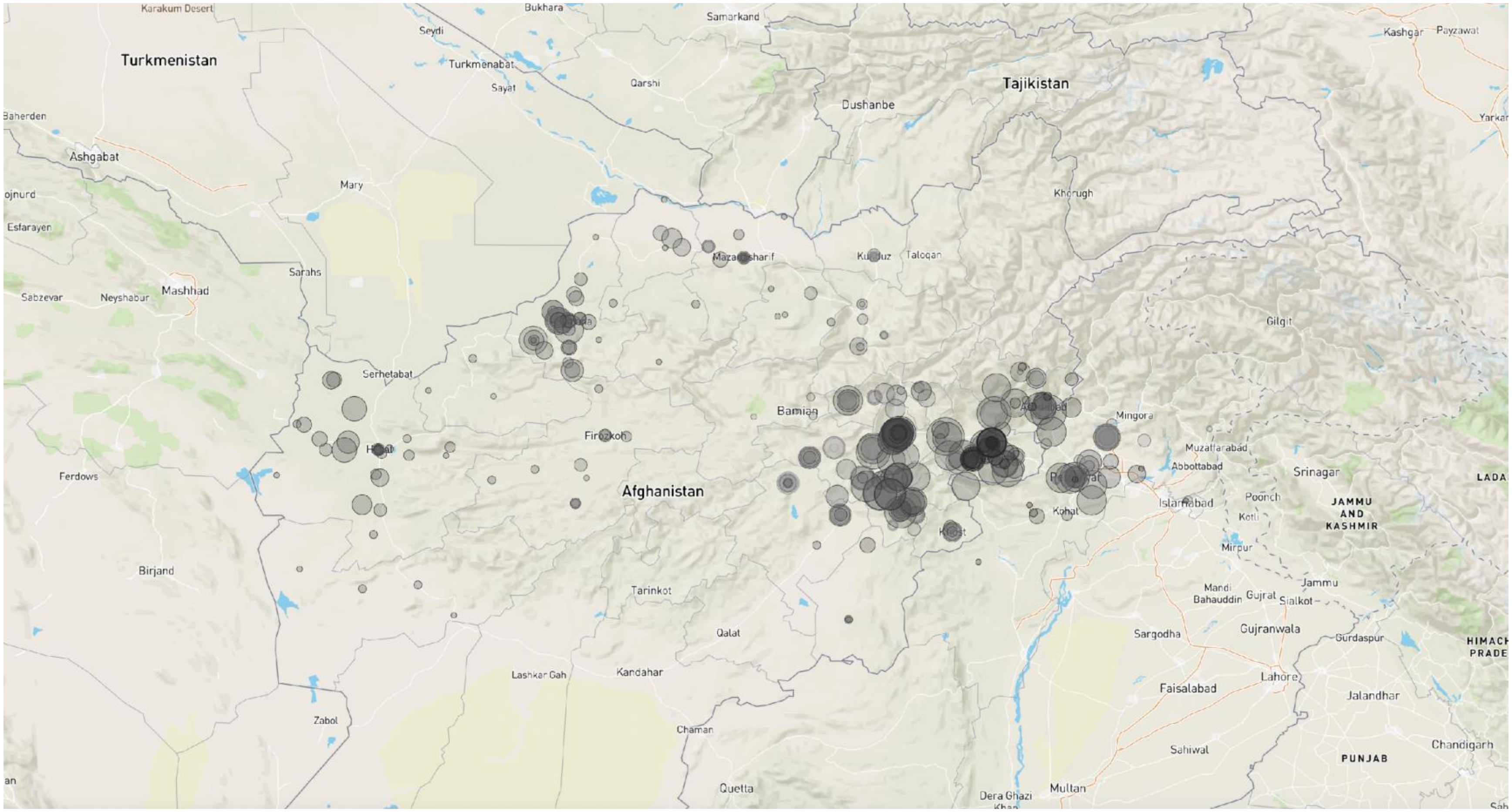

Figure 18. Spatiotemporal density of attacks in Afghanistan and Pakistan from 2001 to 2021 (N=413)

South Sudan from 2011 to 2021 (N=380). The second largest spatiotemporal cluster of attacks against humanitarian assistance ranged from South Sudan to Sudan, Uganda, and Democratic Republic of the Congo from 2011 to 2021 (see Figure 19 and Figure 20). The majority of attacks in this cluster were either shootings (44%) or bodily assaults (38%). Less common means of attack in this cluster included kidnappings (6.58%) and unknown means of attack, which represented 9 percent of events in the cluster. Almost 41 percent of attacks were perpetrated in the context of an ambush, 21 percent were individual attacks, and 18 percent occurred in the context of a raid. Less common contexts in which attacks occurred in this cluster were combat/crossfire (6.58%), mob violence (4.2%), and detention (1.8%). The context of the attack was unknown in almost 7% of incidents. Similarly, the perpetrator was unknown in more than half of attacks in this cluster. Of the remaining half where the perpetrator was known, perpetrators were unaffiliated in 15 percent of attacks, non-state armed group actors (nearly 20%), criminal (3.7%), police or paramilitary (3.16%). Less common perpetrator types included aid recipients (1.8%) and staff members (1.8%). As observed in Figure 19 and Figure 20, attacks occur in microcycles along road passages through South Sudan and concentrating in major population centers like Juba, Bor, and Bantiu.

Figure 19. Diffusion dynamics of attacks through hotspot in South Sudan from 2011 to 2021 (N=380)

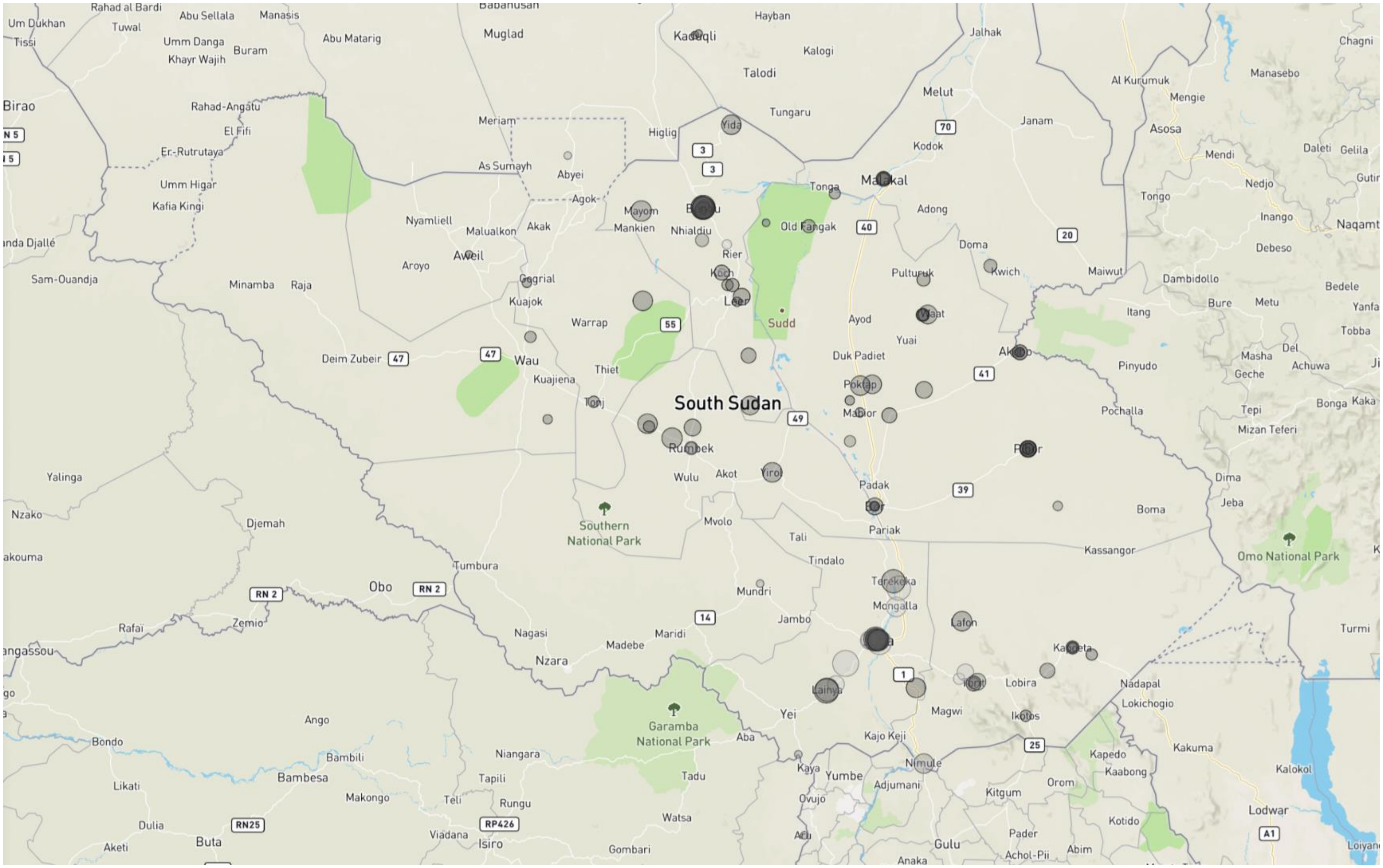

Figure 20. Spatiotemporal density of attacks in South Sudan from 2011 to 2021 (N=380)

Syria from 2011 to 2021 (N=296). The third largest spatiotemporal hotspot identified in this study was located in Syria, with a few attacks also trickling into neighboring countries and territories (see Figure 21 and Figure 22). This cluster emerged in 2011 and continued until 2021, representing 296 attacks. Russia and its coalition with the Syrian government are responsible for 107 aid worker deaths via aerial bombardment in 82 separate airstrikes, as well as 8 aid worker deaths in shellings. The Syrian Armed Forces are responsible for 12 airstrikes, as well, resulting in 5 aid worker deaths. The Russian Syrian Military Coalition and the Syrian Armed Forces are responsible for 93 airstrikes and 22 shellings of aid operations between 2011 and 2021. In addition to airstrikes and shelling, this cluster also includes 49 kidnappings, 8 kidnap-killings, and 28 aid worker shootings. The majority of attacks in this cluster occurred in Northwest Syria, with diffusion to the southwest, in addition to a handful of attacks in the central part of the country that occurred toward the start of the cluster.

Figure 21. Diffusion dynamics of hotspot in Syria from 2011 to 2021 (N=296)

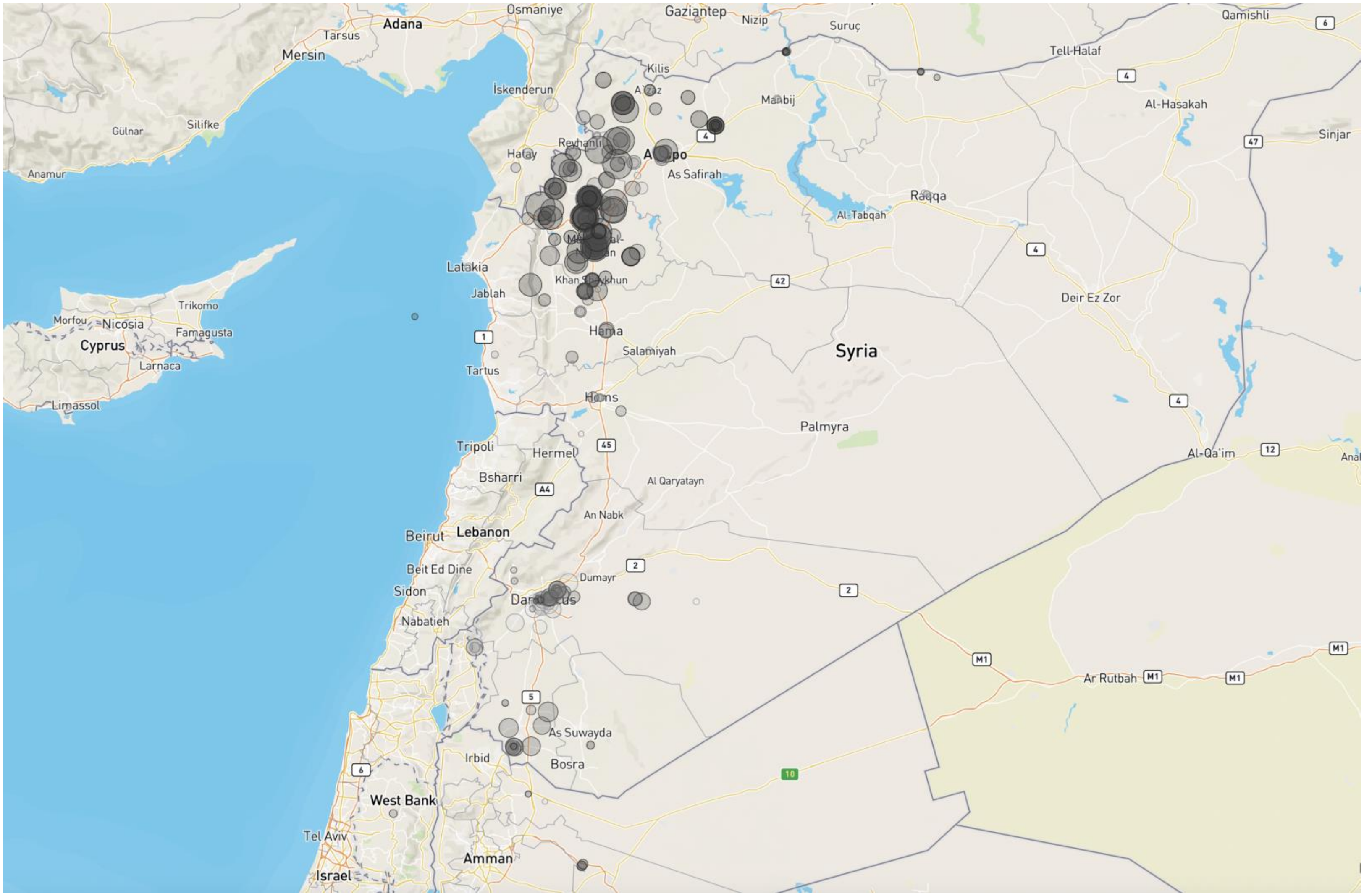

Figure 22. Spatiotemporal density of attacks in Syria from 2011 to 2021 (N=296)

Somalia from 2007 to 2021 (N=105). The fourth largest spatiotemporal hotspot detected by the ST-DBSCAN clustering analysis emerged in southern Somalia in 2011 and continued until 2021, representing 105 attacks (see Figure 23 and Figure 24). In this hotspot, 59 attacks were perpetrated by unknown actors, 24 attacks were perpetrated by Al-Shabaab, 4 attacks by the Somali Armed Forces, 2 attacks by the Transitional Federal Government Forces (TFR), one attack by the police, one by a security guard, one by pirates, one by local government actors, one by clan leaders, and one by the Alliance for the Re-liberation of Somalia. Of the 106 aid workers killed in attacks in this cluster, 99 were nationals and 7 were internationals. In addition to those killed, attacks in this cluster resulted in 64 people being wounded, and 35 kidnappings. Nearly half of the attacks in the cluster were shootings, most by unknown perpetrators. In addition to shootings by unknown perpetrators, this cluster also had 25 kidnappings and 15 kidnap-killings of national and international aid workers.

In the spatiotemporal hotspot in Somalia from 2007 to 2021, Al-Shabaab perpetrated four ambush attacks, six individual attacks, five attacks in combat and crossfire, and nine raids on UN, INGO, and NGO operations. The majority of those killed and wounded in attacks by Al-Shabaab were national staff of the affected organizations. The most common location of Al-Shabaab attacks was in public (N=15), followed by the road (N=4), and the organization's office (N=2). Most of the attacks by Al-Shabaab were complex attacks (N=8) and body-borne and vehicle-borne IEDs (N=11). Additionally, Al-Shabaab was responsible for four shootings in this spatiotemporal cluster, as well as an attack where four national staff of an NGO were beheaded by militants, allegedly for being Christian. The Somali Armed Forces is also responsible for four attacks in this hotspot, including a national NGO ambulance driver who was shot in the leg, an INGO worker who was killed during clashes between two groups of Somali National

Government forces, a UN driver of a truck carrying food aid who was shot and killed, and two local NGO aid workers who were shot at a military checkpoint because their vehicle did not stop when ordered to.

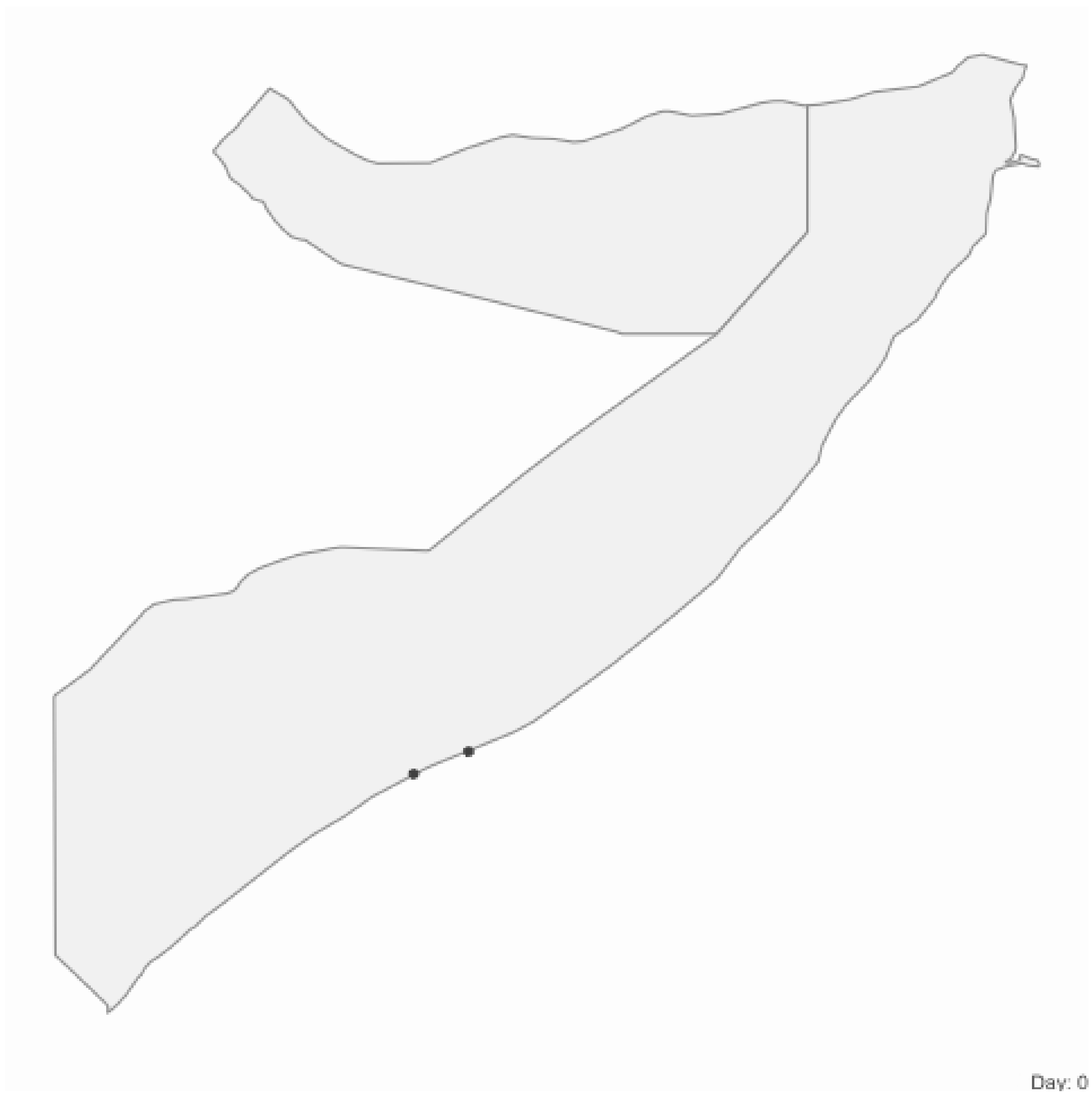

Figure 23. Diffusion dynamics of attacks through hotspot in Somalia from 2011 to 2021 (N=380)

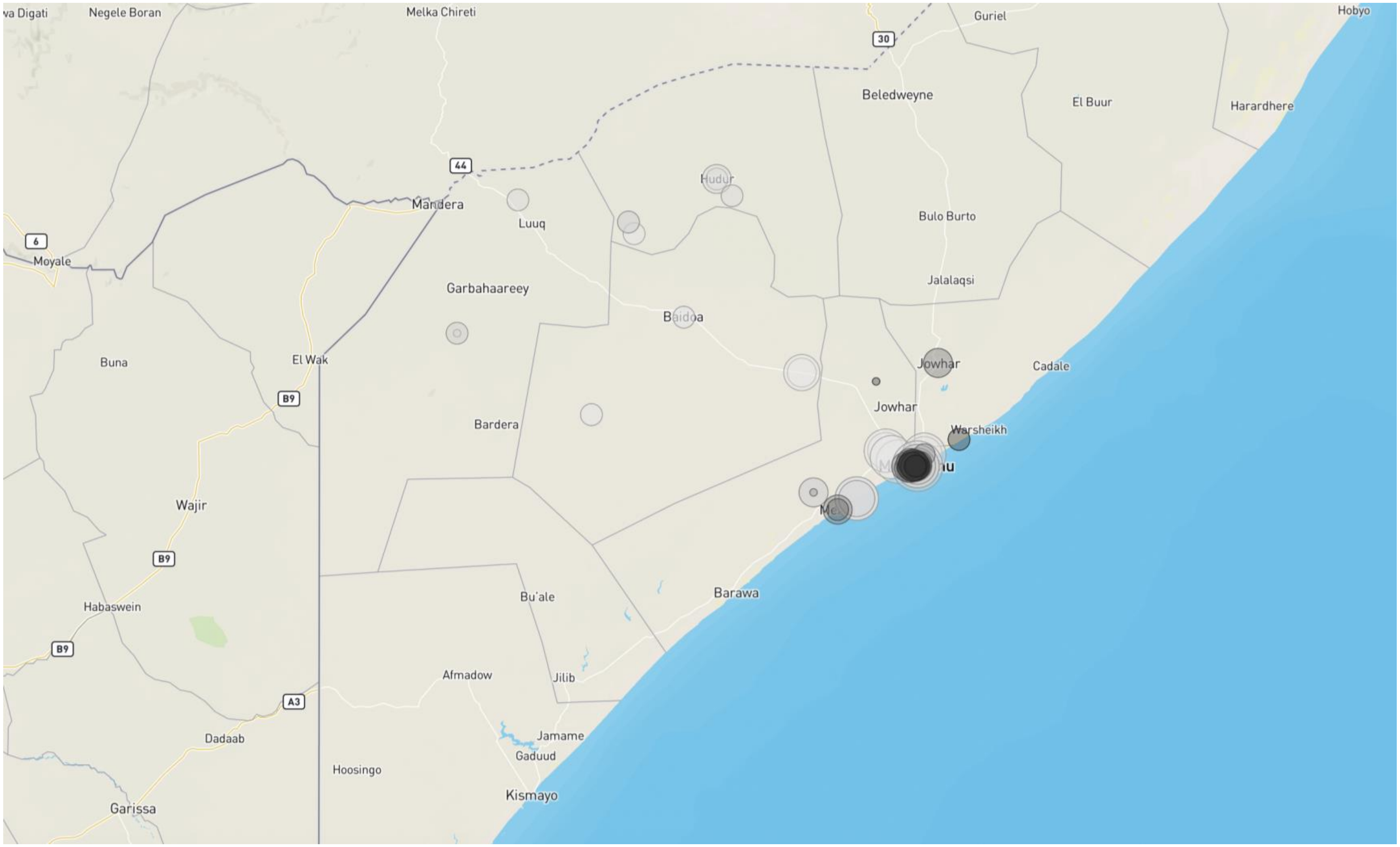

Figure 24 Spatiotemporal density of attacks in Somalia from 2007 to 2021 (N=105)

Discussion

The present study examined the spatiotemporal distribution and clustering of attacks against humanitarian assistance, globally, from 1997 to 2021. The first key finding from this study is clear evidence that attacks against humanitarian assistance are far from randomly distributed over time and space—rather, they tend to cluster and form patterns that evolve across time and space. There were several key areas that emerged as being especially emblematic of this clustering. Our analysis detected major hotspots—defined as clusters with 100 or more attacks—in Afghanistan, Somalia, South Sudan, and Syria. Each of these countries was the site of intense, persistent armed conflict during the observation period—in the form of insurgencies and disputes over government control. In settings like Afghanistan and Syria, conflict disproportionately affects and, in some cases, targets civilian areas, which has placed humanitarian operations in the crosshairs. Likewise, conflicts in Afghanistan and Syria have attracted international military intervention by the likes of NATO allies and Russian coalition forces, which has exacerbated the issue of both indiscriminate and strategic attacks in the form of aerial bombardment and intentional, directed targeting of aid workers in order to undermine the humanitarian imperative. These findings are consistent with the breadth of qualitative literature from Syria and Afghanistan which underscores intentional targeting in addition to collateral damage incurred from conflict conditions (see Abdelrahman et al., 2018; Bouchet-Saulnier et al., 2018).

While there are similarities between these clusters in terms of the number of attacks observed and the means of attack used, they differ in terms of spatial distribution and diffusion characteristics. For example, in Somalia, although there are a few attacks that are observed in the Southwest part of the country, the majority of events—especially in more recent history—are observed in and around urban Mogadishu. This is compared to the other hotspots—Syria, South

Sudan, and Afghanistan—where attacks are more widely distributed across geographic locations over the life of the cluster. In addition, while Afghanistan and Syria exhibit a broader range of attack types (i.e., aerial bombardment, kidnapping, IED and other explosive detonation), the violence in South Sudan and Somalia occurs disproportionately in the form of individual attacks and shootings with small arms, often in urban centers, project sites, and along transit routes.

Another key finding points back to the literature on crime and terrorism as processes of spatiotemporal risk contagion, in that attacks against humanitarian assistance, in certain settings (e.g., Afghanistan and Somalia) occur in cycles of diffusion and repetition across time and space. That is, certain operational locations (e.g., hospitals, health centers, and transit routes) are repeatedly targeted in attacks by multiple types of perpetrators. For example, in certain areas of the cluster in Afghanistan and Pakistan, Taliban actors and US coalition forces carry out attacks against targets, and the perpetration of attacks forms localized bursts of violence—spatiotemporal "cycles" or "loops" of repetition. This may be attributed to the fact that conflict is a dynamic process, and, as conflict changes and moves, the incentives and opportunities of attacks against humanitarian assistance change, as well. On the other hand, repeated targeting of certain sites emphasizes how these locations persist as targets of perpetrators over time. Future research should examine the unique dynamics of perpetration in these persistent high-risk locations.

Limitations

Study findings should be interpreted in light of several limitations. For example, since spatial analysis is not a common use of the evidence in this dataset, the quality and completeness of the data used in this study could be improved. Similarly, general underreporting of attacks may bias the dataset where attacks in higher-risk areas are less likely to be reported. Third, although a strength of this paper is that it focuses on the major hotspots of attacks against humanitarian assistance over the past twenty-five years, we cannot overlook the fact that—in some countries/regions—important attacks against humanitarian assistance have taken place at a lower intensity or less frequent rate to be detected as a cluster. We do not want to lose sight of these smaller, more diffused clusters, as their existence may indicate different dynamics of perpetration compared to larger or more dense clusters. Along these lines, while these data cover a relatively long window of time, they cannot speak to dynamics before the late 1990s.

Conclusion

The findings from this study have important implications for monitoring risk and for evidence-based protection and prevention efforts, in that they illustrate where and under what conditions aid workers are targeted in attacks. Indeed, we see strong evidence that—rather than being randomly distributed across time and space—clear "hotspots" are observed in which attacks are exceedingly common. Interestingly, within these hotspots we see that different types of conflicts seem to result in different patterns of attacks. In broadly internationalized conflicts like those of Afghanistan and Syria, which included aerial bombardment and frequent use of large-scale explosives, attacks were distributed widely across space. In contrast, in South Sudan and Somalia, attacks against humanitarian aid manifested primarily in individual attacks and tended to be limited to particular urban centers and transit routes. Additionally, when we consider patterns across time, we see evidence that important strategic locations (e.g., hospitals, transit routes) are repeatedly the sites of humanitarian assistance attacks, often by multiple actors within a conflict. Given the clear evidence of numerous distinct spatial patterns of attacks against aid workers, future research should examine how, if at all, attacks against humanitarian assistance cluster with conflict events, and how spatiotemporal conflict patterns relate to and/or shape cluster emergence and the diffusion dynamics of attacks against humanitarian assistance in spatiotemporal hotspots.

Chapter IV. Examining Emerging Clusters and Diffusion Dynamics of Attacks Against Humanitarian Assistance Using a Density-Based Clustering Algorithm for Applications with Noise (ST-DBSCAN)

Abstract

The need for humanitarian assistance, globally, is growing rapidly due to emerging conflicts, natural disasters, and other dynamic emergency situations. In light of the growing presence of aid workers in communities around the world and the devastating impacts of attacks against aid workers in affected communities, in this study we explore new ("emerging”) clusters of attacks against aid workers. During the period from July 2021 to July 2022, the spatiotemporal analysis detected eleven new clusters of attacks against humanitarian assistance; in Mali, Burkina Faso, Central African Republic, Myanmar and Ethiopia. These clusters represent locations where no attacks were observed in the adjacent 365-day period within 100 kilometers. Attacks in these clusters indicate a change in incentives or opportunities for attacks against humanitarian assistance and thus a change in the dynamics of risk experienced by aid workers. We present emerging clusters in the form of a web-based tool which demonstrates new spatiotemporal patterns and details of attacks against humanitarian assistance as they emerge.

We live in a dynamic world where new conflicts emerge often. The need for humanitarian assistance in affected communities is constantly growing and changing—in and across national boundaries. In this dynamic global context, risks facing aid workers are complex and likewise dynamic, which makes it especially important for humanitarian policymakers and aid practitioners to pay close attention to how risk changes over time and space. Promisingly, the growing power and accessibility of geographic information systems, georeferenced datasets, and associated software technologies have made it possible for us to track, in real-time, the changing conditions of humanitarian operations. These technological advancements offer a unique avenue for monitoring humanitarian security across global geographies over time. Thanks to advancements in applied geographic information systems and statistical modeling technologies, spatiotemporal clustering techniques are becoming more widely used in criminology, counterterrorism, public health, epidemiological research, and other social and natural sciences to understand how the phenomena like terrorist attacks and disease outbreaks form patterns across space and time. Evidence on a global scale demonstrates, for example, that conflict events, organized crime, and terrorism occur in a process of diffusion or risk contagion, where the likelihood of an event increases if a neighboring location experienced an event recently (Midlarsky et al., 1980).

The emergence of new patterns of attacks against humanitarian assistance, and settings that become active after a period of inactivity, indicate a change in the security environment for humanitarian operations. In areas that had been relatively secure, attacks indicate that the incentives and/or opportunities for attacks against humanitarian assistance have shifted. Effective mapping of emerging clusters and their evolution patterns can be informative to decision-making and risk adjustment in the diverse, complex security environments where humanitarian assistance is provided. In this study, we present a web-based tool which may be used to monitor and investigate emerging clusters of attacks against humanitarian assistance, and their diffusion across time and space, using continuously-updated data from the Aid Worker Security Database. For a geographical area to be flagged as an emerging cluster, an attack must occur where no attack was reported in the preceding 365 days within 100 kilometers, and a second attack must be observed within 100 kilometers and 365 days of the first event, creating a new attack network ("cluster").

Theoretical Framework

Tobler's First Law of Geography

A common concern in regression-based statistical learning models is spatial autocorrelation, or the tendency for events to "cluster" in particular areas. This chapter explores the spatial autocorrelation of attacks against humanitarian assistance through the combined theoretical lens of Tobler's First Law of Geography, network theory, and diffusion theory. Tobler's First Law of Geography states that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970; p. 234). Applied in the context of the spatiotemporal point process of attacks against humanitarian assistance, Tobler's Law would suggest that two attacks that occur closer together are more likely to be similar that those occurring in different clusters. The spatiotemporal analyses in chapters 3 and 4 examine how, if at all, Tobler's Law is validated in the case of attacks against humanitarian assistance, through the applied lens of network theory and diffusion theory.

Network theory. Network theory examines how objects or, in the case of this dissertation, events, create a network or pattern across dimensions (i.e., space and time). Diffusion theory considers how, over time, phenomena spread from their origin (core) point.

According to Meade and Emch (2010), three different types of spatiotemporal diffusion patterns exist – including hierarchy, contagion, and relocation. Each diffusion pattern has a different epidemiological meaning and exhibits a distinct evolution mechanism. Point networks, their diffusion dynamics, and interaction relationships are depictured in figures 25 and 26, where A:1 refers to the first attack at a certain point on the Earth's surface, and subsequent attacks over time occur around it ("neighboring"). Time is a key parameter implied in this theory, where attacks that occur closer together in time and space are more similar to or even related to one another. In this dissertation, this theory is operationalized where two or more attacks are considered part of the same spatiotemporal cluster if they occur within 365 days and 1-degree longitude/latitude (approximately 100 kilometers; see Figure 25).

Diffusion theory. According to Midlarsky (1980), on a global scale, violence occurs in a process of risk contagion where the likelihood of events increases if neighboring locations experienced violence recently. We examine this risk contagion process through the perspective of diffusion theory, which is described in detail, below. Cluster growth, movement, and change are referred to as diffusion, which is an important theoretical process in chapter 4, where a spatiotemporal density-based clustering algorithm is used to detect emerging clusters. Emerging clusters are new networks of attacks against humanitarian assistance that have emerged since July 2021.

The full series of attacks that emerge in the spatiotemporal network is referred to as a spatiotemporal cluster, and clusters can grow and interact over time, as depicted in Figure 25 and Figure 26 (adapted for the present study from Kuo et al., 2021). Numbers behind points indicate the appearance time of the event. Line weights represent relationships among points in the cluster ($EpsT$=Maximum time distance between consecutive attacks to consider as the same cluster; $EpsS$=Maximum spatial distance to consider as cluster). Diffusion of attacks is understood as a process determined by actor decision-making and actions that play out across time and space, which are modeled as a diffusion process through a spatiotemporal network (see Figure 25). Clusters move, grow, and reduce in density across different points in time and space.

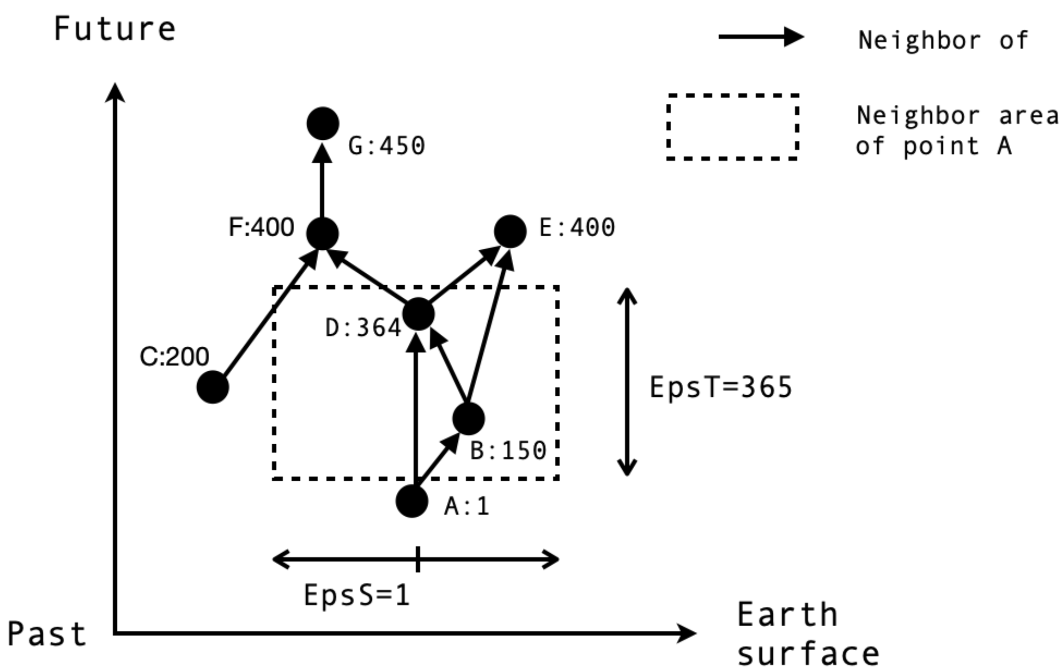

Figure 25. Context of a spatiotemporal relationship

There are four types of spatiotemporal relationships in the diffusion process framework (see Figure 27; adapted for the present study from Kuo et al., 2021). Panel A depicts direct spatiotemporal reachability. Given the two points p and q, if p is a neighbor of q and q is a core point, then p is directly spatiotemporally reachable from q (Ester et al. 1996). Panel B depicts spatiotemporal reachability. Given a point p and a core point q, p is spatiotemporally reachable from q if there is a chain-like series of events — $k_1, k_2, \ldots, k_n$, where $k_1 = q$ and $k_n = p$, such that k_{i+1} is directly spatiotemporally reachable from k_i, for $1 \leq i < n$, $k \epsilon D_t$. Panel C depicts spatiotemporal connectedness, where given two points, p and q, if a point k exists and both p

and q are spatiotemporally reachable from k, then p and q are spatiotemporally connected. Panel D demonstrates indirect spatiotemporal connectedness, where given two events p and q, if a core point k exists such that p and q are spatiotemporally connected to k simultaneously, then p and q are indirectly spatiotemporally connected. Figure 26 depicts these relationships, illustrating how specific points are characterized by both single patterns as well as interaction patterns among points.

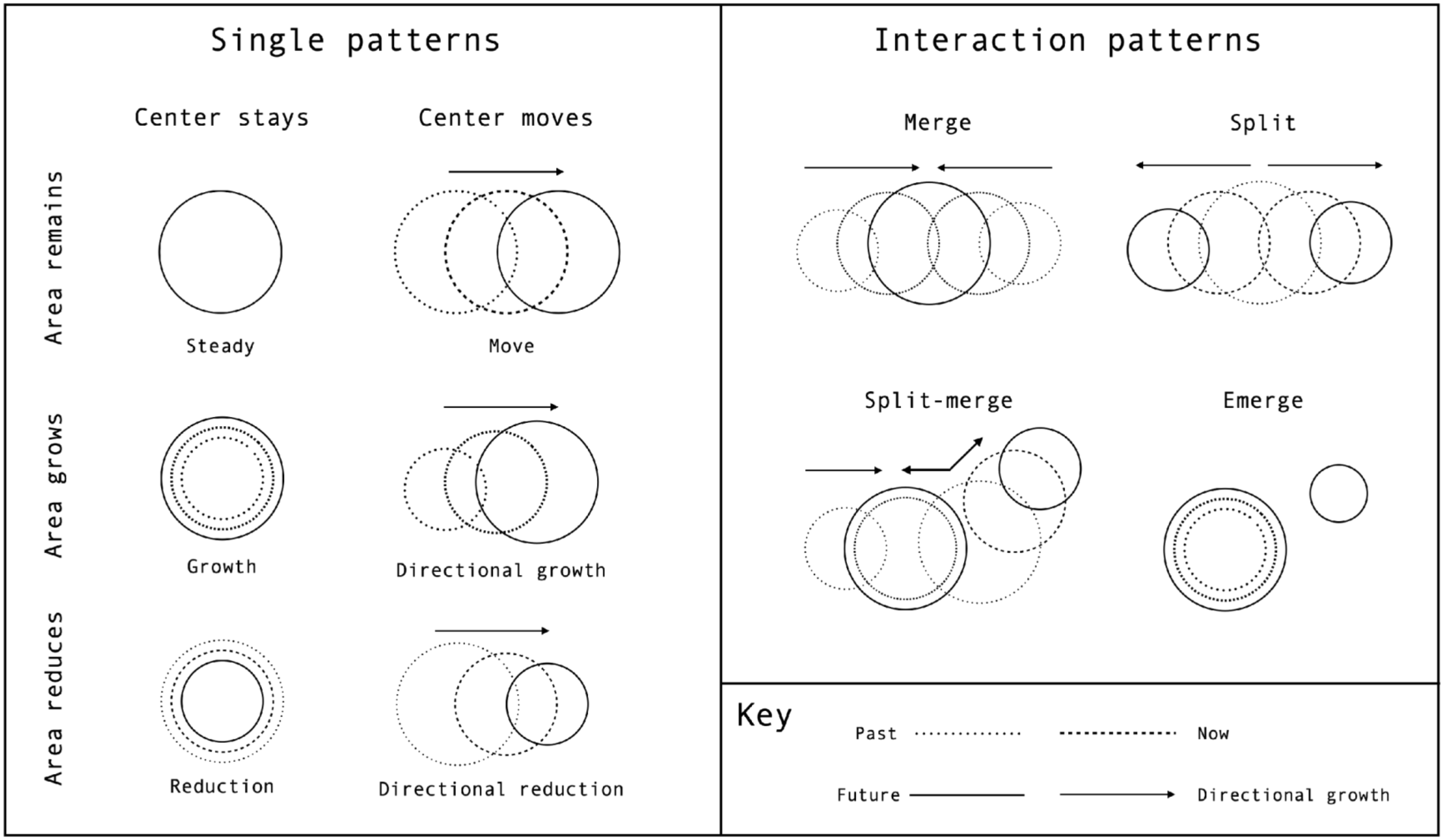

Figure 26. Point patterns and spatiotemporal cluster evolution

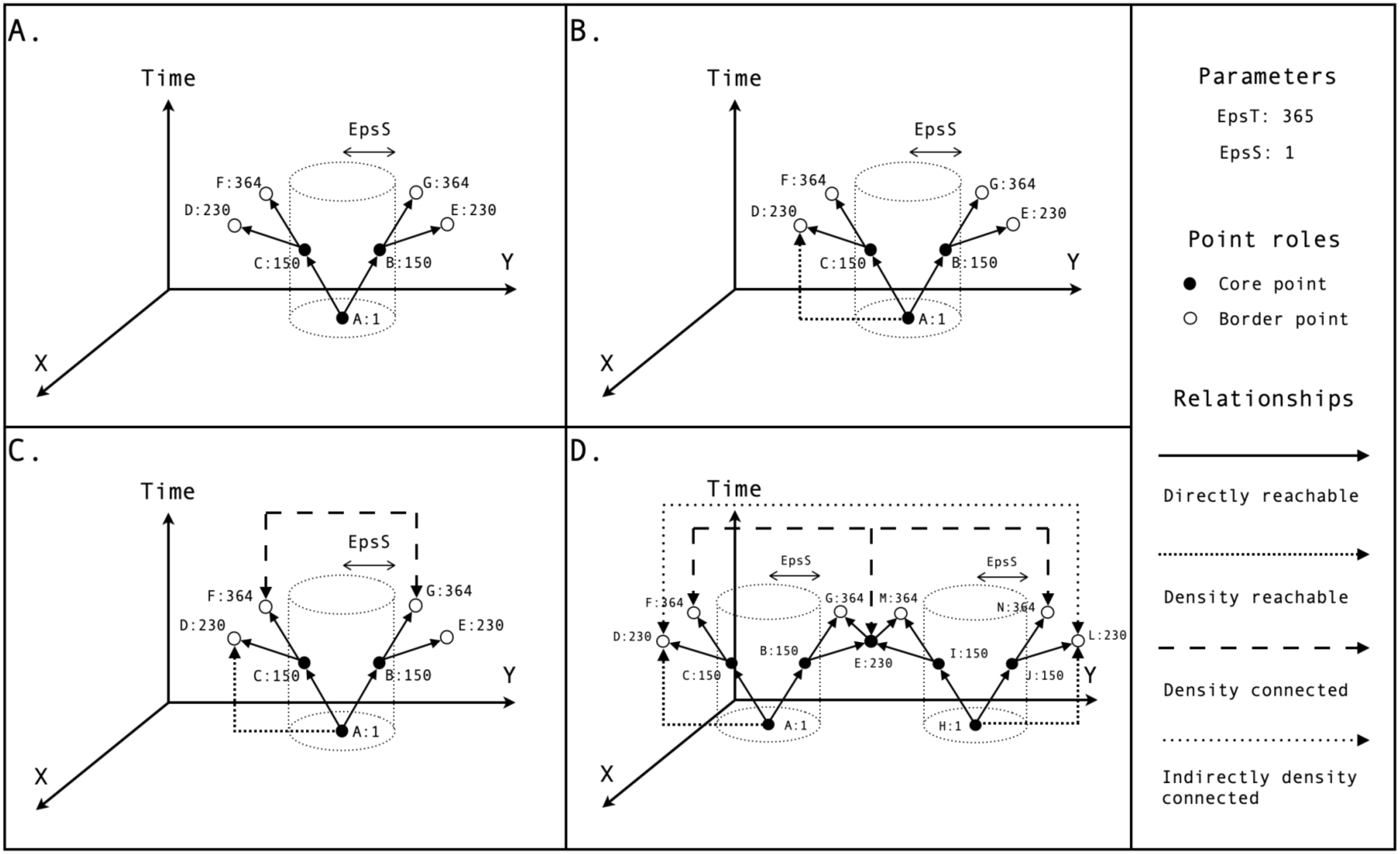

Figure 27. Point roles, patterns, density, and relationships in a spatiotemporal cluster

Methods

Design and sample

The present study uses a spatiotemporal density-based clustering algorithm for applications with noise (ST-DBSCAN) to detect and explore the diffusion dynamics of new clusters of attacks against humanitarian assistance that emerged between July 1, 2021 and July 31, 2022. The full spatiotemporal clustering algorithm (implemented in R) is presented with annotations in Appendix A. The Aid Worker Security Database (AWSD) is a publicly available dataset that includes 3,378 geo-referenced attacks against humanitarian assistance from 1997 to 2022. Geospatial data elements include a map with basic geographic features for overlaying the attack data, geo-referenced attack locations in the form of geographical information system (GIS) coordinates, attributes, and time variables. Spatial data are aggregated using a polygon grid of thematic units based on internationally-recognized geopolitical boundaries. To prepare data for the spatiotemporal analyses, we curated the geo-referenced attack data sets and identified appropriate geopolitical maps for the indicated years to layer under attack clusters.

Measurement

Outcomes of attacks against humanitarian assistance. Attacks are described in terms of density; lethal and non-lethal outcomes; the organization, nationality, and sex of those affected; the means, context, and location of the attack; and the perpetrator responsible for the incident (see table 5).

Table 5
Outcomes of attacks against humanitarian assistance

Outcome	Operationalization
Lethal	1 or more persons killed in attack
Non-lethal	Kidnapping
	Wounding
Organization	United Nations (UN)
	International Non-governmental Organization (INGO)
	International Committee of the Red Cross
	National Red Cross (NRCS)/International Federation of the Red Cross (IFRC)
	National Non-governmental Organization (NNGO)
	Other
Nationality	National aid worker
	International aid worker
Sex	Total females affected
	Total males affected
	Total unknown gender affected
Means	Aerial bombardment (attacks via aircraft, including UAVs)
	Bodily assault (beating without firearms, attack with a knife or club)
	Body-borne IED (body-borne improvised explosive device)
	Complex attack (explosives in conjunction with small arms)
	Kidnapping (released, rescued, or escaped)
	Kidnap-killing
	Landmine or UXO detonation
	Other explosives (i.e., set explosives with a stationary target, lobbed grenade)
	Rape or serious sexual assault
	Roadside IED
	Shelling (mortar, artillery fire, RPG)
	Shooting (small arms/light weapons; e.g., pistols, rifles, machine guns)
	Vehicle-borne IED (unknown whether remote or suicide)
	Unknown
Location	Home (private home, not compound)
	Office (organization compound or project site)
	Project site (village, camp, distribution point, hospital, etc.)
	Public location (street, market, restaurant, etc.)
	Road (vehicle in transit)
	Custody (official forces or police)
Context	Ambush
	Combat/crossfire
	Individual attack or assassination

		Mob violence
		Raid (armed incursion)
		Detention (by government forces or police, where abuse occurs)

Perpetrators of attacks against humanitarian assistance include individuals and organized groups. Individual perpetrators refer to aid recipients, staff members, and unaffiliated persons who carry out attacks against aid workers. Organized groups include criminal organizations (i.e., pirates, cartels, syndicates, mafia, and gangs), states (i.e., foreign or coalition forces, host state, police or paramilitary), non-state armed groups (e.g., global, regional, national, and subnational), and unknown actors (see table 6).

Table 6
Levels of perpetrator organization, types, and descriptions with examples

Level	Type		Description
Individuals	Unaffiliated		Perpetrator acted alone and was most likely not affiliated with an official group or organization (e.g., petty crime, attacks by civilians, mob violence)
	Aid recipient		Aid beneficiary with grievance that results in a confrontation with staff member(s) of supporting aid organization, which leads to a serious attack
	Staff member		A current or former employee with a grievance that results in a confrontation that leads to a serious attack
Organized groups	Criminal		Organized criminal groups like "pirates", cartels, syndicates, mafia, gangs. (including groups whose primary motivation is criminal activity funding and commercial enterprise)
	State	Foreign or coalition forces, host state, police, paramilitary	Officially recognized authorities, representatives, or groups authorized by the state (including military forces, law enforcement and security forces, paramilitary forces like militias, and foreign state entities
	Non-state armed groups	Global	Global in operational scope and scale of ambition (e.g., Al-Qaeda core, Al-Qaeda in the Arabian Peninsula, and the Islamic State)
		Regional	Control or influence over a territory overlapping current national boundaries on ethnic or

		ideological grounds (e.g., Al-Qaeda in the Islamic Maghreb, Boko Haram)
	National	Insurgent groups fighting national government (e.g., the Taliban)
	Subnational (local)	Smaller groups seeking autonomy or control over areas within the existing state (e.g., the Mai Mai militias)
Unknown		Perpetrator cannot be discerned through available information or through verification

Hypotheses

Through the combined theoretical lens of Tobler's Law, network theory, and diffusion theory, this study examines and illustrates the local dynamics of attacks against humanitarian assistance in emerging spatiotemporal clusters. We explore new clusters of attacks that have emerged since July 2021, in terms of the aforementioned attack outcomes as well as the actors involved. An emerging cluster is defined as a series of two or more attacks where no attacks were observed within 100 kilometers in the preceding 365 days or ever before. The null hypothesis of the spatiotemporal analysis is that the distribution of attacks is random—that is, no new clusters ("patterns") were observed in the spatiotemporal observation window. We expect new clusters will have emerged during the observation period, and that they may or may not be related to conflict conditions depending on the location. That said, we expect to find larger clusters in settings where conflict has emerged and intensified during the period (e.g., Ukraine and Ethiopia).

Results

We reject the null hypothesis that the spatiotemporal distribution of attacks against humanitarian assistance between July 2021 and July 2022 was random. The analysis detected eleven emerging clusters, which are depicted in Figure 28. The temporal dimension is represented by a point's color, with the darker points being the more recent attacks in a cluster, and lighter points indicating attacks that are older. The size of the point is scaled to the density of attacks at that point in time and space (i.e., the number of spatiotemporal neighbors the point has).

Emerging clusters of attacks

In the order that they first appeared, new clusters of attacks against humanitarian assistance between July 2021 and July 2022 were detected in Mali; Burkina Faso; Chad, Nigeria, and Cameroon; Uganda; Central African Republic; Democratic Republic of the Congo; Myanmar; Ukraine; and Ethiopia. The cascade of attacks that was observed in each emerging cluster is presented in detail, below, focusing on the local dynamics of outcomes and perpetration. Local profiles offer a better understanding of the factors driving ongoing trends in hotspots and to contextualize the latest developments across regions. In addition, regional profiles afford insights into the proliferation of attacks, affording the ability to monitor and evaluate the continued risk that emerging clusters of attacks pose to humanitarian assistance.

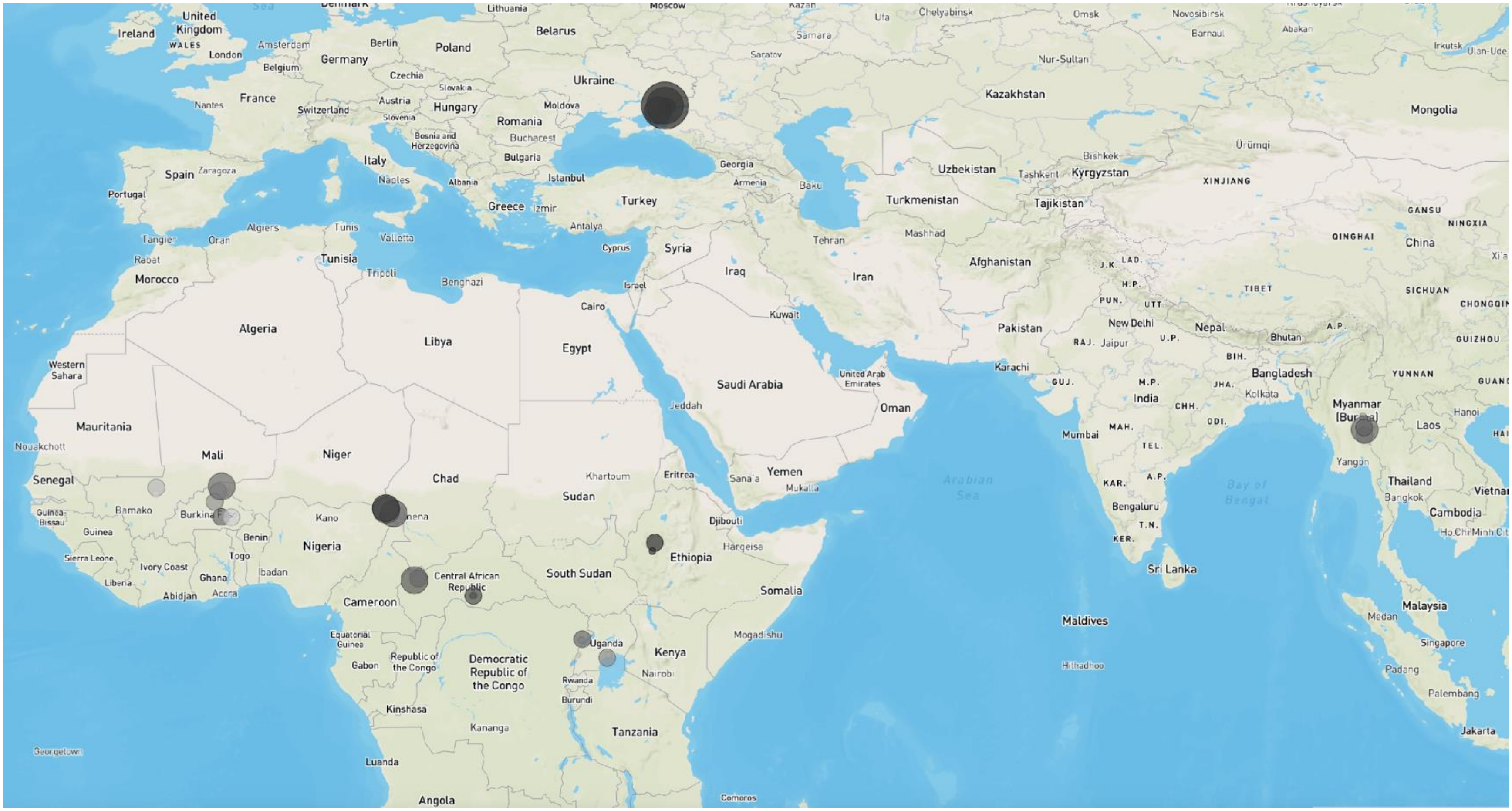

Figure 28. Emerging clusters of attacks against humanitarian assistance, July 2021 to July 2022 (N=11; click image to view clusters)

Mali (N=2). The first cluster to emerge in the period from July 2021 to July 2022 is located in the Segou region of Mali (see Figure 29). The first attack in this emerging cluster took place on July 13, 2021 and involved 4 INGO staff members who were kidnapped by an unknown non-state armed group while travelling 34 kilometers south of Niono, near Heremakono. They were released with their vehicles the following day. The second attack took place on October 7, 2021, close to where the first attack occurred, involving two national NGO staff members who were kidnapped in Niono by members of an unknown non-state armed group. The second attack in this cluster represents directional growth of the cluster to the north.

Burkina Faso (N=7). Two clusters of attacks emerged in separate areas of Burkina Faso during the observation period (see Figure 30 and Figure 31). The first of three attacks by the Jama'at Nasr al Islam wal Muslimin (JNIM) militia in Gourma districts took place on July 14, 2021. In this attack, an INGO staff member was kidnapped along with a teacher and a government official by JNIM Militia while travelling between Matiakoali and Fada Ngourma. On September 11, 2021, a national NGO staff member was kidnapped travelling between Fada Ngourma and Kikideni and subsequently released on September 18, 2021. Most recently, on January 25, 2022, a male NGO staff member was kidnapped by JNIM militia on the road between Fada Ngourma and Pama, Gourma. At the time of the most recent attack, the cluster was expanding directionally westward and appears to be in a split-merge interaction pattern with a nearby cluster in the Sahel region.

On October 13, 2021, a new cluster of attacks against humanitarian assistance emerged in the Sahel region of Burkina Faso, when an NGO driver was injured by unknown criminals who fired shots at the organization's vehicle while they were leaving the Goudebo refugee camp. A month later, on November 15, 2021, another attack occurred when Dawlatoun militia members kidnapped two national INGO staff members while they were travelling in the Sahel region. They were later released. On November 21, 2021, an INGO staff member was injured after the health center in Foube, Sanmatenga was burned down by unknown armed perpetrators. Official reports indicate that this attack was likely collateral damage after the initial attack targeted a nearby police station. Like the other emerging cluster in Burkina Faso, this cluster demonstrates directional growth and spatial expansion over time, and may be in the early stages of an interaction pattern with the cluster in Gourma district.

Chad, Nigeria, and Cameroon (N=5). A cluster of attacks against humanitarian assistance emerged on August 21, 2021 in Chad, Nigeria, and Cameroon—beginning with an attack in N’Djamena where one UN national staff member was injured after he was assaulted during a road robbery on a motorbike in N’djamena. Nine days later, on August 30, 2021, another attack occurred 95 kilometers West of N’Djamena in Nigeria, when a national INGO aid worker was killed along with 16 civilians in a raid by Boko Haram on the organization’s compound in the Borno region. The third attack in this emerging cluster took place on February 24, 2022 in the Logone-et-Chari district of Cameroon, where Boko Haram kidnapped 5 national NGO staff members from their organization residence. The whereabouts of these aid workers and the motive for the attack remain unknown. The attack on February 24 represents directional growth of the cluster in terms of density and diffusion—with the cluster’s center moving north and west (see Figure 32).

On March 10, 2022, a second Boko Haram attack was reported 50 kilometers West of Logone-et-Chari, in the Borno region of Nigeria. In this attack, one national NGO staff member and two guards were kidnapped when their compound was raided. The most recent attack in this emerging cluster took place on June 18, 2022, when three national aid workers were kidnapped from their organization's compound in Gana Ari, in also in Borno region. This attack occurred when militants raided the village and killed 16 civilians in the community before abducting the staff.

Uganda (N=2). A small cluster of two attacks in Uganda emerged on August 27, 2021, when one national NGO staff member was assaulted by unknown criminals in Kampala (see Figure 33). Several months later, on November 30, 2021, a UN staff member was stabbed by a criminal on the way to work in Entebbe, a suburb of Kampala. The new cluster that was detected in Uganda underscores how dynamics of attacks *differ* across emerging clusters. Although several of the other emerging clusters are experiencing conflict conditions, Uganda has not experienced armed conflict during the observation period, and the two observed attacks may be better explained as an issue of localized crime in the urban and suburban areas around the capital, Kampala.

Central African Republic (N=5). In Central African Republic, two distinct clusters emerged between July 2021 and July 2022 (see Figure 34). The first cluster includes a series of three roadside IED attacks by the Retour, Reclamation, et Rehabilitation (3R) armed group and others in the northwest on September 9, 2001; December 16, 2021; and March 2, 2022. In April, two attacks were observed two days apart, when armed actors ambushed aid convoys and robbed national staff members, injuring some critically. This cluster emerged separately from a second cluster to the southeast, in south-central Central African Republic. In the cluster in the south-central part of Central African Republic, the first attack was observed on April 7, 2022, when an INGO convoy was ambushed by Unite pour la Paix en Centrafrique (UPC) actors between Alindao and Bambari (see Figure 35), The aid workers were robbed and five were injured, one critically. The second attack took place two days later only a few kilometers away, when an aid convoy was attacked by unknown gunmen, injuring two national NGO staff. The emergence of two distinct clusters in separate regions of Central African Republic demonstrates how multiple dynamics of emergence and cluster diffusion can occur in the same country over the same time period, with different outcomes for the affected aid workers. Monitoring these trends may be informative to targeted protection and prevention efforts.

Democratic Republic of the Congo (N=2). In Democratic Republic of the Congo, a new cluster of two attacks was detected during the observation period (see Figure 36). In the first attack on October 19, 2021, an NGO staff member was shot and killed in Bunia for unknown reasons. A few months later, in attack nearby on January 18, 2022, an NGO staff member was injured when a gang of commercial motorcyclists threw stones at the INGO vehicle in Ituri. Compared to historical clusters of attacks observed in Democratic Republic of the Congo, this cluster represents a new trend of attacks in the North, suggesting the risk of attacks against humanitarian assistance may be expanding over time and space in this region.

Myanmar (N=4). A cluster of attacks against humanitarian assistance in Myanmar emerged in late 2021 (see Figure 37). The first attack in this cluster took place on November 15, 2021, when a national aid worker was arrested by Myanmar Armed Forces in Pekon, Shan. Reports indicate that he was tortured and killed in custody. Following this attack, on December 24, 2021, two national INGO staff were among at least 35 people found dead following a military massacre by the Myanmar Armed Forces in the So village, Kayah state. The most recent attack in this cluster was observed on February 24, 2022, when a national INGO staff member and two others were injured when airstrikes led by the Myanmar Armed Forces hit the town of Thar Yat while they were attempting to evacuate civilians.

Ukraine (N=5). A cluster of attacks in Ukraine emerged in March 2022, beginning with an attack on March 15, 2022 in which two national aid workers were killed when their office in Mariupol was shelled by Russian tanks (see Figure 38). Subsequently, on March 22, 2022, 15 aid workers were kidnapped when they were ambushed by the Russian Armed Forces as they approached the besieged city of Mariupol. This was the first in a series of three separate kidnapping attacks by the Russian Armed Forces: on April 6, 2022, four aid workers were kidnapped in Berdiansk, and on April 10, 2022, nine NGO bus drivers were kidnapped while attempting to enter Mariupol to evacuate civilians. Reports indicate that the reason the aid workers were kidnapped was that they refused to redirect their evacuation efforts into Russia.

Ethiopia (N=2). The first attack in the most recent cluster to emerge between July 2021 and July 2022 was observed on June 27, 2022, when a national INGO aid worker was kidnapped by an unknown armed group and interrogated about his organization's operation (see Figure 39). He was released the following day. In the second attack, on July 16, 2022, a national NGO worker was kidnapped by unknown actors while traveling in East Wellega, close to the same location as the first incident on June 27. The whereabouts of this aid worker remain unknown.

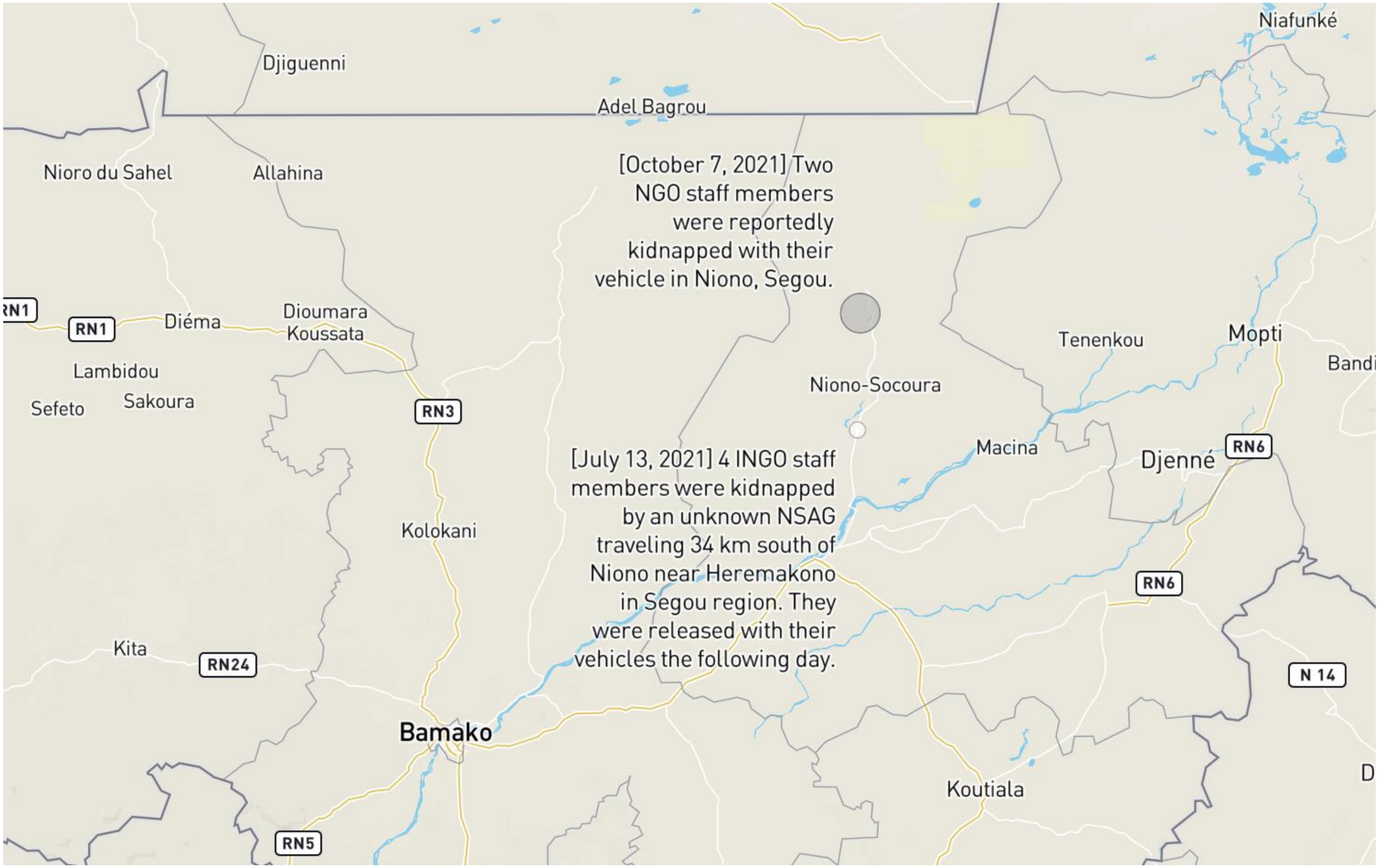

Figure 29. Diffusion dynamics of emerging cluster in Mali (N=2)

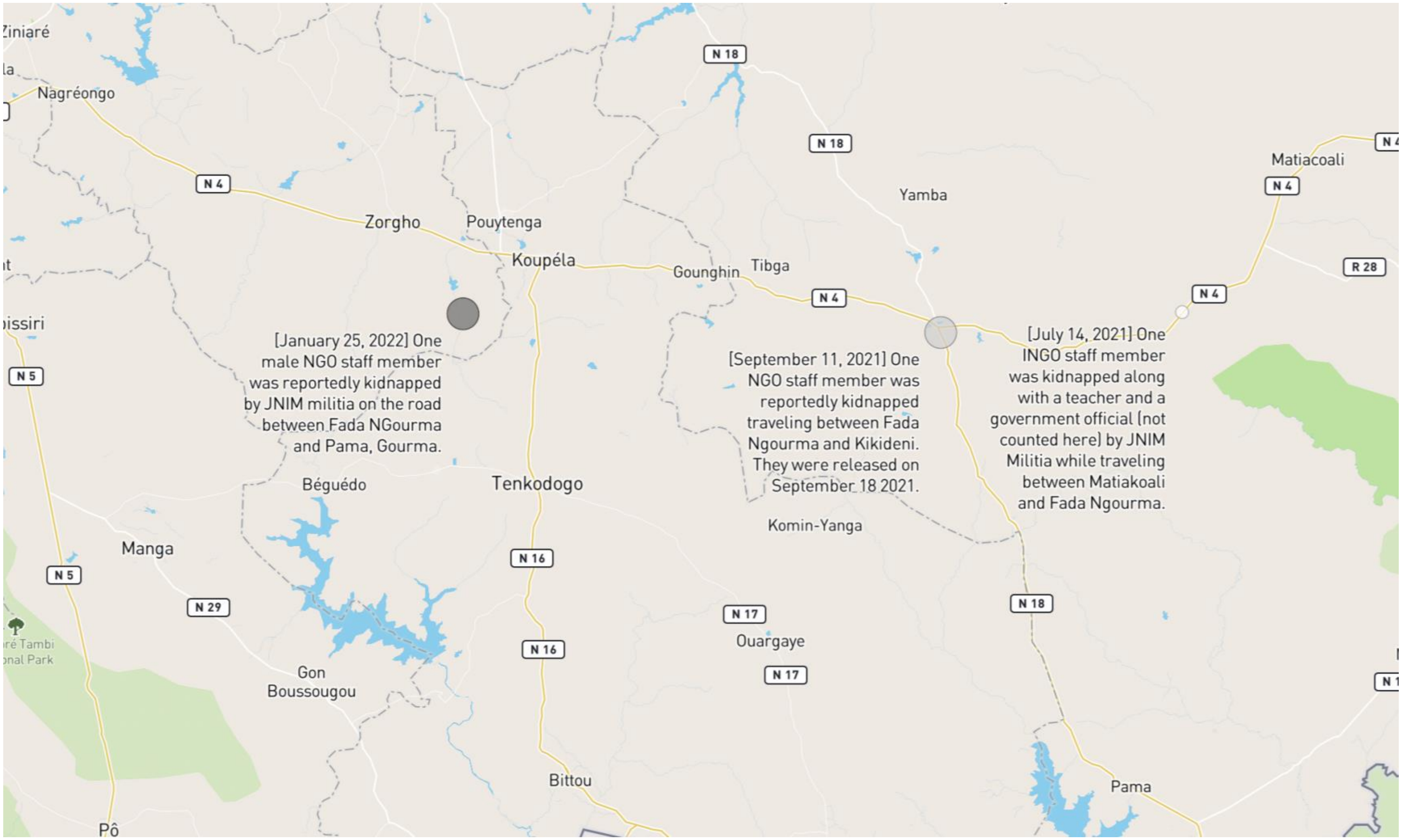

Figure 30. Diffusion dynamics of emerging cluster in Gourma region, Burkina Faso (N=3)

Figure 31. Diffusion dynamics of emerging cluster in Sahel and Centre Nort regions, Burkina Faso (N=4)

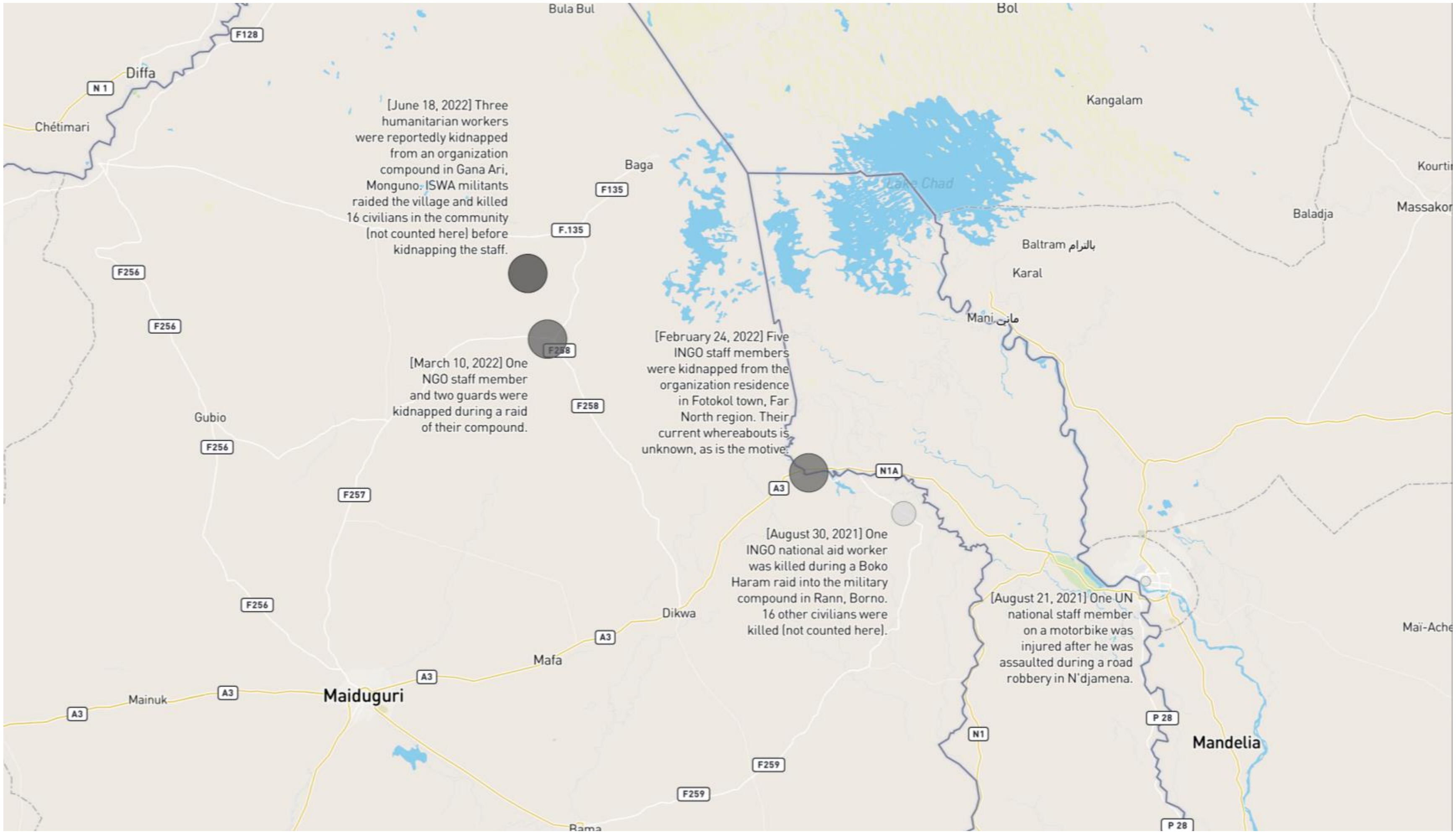

Figure 32. Diffusion dynamics of emerging cluster in Chad, Cameroon, and Nigeria (N=5)

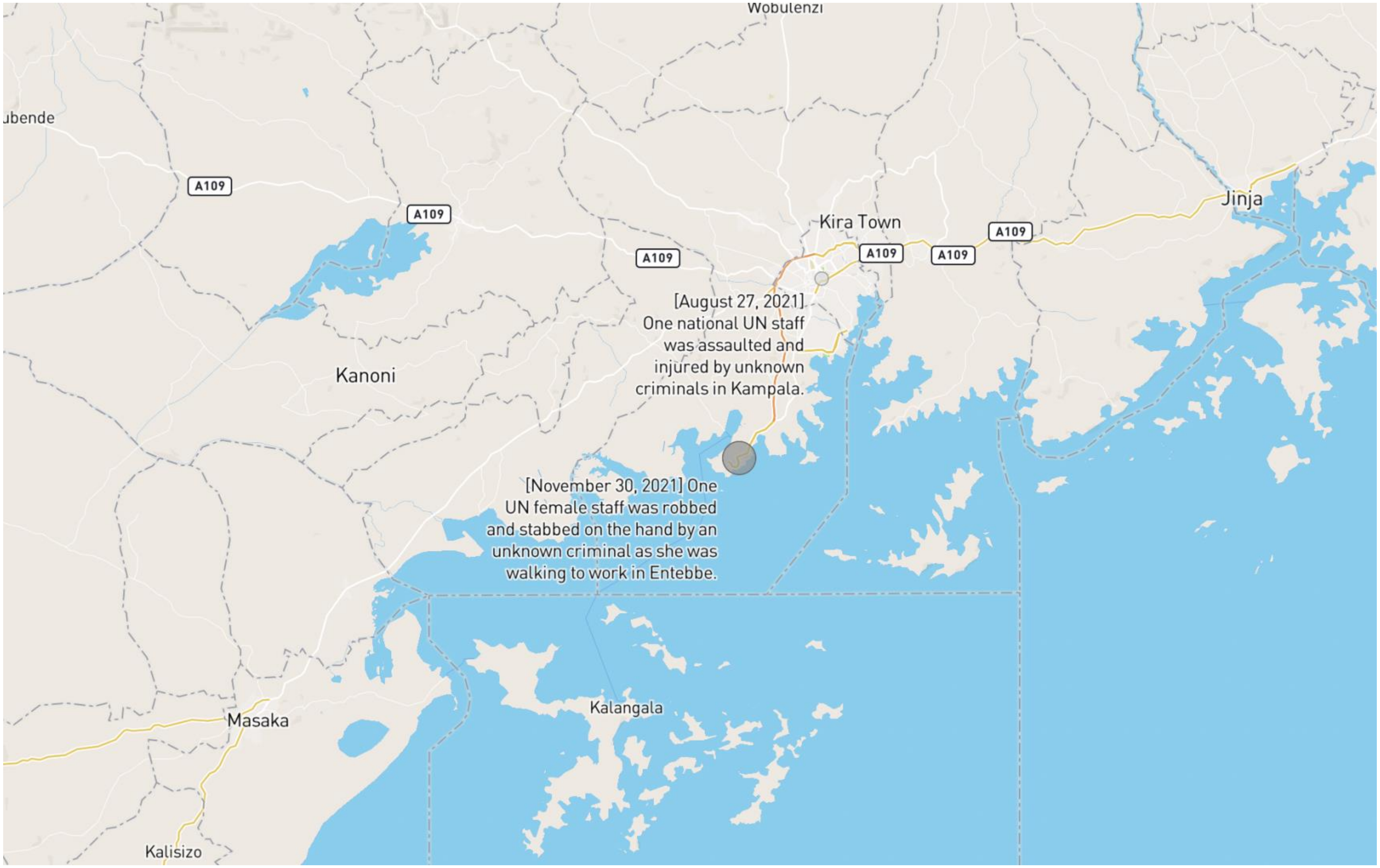

Figure 33. Diffusion dynamics of emerging cluster in Uganda (N=2)

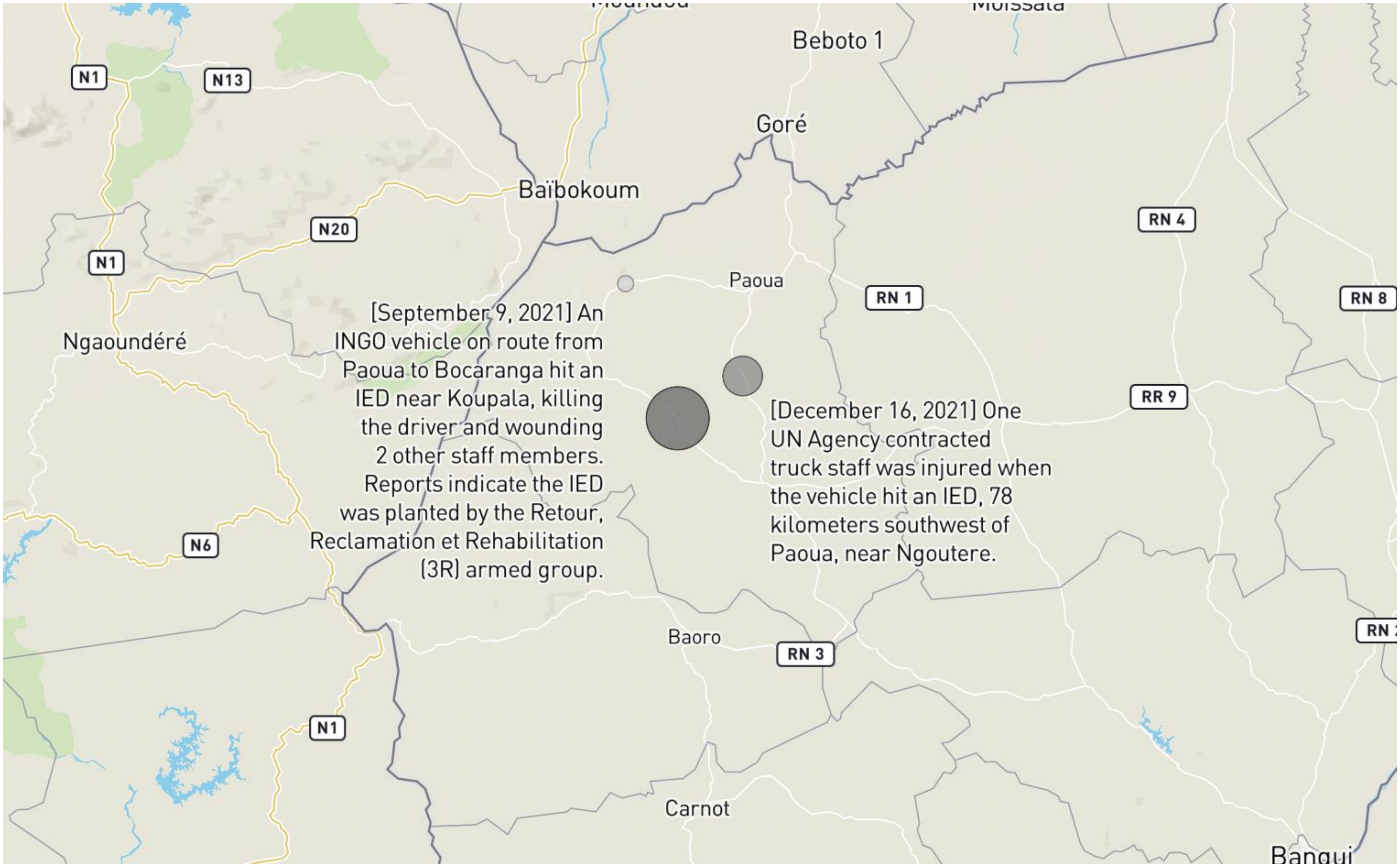

Figure 34. Diffusion dynamics of emerging cluster in northwest Central African Republic (N=5)

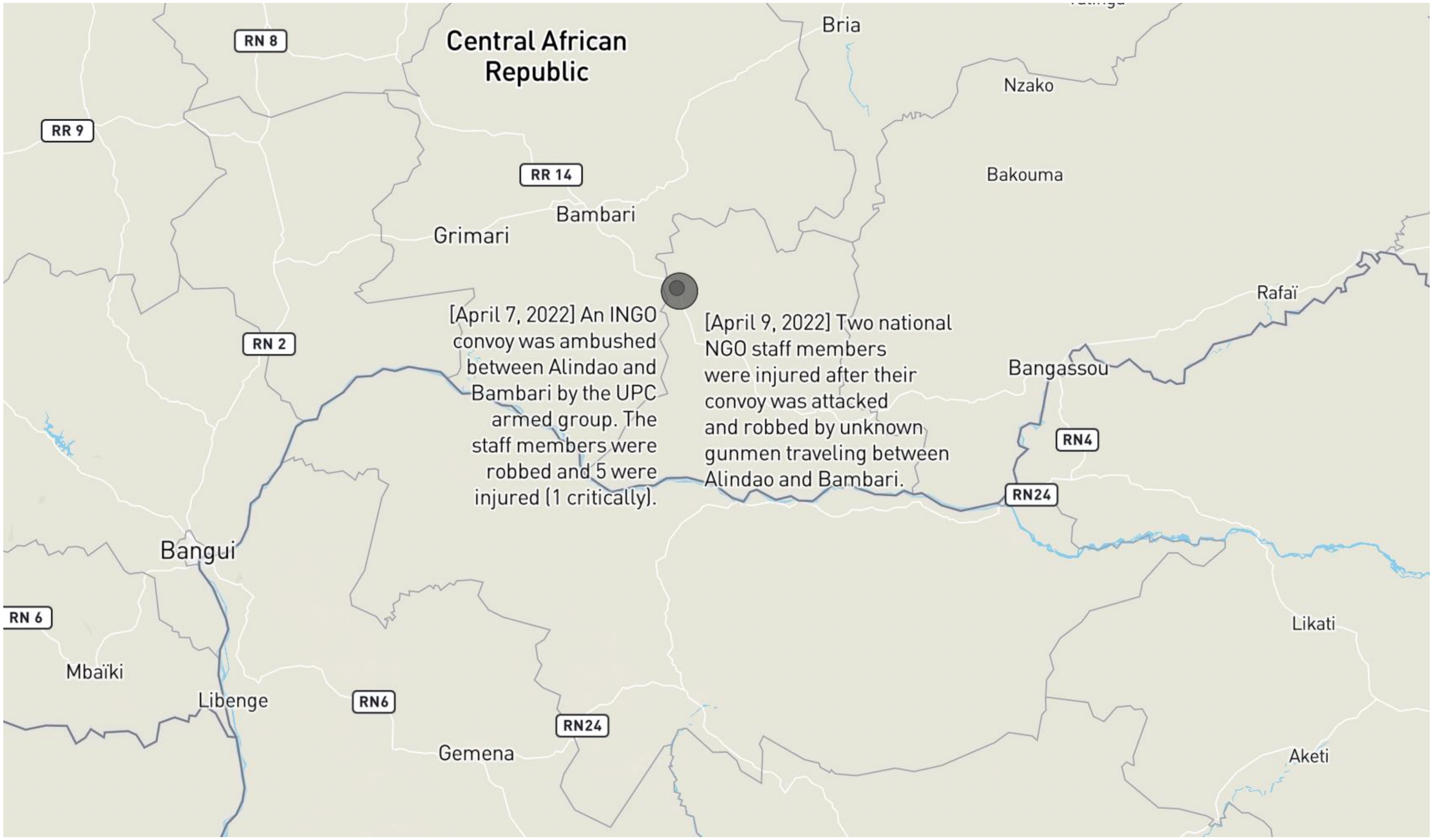

Figure 35. Diffusion dynamics of emerging cluster in south-central Central African Republic (N=5)

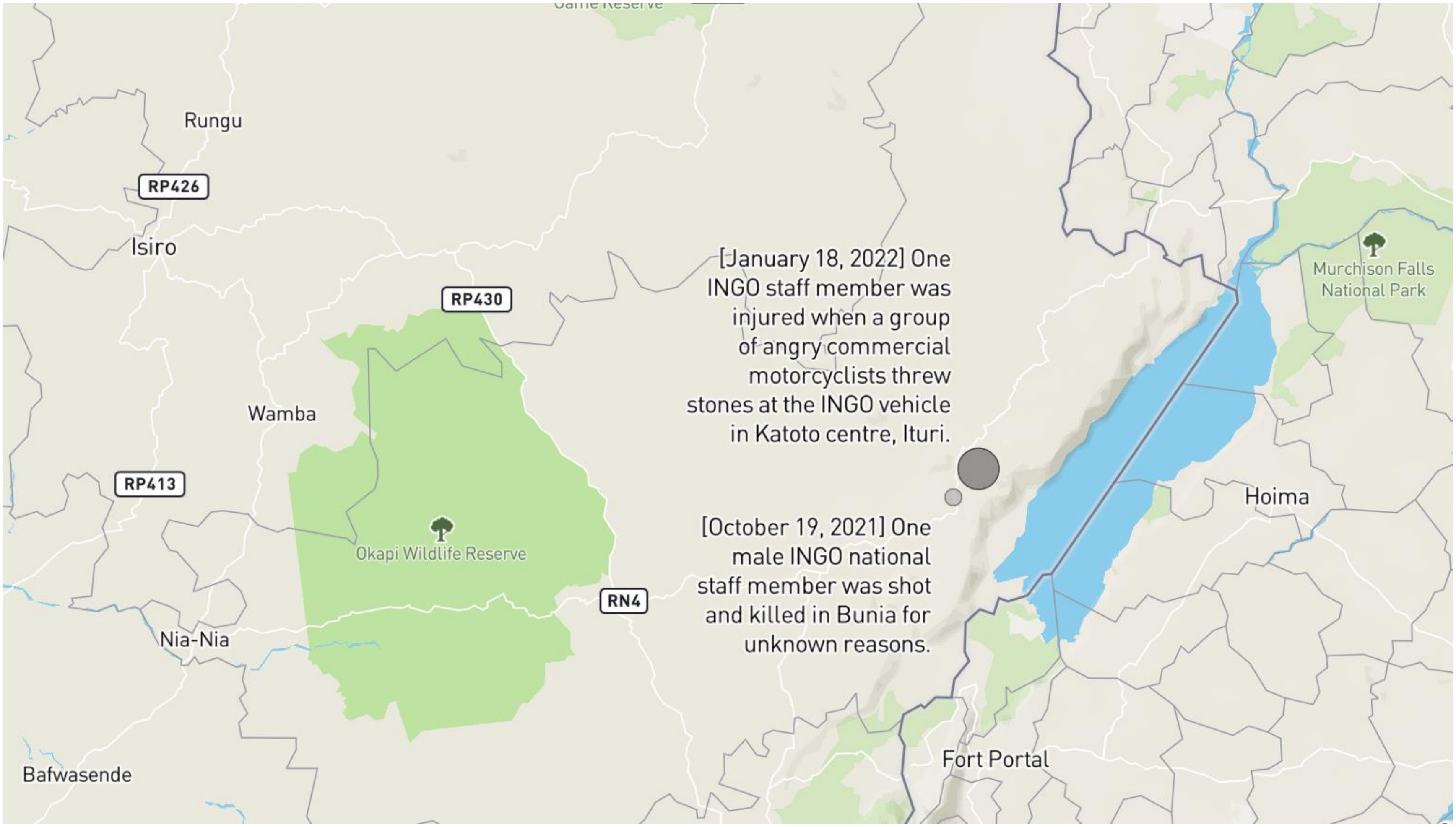

Figure 36. Diffusion dynamics of emerging cluster in Democratic Republic of the Congo (N=2)

Figure 37. Diffusion dynamics of emerging cluster in Myanmar (N=4)

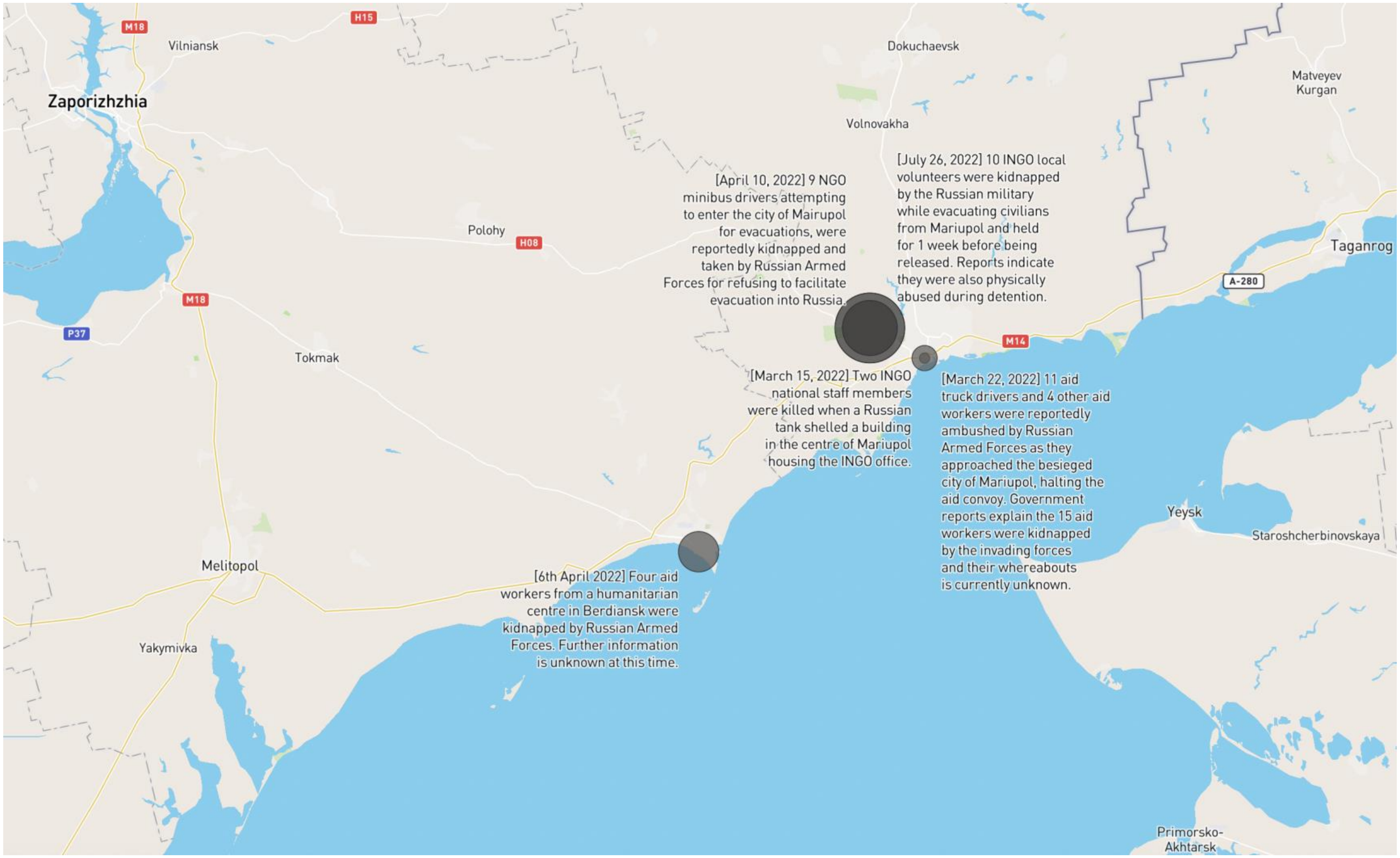

Figure 38. Diffusion dynamics of emerging cluster in Ukraine (N=4)

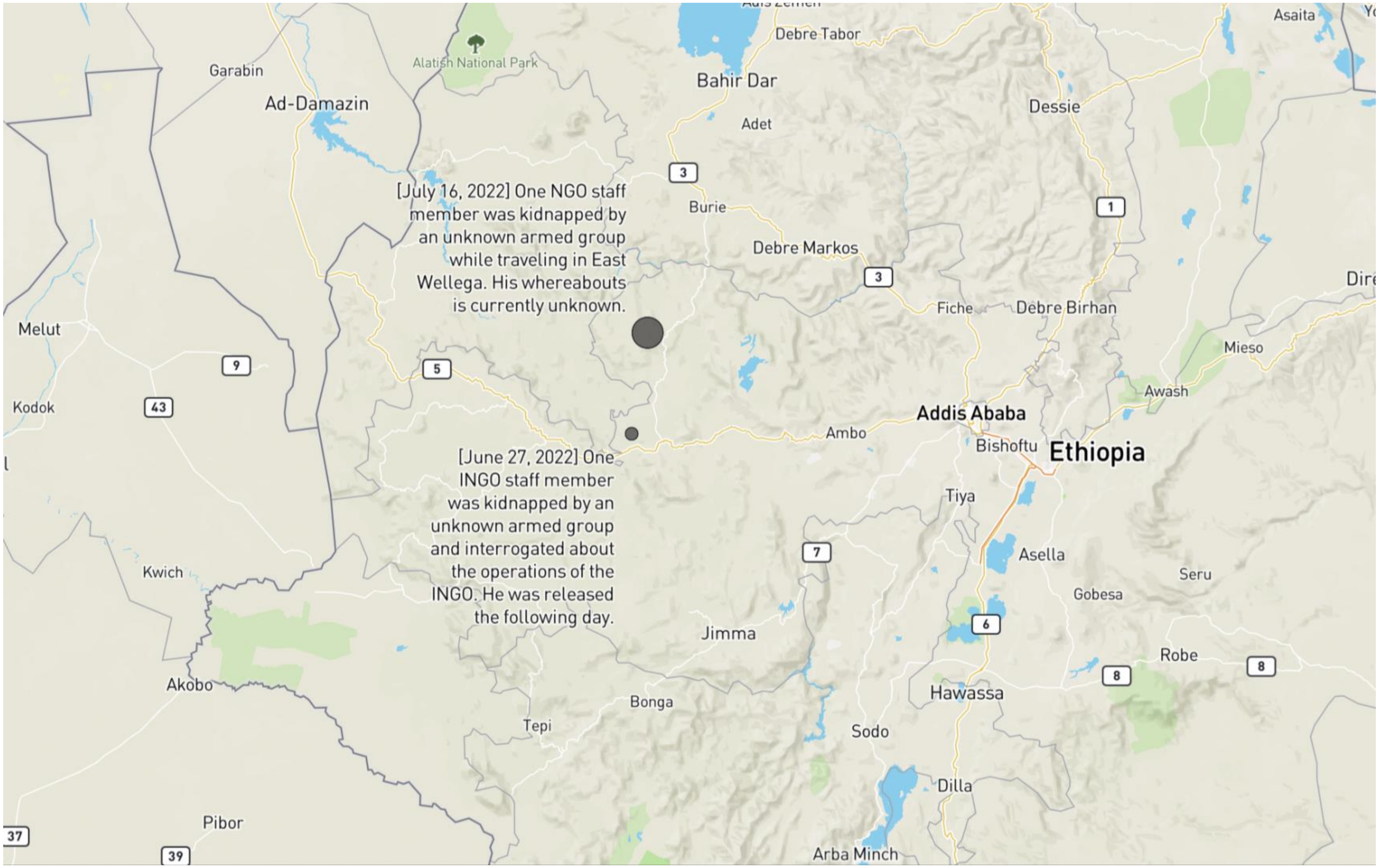

Figure 39. Diffusion dynamics of emerging cluster in Ethiopia (N=2)

Discussion

Attacks against humanitarian assistance are an insidious global problem facing aid workers and the communities where they provide live-saving assistance. In order to better understand the conditions under which humanitarian aid workers are most at risk, this study illustrated new clusters of attacks against humanitarian assistance that emerged between July 2021 and July 2022. The first major finding from the study was that the spatiotemporal distribution of attacks is not randomly distributed and that new clusters of attacks against humanitarian assistance are emerging: the present study found eleven emerging clusters of attacks against humanitarian assistance.

The second major finding from the study was that in certain areas, the spatial distribution of new clusters of attacks against humanitarian assistance spans multiple countries. A trans-national pattern of attacks was observed in Chad, Cameroon, and Nigeria where attacks by Boko Haram were documented in the same cluster, but in different countries (i.e., Cameroon and Nigeria). Given the fact that conflict dynamics (i.e., the regional Boko Haram insurgency) do not respect national borders, it makes sense that the perpetration of violence against humanitarian assistance would follow a similar range. On the other hand, findings from the study also demonstrate how distinct clusters of attacks can occur in the *same* country, but in different regions.

Third, findings from the present study underscore the importance of closely monitoring these clusters in terms of the local characteristics of attacks, repeated targets, and changes in the actors involved—in the interest of protection and efforts to strengthen humanitarian security. In order to ensure our ability to meet growing humanitarian needs around the world, it will be salient to track the emergence and diffusion dynamics of clusters across time and space in order to develop targeted, evidence-based efforts to prevent growth (i.e., to stop or reduce the impact of additional attacks in the detected clusters) and encourage behavior change among perpetrators. Interestingly, in countries where multiple new clusters were detected, attacks demonstrate different patterns of perpetration based on the cluster they belong to. In Central African Republic, for example, attacks in the cluster in the Northwest part of the country were disproportionately roadside IED detonations, compared to attacks in the cluster to the Southeast, where the observed attacks were individual ambush attacks. This finding highlights how, even in the same country, the dynamics of cluster emergence and perpetration vary.

Implications for Humanitarian Assistance Practice and Security

The present study, and the method it used, may have relevance to efforts to bolster the security of humanitarian assistance in affected areas. Indeed, there is great utility in being able to track—potentially in real time—changes in the risk dynamics at play in emerging clusters of attacks against humanitarian assistance. The spatiotemporal clustering approach used in this study is suitable because it captures critical nuance across and within countries, which puts into question the real-life applicability of cross-national analyses of attacks against humanitarian assistance. It is plausible that future researchers could adopt a non-parametric approach to understanding the complexity and dynamic nature of attacks within their local contexts. Web-based mapping tools like the one presented in this study may be used to monitor attacks and changes in the risk environment, indicating key target areas where policies and practices may be implemented to protect and support uninhibited humanitarian operations. Ultimately, stronger monitoring and evidence-based prevention strategies will be instrumental for protecting aid workers and reducing the impacts of attacks in communities affected by humanitarian emergencies, globally.

Limitations

Although this study sheds light on important areas of emerging risk for humanitarian assistance, it is not without limitations. First, data used in the present study are collected by human operators and, as a result, there is a chance that errors or issues with incomplete data affecting the reported findings. In addition, our specification of an attack cluster (i.e., detecting any pattern of two or more events within 100 kilometers in the same 365-day period) was selected based on our understanding of the problem, not by a defined rule or theory. If we had specified these definitions differently, attacks like those observed in Uganda, where humanitarian actors are attacked in public and individually, may have instead been classified by the clustering algorithm as isolated events and not detected as an emerging cluster. This raises an important consideration for future research to disentangle, related to the differences in targeting and attack outcomes in smaller clusters with fewer attacks, compared to larger clusters in settings with ongoing conflict.

Conclusion

Findings from the present study provide up-to-date evidence regarding emerging clusters of attacks against humanitarian assistance worldwide. We found that emerging spatial clusters were observed both in distinct regions within specific countries and across multiple countries, underscoring the importance of using a broad lens in the monitoring of the risks faced by humanitarian assistance. This study lays the theoretical and methodological groundwork for future efforts to track emerging sources and locations of risk to the humanitarian aid workers dedicated to helping the world's most vulnerable communities.

Chapter V. Discussion and Conclusion

In light of the growing global need for humanitarian assistance due to armed conflict, accelerating climate crises, and other emergencies researchers, policymakers, and practitioners are compelled to critically evaluate evidence on the dynamics of conflict and the specific contexts in which aid workers are at greatest risk. Findings from the three dissertation studies underscore the importance of armed conflict as a predictor of attacks against humanitarian assistance across the past 25 years. Results presented in chapter 2 indicate that the presence of an armed conflict—particularly over a governance dispute—and intensity of conflict are key factors in understanding the probability of attacks against humanitarian assistance. Although the first study found that on a global scale, contextual factors, such as sociopolitical and macroeconomic indicators, were relatively less important in terms of predicting attacks, findings from the second and third studies detected profound contextual differences in attack outcomes both in and between spatiotemporal hotspots.

Across the three studies, we also confirmed the utility of nonparametric modeling approaches in comparison to traditional regression modeling, which has both scientific and, critically, real world relevance. The second and third dissertation studies (chapters 3 and 4) laid the theoretical and methodological groundwork for future efforts to track historical and emerging areas of risk to humanitarian aid workers, which may support their protection and the prevention of casualties as a result of attacks against humanitarian assistance. Three dissertation studies also offer a number of practical uses. Results from chapter 2 highlight conflict-related and contextual risk factors that may be instrumental for planning and coordinating protection and prevention strategies. Likewise, the emerging cluster tool from chapter 4 may be a useful instrument for humanitarian practitioners for planning operations strategically around emerging risks (e.g., evidence-based security trainings and protocols, resource allocation to high-risk operations). The three studies also offer evidence which may be used to inform stronger, evidence-based policies to protect aid workers and prevent attacks. Our finding that national aid workers are among the most affected in historical hotspots and emerging clusters of attacks warrants the consideration of INGOs and NGOs, whose policies and practices could better cater to the needs and protection of local aid workers. Results from the three studies may also be used by policymakers to target protection and integrated prevention strategies across levels of the humanitarian system.

With respect to theoretical contributions of the three studies, the use of nonparametric statistical learning techniques allowed us to examine the relative importance of conflict-related and contextual predictors that have been theorized to predict attacks against humanitarian assistance on a global scale. Likewise, the applied theoretical lens of network theory and diffusion theory were found to be especially suitable to test Tobler's Law in the context of attacks against humanitarian assistance. We found that these theories were instrumental and supported a more nuanced understanding of the complex dynamics of attacks against humanitarian assistance across time and space. Findings from the three studies lend support to more research predicated on predictive nonparametric statistical learning models, network theory and related methods to disentangle complex, large-scale social and political phenomena.

The three studies also raise a number of important questions for future research to pursue. For example, common parametric modeling techniques in our study performed worse at the task of predicting attacks compared to nonparametric models, which has implications for future studies researching rare outcomes like attacks against humanitarian assistance and their etiology. The dissertation findings also raise numerous new questions for research: for example, what, if any, are the unique characteristics of spatiotemporally isolated attacks against humanitarian assistance, compared to attacks that occur in a defined spatiotemporal pattern or "cluster"? What are the most important factors that influence perpetrator behavior, and to what extent can their behavior be modified? Another important area of inquiry raised in chapter 3 relates to the diffusion dynamics in major hotspots with more than 100 attacks. In particular, to what extent are there clusters *within* the detected clusters of attacks; and how, if at all, do attacks in hotspot micro-clusters relate, spatially and etiologically, to conflict events?

Limitations

While the findings from the three studies are informative for researchers and useful for policymakers and practitioners, they are not without limitations. Most notably, attacks against humanitarian assistance often go unreported, and there may be disproportionate underreporting of attacks in some of the most remote and risky operational contexts. Similarly, general underreporting of attacks may bias the dataset where attacks in higher-risk areas are less likely to be reported. Thus, our data may be missing important information about some of the riskiest operational contexts for humanitarian assistance. Similarly, what data are available are incomplete for some cases, which may limit the extent of what we are able to learn from our analyses. Since spatial analysis is not a common use of the evidence in this dataset, the quality and completeness of the spatial data, in particular, could be improved. Finally, while these data cover a relatively long window of time, they cannot speak to dynamics before the late 1990s. These challenges are directly related to the nature of the dataset, which is not a formal surveillance system for attacks, but rather a publicly available database of events that are documented and verified by a team of humans.

In terms of the methods used in each of the studies, the predictive modeling approach in chapter 2 is limited in its ability to inform us about the underlying mechanisms that drive the observed prediction relationships. In chapters 3 and 4, a limitation of the network analysis approach is that such an algorithmic clustering approach is prone to spurious associations. Similarly, spatiotemporal analysis can produce different results from different methods, as the pre-defined algorithmic thresholds for radius, distance, and density are based on rules or knowledge on the topic of study. Our specification of an attack cluster (i.e., detecting any pattern of two or more events within 100 kilometers in the same 365-day period) was selected based on our understanding of the problem, not by a defined rule or theory. If we had specified these definitions differently, attacks like those observed in Uganda, where humanitarian actors are attacked in public and individually, may have instead been classified by the clustering algorithm as isolated events and not detected as an emerging cluster. Although a strength of this paper is that it focuses on the major hotspots of attacks against humanitarian assistance over the past twenty-five years, we cannot overlook the fact that—in some countries/regions—important attacks have taken place at a lower intensity or less frequent rate to be considered as a major cluster. We do not want to lose sight of these smaller, more diffused clusters, as their existence may indicate different dynamics of perpetration compared to larger or more dense clusters.

Conclusion

Three dissertation studies presented in chapters 2, 3, and 4 illustrate where, when, and under what operational conditions aid workers are targeted in attacks. Findings from chapter 2 emphasize how the presence, intensity, and type of armed conflict predict attacks against humanitarian assistance; and, in contrast to existing theory, indicate that the presence of one-sided violence against civilians *is* an important predictor of attacks against humanitarian assistance. In chapters 2 and 3, we found strong evidence that—rather than being randomly distributed across time and space—clear “hotspots” are observed where attacks are exceedingly common. Another important finding is that different types of conflicts appear to result in different patterns of attacks. Given the clear evidence of conflict-related dependencies and spatial patterns of attacks against aid workers over time, it is important for future research to disentangle how, if at all, attacks against humanitarian assistance coincide with armed conflict events in spatiotemporal hotspots, and how the movement of displaced persons and spatial exigencies of conflict and relate to cluster diffusion. Overall, findings presented in the three dissertation studies lay a critical foundation for monitoring attack risk across global contexts, and creating evidence-based protection and tools to reduce the preventable impact of attacks against humanitarian assistance.

References

Abdelrahman, S., & Haar, R. (2021). The Effect of Conflict on Healthcare Workers in Syria: Results of a Qualitative Survey. *medRxiv*.

Barreto, F. R., Teixeira, M. G., Costa, M. D. C. N., Carvalho, M. S., & Barreto, M. L. (2008). Spread pattern of the first dengue epidemic in the city of Salvador, Brazil. *BMC Public Health*, *8*(1), 1-20.

Benton, A., & Atshan, S. E. (2016). "Even war has rules": On medical neutrality and legitimate non-violence. *Culture, medicine, and psychiatry*, *40*(2), 151-158.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health*, *6*, 149.

Breiman, L. (2001a). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, *16*(3), 199-231.

Breiman, L. (2001b). Random forests. *Machine learning*, *45*(1), 5-32.

Briody, C., Rubenstein, L., Roberts, L., Penney, E., Keenan, W., & Horbar, J. (2018). Review of attacks on health care facilities in six conflicts of the past three decades. *Conflict and health*, *12*(1), 1-7.

Bouchet-Saulnier, F., & Whittall, J. (2018). An environment conducive to mistakes? Lessons learnt from the attack on the Médecins Sans Frontières hospital in Kunduz, Afghanistan. *International Review of the Red Cross*, *100*(907-909), 337-372.

Boutton, A., & Pascoe, H. (2018). Do Foreign Aid Projects Attract Transnational Terrorism?. *Peace Economics, Peace Science and Public Policy*, *24*(4).

Chainey, S., Tompson, L., & Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security journal*, *21*(1), 4-28.

Cliff, A. D., Ord, J. K., Haggett, P., & Versey, G. R. (1981). *Spatial diffusion: an historical geography of epidemics in an island community* (Vol. 14). CUP Archive.

Crombé, X., & Kuper, J. (2019). War breaks out: interpreting violence on healthcare in the early stage of the south Sudanese civil war. *Journal of Humanitarian Affairs*, *1*(2), 4-12.

Davies, S., Pettersson, T., & Öberg, M. (2022). Organized violence 1989–2021 and drone warfare. *Journal of Peace Research*, *59*(4), 593-6

Eisenberg, C., Halperin, D., Hargreaves, A., Hubbard, F., Mittleberger, J., Palmisano, J., & Stanbury, J. (1983). Health and human rights in El Salvador. *The New England journal of medicine*, *308*(17), 1028-1029.

Elamein, M., Bower, H., Valderrama, C., Zedan, D., Rihawi, H., Almilaji, K., ... & AbouZeid, A. (2017). Attacks against health care in Syria, 2015–16: results from a real-time reporting tool. *The Lancet*, *390*(10109), 2278-2286.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd* (Vol. 96, No. 34, pp. 226-231).

Fardousi, N., Douedari, Y., & Howard, N. (2019). Healthcare under siege: a qualitative study of health-worker responses to targeting and besiegement in Syria. *BMJ open*, *9*(9), e029651.

Fast, L. A. (2010). Mind the gap: Documenting and explaining violence against aid workers. *European Journal of International Relations*, *16*(3), 365-389.

Felson, M., & Eckert, M. A. (2018). *Crime and everyday life: A brief introduction*. Sage Publications.

Footer, K. H., Clouse, E., Rayes, D., Sahloul, Z., & Rubenstein, L. S. (2018). Qualitative accounts from Syrian health professionals regarding violations of the right to health, including the use of chemical weapons, in opposition-held Syria. *BMJ open*, *8*(8), e021096.

Garfield, R. M., Frieden, T., & Vermund, S. H. (1987). Health-related outcomes of war in Nicaragua. *American Journal of Public Health*, *77*(5), 615-618.

Geiger, J., Eisenberg, C., Gloyd, S., Quiroga, J., Schlenker, T., Scrimshaw, N., & Devin, J. (1989). A new medical mission to El Salvador. *New England Journal of Medicine*, *321*(16), 1136-1140.

Gellhorn A. (1983). Medical mission report on El Salvador. *The New England journal of medicine*, *308*(17), 1043–1044. https://doi.org/10.1056/NEJM198304283081736

Geiger, J., Eisenberg, C., Gloyd, S., Quiroga, J., Schlenker, T., Scrimshaw, N., & Devin, J. (1989). A new medical mission to El Salvador. *The New England journal of medicine*, *321*(16), 1136–1140. https://doi.org/10.1056/NEJM198910193211629

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed conflict 1946-2001: A new dataset. *Journal of peace research*, *39*(5), 615-637.

Hague Regulations (1907). Annex to the Convention: Regulations Respecting the Laws and Customs of War on Land.

Haar, R. J., Read, R., Fast, L., Blanchet, K., Rinaldi, S., Taithe, B., ... & Rubenstein, L. S. (2021). Violence against healthcare in conflict: a systematic review of the literature and agenda for future research. *Conflict and health*, *15*(1), 1-18.

Hoelscher, K., Miklian, J., & Nygård, H. M. (2017). Conflict, peacekeeping, and humanitarian security: Understanding violent attacks against aid workers. *International Peacekeeping*, *24*(4), 538-565.

Humanitarian Outcomes (2022). Aid Worker Security Database.

International Committee of the Red Cross (1949, August 12) The Fourth Geneva Conventions. Geneva, pp.153-221.

Kuo, F. Y., Wen, T. H., & Sabel, C. E. (2018). Characterizing diffusion dynamics of disease clustering: a Modified Space–Time DBSCAN (MST-DBSCAN) algorithm. *Annals of the American Association of Geographers*, *108*(4), 1168-1186.

Lian, M., Warner, R. D., Alexander, J. L., & Dixon, K. R. (2007). Using geographic information systems and spatial and space-time scan statistics for a population-based risk analysis of the 2002 equine West Nile epidemic in six contiguous regions of Texas. *International Journal of Health Geographics*, *6*(1), 1-10.

Madhiwalla, N., & Roy, N. (2009). Bombing medical facilities: a violation of international humanitarian law. *Indian J Med Ethics*, *6*(2), 64-65.

Mahieu, A., Blanchet, K., Haar, R.J. Fast, L., Kostandova, N., Rubenstein, L.S. (2020) Conceptual Issues and methodological approaches to evaluating the wider and longer-term impacts of attacks on healthcare in conflict. Manchester, UK.

Marshall, M. G., & Gurr, T. R. (2021). Polity5: Political Regime Characteristics and Transitions, 1800-2018 Dataset. *Center for Systemic Peace, Vienna*.

Meade, M. S., & Emch, M. (2010). *Medical geography*. Guilford Press.

Michlig, G. J., Lafta, R., Al-Nuaimi, M., & Burnham, G. (2019). Providing healthcare under ISIS: A qualitative analysis of healthcare worker experiences in Mosul, Iraq between June 2014 and June 2017. *Global public health*, *14*(10), 1414-1427.

Midlarsky, M. I., Crenshaw, M., & Yoshida, F. (1980). Why violence spreads: The contagion of international terrorism. *International Studies Quarterly*, *24*(2), 262-298.

Narang, N., & Stanton, J. A. (2017). A strategic logic of attacking aid workers: Evidence from violence in Afghanistan. *International Studies Quarterly*, *61*(1), 38-51.

Nickerson, J. W. (2015). Ensuring the security of health care in conflict settings: an urgent global health concern. *Cmaj*, *187*(11), E347-E348.

Poldrack, R. A., Huckins, G., & Varoquaux, G. (2020). Establishment of best practices for evidence for prediction: a review. *JAMA psychiatry*, *77*(5), 534-540.

Rubenstein, L. S., & Bittle, M. D. (2010). Responsibility for protection of medical workers and facilities in armed conflict. *The Lancet*, *375*(9711), 329-340.

Rubenstein, L. S. (2013). A way forward in protecting health services in conflict: moving beyond the humanitarian paradigm. *International Review of the Red Cross*, *95*(890), 331-340.

Safeguarding Health Care Coalition (2020). Health care workers at risk: Violence against health care. https://www.safeguardinghealth.org/sites/shcc/files/SHCC2020final.pdf

Shmueli, G. (2010). To explain or to predict?. *Statistical science*, *25*(3), 289-310.

Singh, N. S., Redman, B., Broussard, G., DeCamp, M., Rayes, D., Ho, L. S., ... & Rubenstein, L. (2022). ‘We will never give up’: a qualitative study of ethical challenges Syrian health workers face in situations of extreme violence. *Disasters*, *46*(2), 301-328.

Tammi, I. M. (2021). Politicians, pathogens, and other threats to aid workers: a material semiotic analysis of violence against health care in the Syrian conflict. *Critical Studies on Security*, *9*(3), 183-195.

Terry, F. (2013). Violence against health care: insights from Afghanistan, Somalia, and the Democratic Republic of the Congo. *International Review of the Red Cross*, *95*(889), 23-39.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, *46*(sup1), 234-240.

Ward, M. D., Greenhill, B. D., & Bakke, K. M. (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of peace research*, *47*(4), 363-375.

Wortley, R., & Mazerolle, L. (2013). Environmental criminology and crime analysis: situating the theory, analytic approach and application. In *Environmental criminology and crime analysis* (pp. 23-40). Willan.

Yousuf, M. H., Jabbar, A., Ullah, I., Tahir, M. J., & Yousaf, Z. (2021). Violence against health care system in areas of conflict: Unveiling the crisis globally. *Ethics, Medicine, and Public Health*, *19*, 100730.

Zhao, L. (2021). Event prediction in the big data era: A systematic survey. *ACM Computing Surveys (CSUR)*, *54*(5), 1-37.

Appendix A: ST-DBSCAN Algorithm

```
### x = x-axis coordinate of point data
### y = y-axis coordinate of point data
### time = time coordinate of point data
### eps = distance maximum for longitude and latitude
### eps2 = distance maximum for date
### minpts = number of points to consider a cluster

stdbscan <- function (x,
                      y,
                      time,
                      eps,
                      eps2,
                      minpts) {

  distdata <- cbind.data.frame(x, y)
  time <- time
  value <- value

  n <- nrow(distdata)

  classn <- cv <- integer(n)
  isseed <- logical(n)
  cn <- integer(1)

  for (i in 1:n) {
    unclass <- (1:n)[cv < 1]

## making distance

      a <- data.frame(x = distdata[i, 1], y = distdata[i, 2])
      fordist <- cbind.data.frame(a, distdata)
      idist <- abs(sqrt((fordist[,1] - fordist[,3])^2 + (fordist[, 2] -
                                                                fordist[,
      4])^2))
      fortime <- cbind.data.frame(time[i], time)
      itimedist <- abs(fortime[, 1] - fortime[, 2])

      if (cv[i] == 0) {

            reachables <- intersect(unclass[idist[unclass] <= eps],
            unclass[itimedist[unclass] <= eps2])
            if (length(reachables) + classn[i] < minpts)
                cv[i] <- (-1)
            else {
                cn <- cn + 1
                cv[i] <- cn
                isseed[i] <- TRUE
                reachables <- setdiff(reachables, i)
                unclass <- setdiff(unclass, i)
```

```
                classn[reachables] <- classn[reachables] + 1
                while (length(reachables)) {
                    cv[reachables] <- cn
                    ap <- reachables
                    reachables <- integer()

                    for (i2 in seq(along = ap)) {
                        j <- ap[i2]

## make distance again when cluster is expanding

      b <- data.frame(x = distdata[j, 1], y = distdata[j, 2])
      jfordist <- cbind.data.frame(b, distdata)
      jdist <- sqrt((jfordist[,1] - jfordist[,3])^2 +
                                             (jfordist[, 2] - jfordist[,
4])^2)
      jfortime <- cbind.data.frame(time[j], time)
      jtimedist <- abs(jfortime[, 1] - jfortime[, 2])
      jreachables <- intersect(unclass[jdist[unclass] <= eps],

unclass[jtimedist[unclass] <= eps2])

      if (length(jreachables) + classn[j] >= minpts) {
           isseed[j] <- TRUE
            cv[jreachables[cv[jreachables] < 0]] <- cn
            reachables <- union(reachables, jreachables[cv[jreachables]
== 0])
      }
      classn[jreachables] <- classn[jreachables] + 1
      unclass <- setdiff(unclass, j)
               }
            }
         }
      }
      if (!length(unclass))
            break
    }

    if (any(cv == (-1))) {
        cv[cv == (-1)] <- 0
    }
    result <- list(cluster = cv, eps = eps,
                   eps2 = eps2,
                   minpts = minpts, density = classn)
    rm(classn)

    class(result) <- "stdbscan"
    return(result)
}
```