Modelling the Spatial Dynamics of Non-State Terrorism: World Study, 2002-2013

André Python

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of
PhD
at the
University of St Andrews

July 25, 2017

1. Candidate's declarations:

I, André Python, hereby certify that this thesis, which is approximately 50,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in August, 2013 and as a candidate for the degree of PhD in April 2014; the higher study for which this is a record was carried out in the University of St Andrews between 2013 and 2017.

I, André Python, received assistance in the writing of this thesis in respect of spelling and syntax, which was provided by Mr James Hansen.

Date: July 24, 2017 signature of candidate …….

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: July 24, 2017 signature of supervisor…….

3. Permission for publication:

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

a) Embargo on all or part of print copy for a period of 2 years on the following ground(s):
  - Publication would preclude future publication

ELECTRONIC COPY

a) Embargo on all or part of electronic copy for a period of 2 years on the following ground(s):
  - Publication would preclude future publication

ABSTRACT AND TITLE EMBARGOES

An embargo on the full text copy of your thesis in the electronic and printed formats will be granted automatically in the first instance. This embargo includes the abstract and title except that the title will be used in the graduation booklet.

If you have selected an embargo option indicate below if you wish to allow the thesis abstract and/or title to be published. If you do not complete the section below the title and abstract will remain embargoed along with the text of the thesis.

a)	I agree to the title and abstract being published	YES
b)	I require an embargo on abstract	NO
c)	I require an embargo on title	NO

Date: July 24, 2017 signature of candidate ….. signature of supervisor …….

Please note initial embargos can be requested for a maximum of five years. An embargo on a thesis submitted to the Faculty of Science or Medicine is rarely granted for more than two years in the first instance, without good justification. The Library will not lift an embargo before confirming with the student and supervisor that they do not intend to request a continuation. In the absence of an agreed response from both student and supervisor, the Head of School will be consulted. Please note that the total period of an embargo, including any continuation, is not expected to exceed ten years. Where part of a thesis is to be embargoed, please specify the part and the reason.

"*Courage is the first of human qualities because it is the quality which guarantees the others*"

- Aristotle, 384-322 BCE -

Acknowledgements

I would like to acknowledge those who have supported me during my PhD study. I could not even have dreamed of starting my PhD at the University of St Andrews without the trust and generous support of Martin Beniston, Gregor Kozlowski, Giuseppe Imparato, Mathias Thoenig, Dominic Rohner and Mathieu Couttenier. I would like to thank my supervisors, Janine Illian and Javier Argomaniz for their immense generosity and incredible guidance. Their tireless dedication to work along with their genuine care for my research pushed me ceaselessly to reach beyond my perceived boundaries. Their ability to bridge eclectic disciplines such as statistics and terrorism studies have been a great source of inspiration throughout my interdisciplinary studies. I am also very grateful to Richard English, Len Thomas and Gillian Brunton, who encouraged me to carry out my research. Through their strong support and wise advices, they convinced me that is not only desirable to believe in the success of innovative ideas, but above all it is an essential requirement to carry out research at the frontiers of knowledge. I would also like to acknowledge the Centre for Research into Ecological and Environmental Modelling (CREEM) for granting scholarship.

I am also very grateful to people who shared their valuable expertise and dedicated time to my progress: Marta Blangiardo, Peter Lehr, Charlotte Jones-Todd, Håvard Rue, Finn Lindgren, Anna Freni Sterrantino, Nils Weidmann, Guido Sanguinetti, George Huffman, Elias Krainski, and Richard Booth. I also would like to thank my friends and colleagues for their kind companionship in London: Hakim Dehbi, James Hansen, Areti Boulieri, Oscar Rodriguez de Rivera, Talip Alkhayer; St Andrews: Manlio Fossati, Alberto Mastellari, Enrico Veneziani, Alessandro Corso, Aliya Tskhay, Elena Zhirukhina, Daria Vorobyeva, Selbi Hanova, Marat Ilyasov, Abel Farkas, Francois Sarah, Alexander Knight; and Switzerland: Xavier Ruffieux, Maurizio Cortelessa, Virgilio Conti, Rossana Fondi-Reboh, Johann Marion, Azmir Huseinovic, Pascal Martelli, Yannick Mettraux and their family.

I would like to address a special thank my mother for her unconditional love, courage and humility. As the eldest of five siblings, she decided to work in Switzerland in order to help her family in Slovenia, ex-Yugoslavia. Through her courage and determination, she raised a family of four children while providing her family in Slovenia with basic needs. Since my childhood, she has conveyed her passion for the study of nature and stimulated my curiosity about the beauty of sciences. She remains my greatest source of inspiration. I also would like to sincerely thank my partner, Angélique. She has bravely endured the hardship of distant relationships and difficulties we faced together. Through her resilience and desire to cultivate enduring love, we have stayed united. I could not conclude without acknowledging the strong support and love I received from my siblings, Anne, Martin and Vladimir, and my other family members, Dounya, Laura, Marie-Claire, Nuša, Pavla, Tina, Saška, Adrien, Arnold, Francis, Mišo, and Marian. Thank you all for contributing to this exciting and rewarding journey!

Abstract

To this day, terrorism perpetrated by non-state actors persists as a worldwide threat, as exemplified by the recent lethal attacks in Paris, London, Brussels, and the ongoing massacres perpetrated by the Islamic State in Iraq, Syria and neighbouring countries. In response, states deploy various counterterrorism policies, the costs of which could be reduced through more efficient preventive measures. The literature has not applied statistical models able to account for complex spatio-temporal dependencies, despite their potential for explaining and preventing non-state terrorism at the sub-national level. In an effort to address this shortcoming, this thesis employs Bayesian hierarchical models, where the spatial random field is represented by a stochastic partial differential equation. The results show that lethal terrorist attacks perpetrated by non-state actors tend to be concentrated in areas located within failed states from which they may diffuse locally, towards neighbouring areas. At the sub-national level, the propensity of attacks to be lethal and the frequency of lethal attacks appear to be driven by antagonistic mechanisms. Attacks are more likely to be lethal far away from large cities, at higher altitudes, in less economically developed areas, and in locations with higher ethnic diversity. In contrast, the frequency of lethal attacks tends to be higher in more economically developed areas, close to large cities, and within democratic countries.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and Research Questions

To this day, terrorism still persists as a worldwide threat, exemplified by the recent deadly attacks in Nice, Istanbul, Brussels, and Paris, as well as the ongoing massacres perpetrated in Iraq and Syria by ISIS, also-known by its Arabic acronym Daesh (داعش). Terrorist attacks may generate economic, environmental, and social damages, including the loss of human lives. In response to this threat, states deploy a combination of various countert-errorism policies, which include the use of criminal justice, military power, intelligence, psychological operations, and preventives measures (Crelinsten, 2009, p. 45). Following the tragic terrorist attacks on September 11, 2001 in New York (9/11), states have tended to increase their spending to counter terrorism. From 2001 to 2008, the expenditure of world-wide homeland security increased by US$ 70 billion (NATO, 2008). In the U.S. alone, the 2013 federal budget devoted to combatting terrorism reached around US$ 17.2 billion (The Washington Post, 2013).

In order to increase the efficacy of preventive measures and therefore reduce the cost of counterterrorism, policy-makers would benefit from accurate models that capture the complexity of the dynamics of terrorism. Despite existing success in modelling crime and conflict in space and time (Lewis et al., 2012; Mohler, 2013; Rodrigues et al., 2010; Zammit-Mangion et al., 2013, 2012), recent spatio-temporal modelling approaches have not been systematically applied in the field of terrorism. Indeed, theoretical and empirical investigations of terrorism have often neglected its spatial component or have been mainly conducted on a macro-scale (country or wider level of analysis), which failed to capture fine-scale processes that generate patterns of terrorism at sub-national levels. Moreover, local-level studies have been restricted to case studies within limited study areas, and hence, drastically reduced their generalisability. As a result, local drivers of terrorism have not been systematically captured and scholars have failed to assess theories at local scale.

In an effort to address these shortcomings, this thesis aims to provide a systematic approach to modelling in space and time and on a local scale: (i) the *probability of lethal attacks*, estimated as the probability that terrorist attacks are lethal (rather than *non-lethal*); (ii) the *number of lethal attacks*, estimated as the *expected number* of lethal terrorist attacks. For this purpose, this work aims at answering the following questions:

1. Why are some areas more prone to encounter *lethal* rather than *non-lethal* terrorist attacks?
   (a) What are the driving forces behind an increase in the probability or in the number of lethal terrorist attacks?
   (b) Where and when do "abnormally" high levels (also called *hotspots*) of the probability or the number of lethal terrorist attacks occur?
2. Why are some areas more prone to encounter a *spread* of high probability or high number of lethal terrorist attacks?
   (a) What are the driving forces behind the spread (also called *diffusion*)?
   (b) Where and when does the spread occur in space and/or in time?

This thesis identifies covariates and their local effects on the probability and number of lethal terrorist attacks that occurred worldwide from 2002 to 2013. The complex spatial dependencies observed in the data are accurately modelled through a Bayesian approach, which provides insight into the spatial dynamics of the probability and number of lethal attacks. Moreover, hotspots (also called *clusters*) and diffusion processes of lethal terrorism are localised in space and time and analysed.

The results of this research offer a rigorous framework to assess, at sub-national level, theories that aim to explain hotspots and diffusion processes of lethal terrorism. Furthermore, the results provide complementary decision-support tools that could be used to enhance the efficacy of preventative counterterrorism policies. Since fine-scale spatial dynamics of terrorism are modelled globally, the scope of this research is not limited to any specific area and/or scale and could therefore benefit policy makers needing a global, national, or a local perspective on terrorism perpetrated by non-state actors (*non-state* terrorism).

This thesis highlights the intrinsic ambiguity of the concept of terrorism and the main issues of defining terrorism, which have been identified by the literature (Section 2.1.1). A working definition is suggested consistently with the definition used by the database that has been selected to model the probability and number of lethal terrorist attacks (Section 5.1.1). In line with most scholars and practitioners, state terrorism is excluded from this research (Clauset et al., 2007; Drakos and Gofas, 2006b; Enders et al., 2011). Geolocalised data on state terrorist attacks with a worldwide coverage are not publicly available, which makes its spatial analysis virtually impossible. This study encompasses both *domestic* and *transnational* terrorism without distinction, in line with Kegley (1990) and Sánchez-Cuenca and De la Calle (2009).

1.2 A Bayesian Stochastic Modelling Approach

1.2.1 Stochastic Models

Mathematical models have previously been used to understand the behaviour of political conflict, including terrorism. In the early 1990s, Lichbach (1992) identified more than 200 scholarly works which used mathematical models to study terrorism and related phenomena such as guerilla wars and insurrections. Mathematical models can be further distinguished into *deterministic* and *stochastic* models (Cressie and Wikle, 2011, p. 59). A well-known set of deterministic models are Newton's three laws of motion and the law of universal gravitation described in his work *Philosophiae Naturalis Principia Mathematica* first published in 1687 (Nauenberg, 2001, p. 194). In this framework, the position of the earth around the sun (Figure 1.1, *left*) is entirely determined and totally predictable[1]

Despite their immense scientific contribution, deterministic models are rarely suitable in the study of social phenomena and terrorism is no exception. The causes of terrorism are multidimensional and operate at the individual, group, sub-national, national, and transnational levels (Brynjar and Skjølberg, 2000; Richardson, 2006, p. 60). More particularly, ideology and belief are crucial factors of terrorism (Crenshaw, 1983, p. 29; Wilkinson, 1990, p. 141), which may considerably vary among individuals and groups of terrorists, and

[1] Note that the use of deterministic models does not necessarily mean that the processes under study are predictable. This is the case for chaotic processes for example (Kellert, 1994, p. 65).

changing over time. Therefore, it would be difficult, if not impossible, to provide a deterministic prediction of the characteristics of each attack at any desired level of accuracy in space (degree, minute, second, etc. of latitude and longitude) and/or time (day, hour, minute, etc.), since it would imply that all processes involved in each terrorist event are known and entirely determined.

While deterministic models predict a unique outcome from a given set of circumstances, stochastic models predict a *set* of possible outcomes through a probabilistic statement (Taylor and Karlin, 2014, p. 2). A famous example of stochastic model is the "electron cloud model" proposed by Schrödinger in 1926. In this framework, the position of the electron surrounding the nucleus of a ground state hydrogen atom (Figure 1.1, *right*) is not known before it is measured. Therefore, one can only attribute a *probability* to find the electron in a so-called "electron cloud" surrounding the nucleus.

Fig. 1.1 Schematic view of a *stochastic* and a *deterministic* model (not to scale). *Left*: *deterministic* model of the orbit of the earth (*gray disk*) around the sun (*black disk*). At any time of the year, the position of the earth is known with absolute certainty. *Right*: *stochastic* model of the position of the electron, represented by an "electron cloud" (*shaded area*) surrounding the nucleus (*black disk*) of a ground state hydrogen atom, with higher probability (*dark colour*) close to the nucleus and lower probability (*bright colour*) at farther distances.

Analogously, the characteristics of terrorist events — geolocalisation, event time, or number of deaths — are not known before they occur and are observed. Nevertheless, one should not equate terrorism with rolling the dice; terrorist attacks often exhibit patterns generated by underlying processes that can be captured. For example, the number of terrorist events per unit of area is not equally distributed across all cities in the world; rather, one observes hotspots, which can be partially explained through the identification of specific factors. In short, terrorism is neither a phenomenon that occurs purely by chance nor the exclusive result of a well-known underlying process; it is rather a combination of both. As a result, stochastic models have proved to be very useful in modelling social phenomena by considering *uncertainty* in the model whilst detecting trends and patterns (Zammit Mangion, 2011).

1.2.2 The Bayesian Framework

Stochastic models for spatial or spatio-temporal data are essentially led by two schools of thought: (i) *frequentist*, which also refers to the *classical* approach and (ii) *Bayesian*, which have been more recently intensively investigated (Cressie and Wikle, 2011, p. 14). In a general framework, for a set of observations $\mathbf{y} = \{y_1, \ldots, y_n\}$, with $i = 1, \ldots, n$, distributed according to a probability model $P(\mathbf{y}; \boldsymbol{\theta})$, a classical approach estimates the set of parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d)$, of dimension d, where each element represents an unknown fixed number[2]. Then, the parameters can be estimated through different methods such as the maximum likelihood estimator (MLE) for example, which aims to find the most likely parameter value given the data (Sørbye, 2014)[3].

The Bayesian approach initiated from the work conducted by the Reverend Thomas Bayes (1701-61). It provides a conceptual framework through which *prior* beliefs are updated by learning from *data*. *Prior knowledge* is formalised as a probabilistic statement on the parameters to be estimated before observing the data (Cressie and Wikle, 2011, p. 32). Diverse sources of information can be used as prior knowledge: (i) past experiments: information from previous experiments is used to guess the possible outcome of a new experiment; (ii) expert: prior knowledge is based on the opinion of one or several experts, which refers to *elicited* priors; (iii) absence of knowledge: in the absence of any source of knowledge, prior knowledge can be expressed through *noninformative* priors (Carlin and Louis, 2008, p. 27-41).

Unlike the frequentist approach, parameters are not considered *fixed* numbers; rather they represent *random* quantities with unknown distributions, from which the mean and standard deviation might be estimated for example (Sørbye, 2014). *Posterior knowledge* is obtained by updating prior knowledge with the observed data. More formally, the Bayesian approach aims to infer knowledge on a parameter set $\boldsymbol{\theta}$ (King et al., 2014). For a set of

[2] By convention, I use boldface lower case letters for vectors, boldface capital letters for matrices, and capital letters for random variables. Particular realisations of a random variable Y are written in corresponding lower case letters $\{y_1, \ldots, y_n\}$, with $i = 1, \ldots, n$. [3] Note that the development of MLE by R. A. Fisher was a considerable contribution to the field of statistics in the last century. For more information on the development of MLE, see for example Aldrich (1997).

parameters $\boldsymbol{\theta}$ and continuous data $\mathbf{y}$, Bayes' theorem states that (Congdon, 2007, p. 2):

$$P(\boldsymbol{\theta}|\mathbf{y}) = \frac{P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{y})}, \tag{1.1}$$

with prior knowledge about the parameters summarised by the prior distribution $P(\boldsymbol{\theta})$. The information contained in the data on the parameters is called *likelihood* $P(\mathbf{y}|\boldsymbol{\theta})$, and the updated knowledge is contained in the *posterior density* $P(\boldsymbol{\theta}|\mathbf{y})$. The *marginal likelihood* $P(\mathbf{y})$ is an integral over all values of $\boldsymbol{\theta}$ of the product $P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})$[4]. Hence, Bayes' theorem can be expressed as:

$$P(\boldsymbol{\theta}|\mathbf{y}) \propto P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta}). \tag{1.2}$$

Equation 1.2 represents the core of Bayesian inference, which states that the posterior density is proportional ($\propto$) to the product of the likelihood and the prior distribution (King et al., 2014). For example, assume a simple linear regression model with Gaussian observations $\mathbf{y} = (y_1, \ldots, y_n)$, where the expected value $\mathbb{E}(y_i) = \alpha + \beta x_i$, with the variance $\mathbb{V}(y_i) = \tau^{-1}$, with $i = 1, \ldots, n$. In a Bayesian approach, priors with a probability distribution will be assigned to the parameters α, β and τ and the (posterior) distribution of the parameters will be estimated instead of a point estimate.

The main advantage of the Bayesian over the frequentist approach is that it provides more intuitive and meaningful inference. Probability statements are made about the parameters of interest rather than using indirect measures such as *p-values* as the probability under hypothesis of data at least as extreme as that actually observed. In Bayesian analysis, *credible* intervals (CI) — the analogue of confidence interval in classical statistics — are directly related to the parameters, which are easily interpretable in term of probability. A 95% CI contains the true parameter value with approximatively 95% certainty. In contrast, a frequentist 95% confidence interval should not be understood as an interval that contains the true value with a 95% probability. Rather, in the long run of a repeated experience, 5% of the credible intervals will not contain the true parameter (Stevens, 2009).

Thus, the accuracy of classical estimation of parameters can be improved by combining prior knowledge with information from the observed data (Congdon, 2007, p. 1). However, fundamental issues remain in the process of quantifying prior knowledge, which is often an

[4] The marginal likelihood $P(\mathbf{y})$ is also called normalising constant as it ensures that $P(\boldsymbol{\theta}|\mathbf{y})$ is a proper density function.

arduous endeavour. Determining a suitable form of the prior has been under intense debate (Simpson et al., 2016), whose review would go far beyond the scope of this study.

1.2.3 Hierarchical Modelling

Most natural and social phenomena operate at different scales and terrorism is no exception. Potential factors that lead to terrorist attacks interact at different levels (Brynjar and Skjølberg, 2000; Richardson, 2006, p. 60): (i) individual (e.g. psychology, age, sex, education, family context) (Crenshaw, 1983, p. 29; Crenshaw, 2000); (ii) terrorist group (e.g. ideology, objectives, structure, size, funding, type of management) (Wilkinson, 1990, p. 141; Hoffman, 2001); (iii) local/regional/national/transnational (e.g. economy, demography, geography, political system, legal framework, environment, transnational collaboration/funding, media) (Abadie, 2006; Piazza, 2006). As a result, terrorism operates on a wide range of spatial scales (e.g. street, city, administrative region, country, international) and temporal scales (from split-second factors leading to bomb detonation to long-term drivers such as social cleavage which increase the risk of terrorism over years or decades).

Hierarchical model (HM) frameworks are suitable to model multi-dimensional phenomena and can accommodate data, model, and parameter uncertainty, as well as complex spatial, temporal, and spatiotemporal dependence. In HM, it is common to define three stages: (i) stage 1: the *data* model, which defines the conditional probability distribution of $\mathbf{y}$ given a "true" unknown process $\boldsymbol{\xi}$ (also-called *latent field*) and a set of parameters $\boldsymbol{\theta}$; (ii) stage 2: the *process* model, which quantifies uncertainty in $\boldsymbol{\xi}$ through a probability distribution; (iii) stage 3: the *parameters* model, where priors are set for the *hyperparameters*[5] of $\boldsymbol{\theta}$:

$$\begin{aligned} \text{stage 1:} &\quad P(\mathbf{y}|\boldsymbol{\xi},\boldsymbol{\theta}); \\ \text{stage 2:} &\quad p(\boldsymbol{\xi}|\boldsymbol{\theta}); \\ \text{stage 3:} &\quad p(\boldsymbol{\theta}), \end{aligned} \tag{1.3}$$

where the posterior probability distribution $P(\boldsymbol{\xi},\boldsymbol{\theta}|\mathbf{y})$ is proportional to the distributions in the three stages. Hence, through its ability to take into account non-linear interactions within the process (Cressie and Wikle, 2011, p. 403), HM provides a genuine framework to modelling phenomena that exhibit complex spatial and temporal interactions.

[5] In Bayesian statistics, a hyperparameter refers to a parameter of a prior distribution. Fixed values or a probability distribution can be attributed to a hyperparameter, the latter refers to *hyperprior* (Bernardo and Smith, 2010) (as cited in Wikipedia, 2015).

1.3 Modelling Lethal Terrorism in Space and Time

1.3.1 Literature gap

Scholars in conflict and crime study have reduced the gap between theories and empirical findings through the application of spatial and space-time stochastic models on a suitable scale (local scale rather then country-level scale) using disaggregated data (Lewis et al., 2012; Mohler, 2014; Raleigh et al., 2010b; Zammit-Mangion et al., 2013, 2012) (more detail in Sections 2.3 and 2.4). In contrast, research in terrorism so far has failed to capture the fine-scale spatial dynamics of terrorism.

In the 1980s, when scholars initiated the development of theories to explain the spatial dynamics of terrorism, they did not have access to large datasets on geolocalised terrorist events. Their focus was based on identifying and explaining mechanisms that might occur among countries. Hence, local variability was ignored despite that most explaining factors (e.g. demography or ethnic diversity) of terrorism may considerably vary within country. Furthermore, there is no single country where the locations of terrorist attacks are evenly distributed throughout its territory (Figure 4.3). An aggregated measure at country-level says little about terrorism, which is materialised into terrorist attacks that occur *locally* and exhibit specific characteristics according to the location in which they take place.

Equally problematic, more recent theories that relate the propensity for terrorism activity to the location (Nunn, 2007), have not been systematically assessed in both the spatial and temporal dimensions yet. For example, the theories of repeat victimization, crime generator, terrorism geography, and flag theories (more detail in Section 2.2.2) assume that the characteristics of the location are crucial and may be a major driver of different types of crime and terrorism (Brantingham and Brantingham, 1995; Eck et al., 2005).

In effect, most stochastic models used so far have focused on the temporal dimension only (Enders and Sandler, 1999, 2000, 2005; Enders et al., 2011; Hamilton and Hamilton, 1983; Porter and White, 2012; Raghavan et al., 2013). Moreover, studies that have explicitly integrated both space and time dimensions have been carried out at a country or higher level of analysis (Enders and Sandler, 2006; Gao et al., 2013; LaFree et al., 2010; Neumayer and Plümper, 2010), or at a regional or local level of analysis but in restricted study areas (Behlendorf et al., 2012; LaFree et al., 2012; Mohler, 2013; Nunn, 2007; Piegorsch et al., 2007).

1.3.2 Introducing the SPDE Approach

This present work uses geolocalised data on worldwide terrorism in combination with recent Bayesian modelling and inferential techniques to fill the literature gap. An elegant approach used to represent complex space-time interactions is the *Stochastic Partial Differential Equation* (SPDE) models (Lindgren et al., 2011). In order to better appreciate Lindgren et al.'s contribution, I provide a brief introduction to the origins and the main development that led to the SPDE approach. The approach takes its origin from the *Partial Differential Equation* (PDE). It is defined as "any equation which involves an unknown function of two or more independent variables and one or more of its partial derivatives" (Evans, 2010, p. 1). A well-known PDE is the (one-dimensional homogeneous) heat equation, with the temperature z, initial condition $z(s,0) = z_0(s)$, and constant $k \in \mathbb{R}^+$:

$$\frac{\partial z(s,t)}{\partial t} = k\frac{\partial^2 z(s,t)}{\partial s^2} \tag{1.4}$$

Equation 1.4 tells us that the rate of change in the temperature (*left-hand side*) is proportional to the curvature of the temperature (*right-hand side*). PDE has been used to describe a wide variety of phenomena continuous in space and time with applications in ecology (Asensio and Ferragut, 2002; Holmes et al., 1994) and oceanography (Bennett, 2005), to name but two. When the processes that generate the phenomenon under study are not well understood — such as most (if not all) social phenomena — PDE is no more applicable. Recall that deterministic models such as PDE cannot properly model stochastic processes, since they do not account for the uncertainty inherent to stochastic processes (Section 1.2.1).

In order to take into account the stochastic aspect of the process under investigation, the SPDE models incorporate a spatio-temporal *Gaussian white noise* process, where each location in space and time is associated with a normally distributed variable with zero mean and finite variance[6] The one-dimension diffusion equation with a random signal is an example of a linear SPDE (Zammit-Mangion et al., 2013, pp. 20-22):

$$\frac{\partial Z(s,t)}{\partial t} = \frac{\partial}{\partial s}\left(D(s)\left(\frac{\partial}{\partial s}Z(s,t)\right)\right) + \varepsilon(s,t), \tag{1.5}$$

[6] Note that white noise is defined as a sequence of uncorrelated random variables with constant mean and variance. If the sequence of the random variables is also independent and identically distributed (i.i.d.) it refers to *strict* white noise. Note that Gaussian white noise is also strict white noise (Harvey, 1993, p.267).

where the phenomenon of interest, $Z(s,t)$, is a *random field*[7] which varies in space s and time t, $D(s)$ is a non-negative spatial parameter, δ is a non-negative number, and $\varepsilon(s,t)$ represents a spatio-temporal Gaussian white noise process. Note that on each time $t \in (0,T]$, the field takes values on a space $s \in \mathscr{D} \subseteq \mathbb{R}^d$ where the domain $\mathscr{D}$ is commonly assumed *bounded*, and $d = 1$ for a one-dimensional diffusion, $d = 2$ for a two-dimensional diffusion, etc.[8]

The SPDE approach is particularly suitable for describing processes continuous in space, which have been traditionally investigated through methods initially developed in the field of *geostatistics* (Krige, 1951; Matérn, 1960; Matheron, 1963). *Geostatistical* models[9] use measures from a finite number of observations in a given spatial domain $\mathscr{D}$ (with usually $\mathscr{D} \subseteq \mathbb{R}^2$, or $\mathscr{D} \subseteq \mathbb{R}^3$) to predict values at any location within $\mathscr{D}$. This approach assumes the existence of an underlying random field, most often specified as a *Gaussian random field* (GRF), a specific class of stochastic processes, which is characterised by Gaussian properties, and therefore, provides computational and analytical advantages (Diggle, 2007, p. 46). In particular, SPDE models — including but not limited to Lindgren et al.'s approach — have been extensively used to model natural phenomena, such as turbulence (Da Prato et al., 1994), seismology (Zhang et al., 2015), neurophysiology (Walsh, 1981), air pollution (Cameletti et al., 2013b), epidemiology and public health (Costa et al., 2015; Lai et al., 2013; Musenge et al., 2013), queuing theory (Kaspi et al., 2013), climatology and meteorology (Bolin, 2012; Geirsson et al., 2015; Ingebrigtsen et al., 2014; Wallin and Bolin, 2015; Zammit-Mangion et al., 2014), weather forecast (Möller et al., 2015), ozone estimation (Richardson et al., 2015)), and, more rarely, social phenomena, including investment decision (Chavanasporn, 2010) and conflict (Zammit-Mangion et al., 2013).

[7] The phenomenon under investigation (Z) is a random field, which consists of a stochastic process that takes values in a multidimensional space (usually in a Euclidean space) and is defined over a parameter space of dimension $d \geq 1$ (Adler and Taylor, 2009, p.1). Note that the random field generalises the concept of random variable. Consistent with the notation used in this thesis, random fields are written in capital letter.

[8] Note that some authors may use $\mathscr{D}_s$ or $\mathscr{D}_{\mathbf{s}}$ to specify a multi-dimensional spatial domain. In this thesis, all domains are written $\mathscr{D}$, regardless of their dimensionality. In equations 1.4 and 1.5, spatial location (s) is in a one-dimensional space and is therefore not written in bold. Further in the text, spatial location will be written in bold (i.e. $\mathbf{s}$) when a multi-dimensional spatial domain is assumed.

[9] More detail on geostatistical models is provided in Section 3.2.2.

1.3.3 Exploiting the Link between GMRF and SPDE

The GRF is a continuous field, which contains an infinite number of elements, and therefore needs to be *discretised* in order to be properly estimated[10]. Rue and Tjelmeland (2002) showed that Gaussian Markov random fields (GMRFs), consisting of GRFs with observations independent conditional on the neighbourhood of the GRF, have a very sparse precision matrix (inverse of the covariance matrix). As a result, the computational costs associated with the approximation of the precision matrix are drastically reduced (Rue et al., 2009). Moreover, GMRFs can be accurately constructed from a restricted class of SPDE models, which have a GF with Matérn covariance function as solution (Lindgren et al., 2011).

For common applications in two dimensions, Lindgren et al. (2011) only require the square-root of the time required by standard approximations of GFs, by replacing GFs by its discrete equivalent, the GMRFs. The authors used the Integrated Nested Laplace Approximation (INLA), which provides fast approximation of the posterior marginals (Rue et al., 2009). SPDE models do not require regular lattice grid data along with interpolation of observations to the nearest grid-point. They allow for finer resolution where details are required. Furthermore, SPDE models can define GRFs on manifolds (e.g. sphere) (Lindgren et al., 2011), which have a direct application in modelling the dynamics of lethal terrorism worldwide (see Chapter 5). Moreover, the approach can handle non-stationary spatial models as well (not discussed here, see Bolin and Lindgren (2011) for further detail) and explanatory variables can be easily included in the dependence structure, which leads to non-stationary second-order behaviour (Ingebrigtsen et al., 2014)[11].

---

[10] Several approaches have been used to approximate GRFs, through Markov representations of the GRF using standard finite element basis of piecewise linear basis functions (Lindgren et al., 2011) or wavelet basis functions. Alternative approaches — usually less accurate than Markov approximations — using process convolution method and covariance tapering methods have been developed as well. For a systematic comparison of the approaches, see Bolin and Lindgren (2011). [11] For a review of second-order non-stationarity models for univariate geostatistical data, see e.g. Fouedjio (2016).

1.4 Summary and Structure

1.4.1 Summary of Contributions

The main contribution of this thesis is the application of a new approach that systematically captures fine-scale spatio-temporal processes of lethal terrorism. Hence, this work provides a better understanding of the observed sub-national spatio-temporal patterns of lethal terrorism and a rigorous framework used to assess theories on a suitable level of analysis. Its contributions could be summarised as follows:

1. Novel implementation of the SPDE approach in the study of lethal terrorism worldwide (Section 5.2), which opens a new path to the investigation of related social phenomena (e.g. crime, conflict);

2. Accurate description of the global (Section 4.3.2) and local (Section 4.3.3) spatial structure of terrorism and identification of the scale on which clustering processes operate in terrorism (Section 4.3.4), and more specifically in lethal terrorism (Section 5.3.1);

3. Accurate measures of the uncertainty in the estimation of the probability and number of lethal terrorist events across the world and at high spatial resolution (Section 5.3.2);

4. Novel approach to detect fine-scale spatio-temporal processes, such as hotspot (Section 6.1.1), escalation (Section 6.1.3), and diffusion (Section 6.2.1) of lethal terrorism;

5. Identification of potential local drivers and quantification of their effects on spatio-temporal processes involved in the probability and the number of lethal terrorist attacks (Section 5.3.1);

6. Identification of locations at risk of lethal terrorism across the entire world and at high spatial resolution (Sections 6.1.2 and 6.2.2);

7. Assessment on a suitable scale of analysis of major theories that explain hotspot (Section 6.3.1) and diffusion (Section 6.3.2) processes of lethal terrorism.

The methods and results described in Chapters 5 and 6 have been reshaped and submitted as a full article in the *Journal of the Royal Statistical Society: Series A (Statistics in Society)* July 13, 2016. The paper is currently in revision and has been written in collaboration with

Dr Janine Illian (PhD thesis first supervisor), Mrs Charlotte Jones-Todd (PhD student at the University of St Andrews, UK), and Dr Marta Blangiardo (supervisor during my visiting period (2015-2016) at Imperial College, London). I am the first author and wrote the entire text. Mrs Charlotte Jones-Todd contributes to the design of the statistical models. Dr Illian and Dr Blangiardo provide support with regard to the specification of the models and review the paper. The other Chapters of this present work have not been submitted for publication.

1.4.2 Structure of the Thesis

In the previous Sections, I have provided a brief summary of the context, the motivation, the research questions, and the main contribution of this PhD thesis. It should be mentioned that this study aims at answering questions about *terrorism*. This very concept has been subject to vigorous debate in the literature, and therefore, needs to be discussed in order to identify the limits of the generalisability of this study. The purpose of Chapter 2 is therefore to briefly introduce the main current debates with regard to the epistemology, causality and rationality of terrorism. Moreover, a review of the literature in terrorism and related areas of research clarifies the concepts related to space-time data, spatial autocorrelation, hotspot and diffusion. The aims of this present study are described, along with an identification of the literature gap.

Chapter 3 provides a gradual introduction to the main concepts of statistical space-time modelling by examining the characteristics from the simplest temporal models to the most sophisticated spatio-temporal models used in the literature in terrorism and related fields of study. This review highlights the limits of each approach used so far and describes recent advances in spatial statistics and inferential techniques that will be applied in this study. Furthermore, it highlights the advantages of using the SPDE approach combined with accurate and fast inference techniques (INLA) in the study of lethal terrorism.

The main databases currently available that provide geolocalised data on terrorism worldwide are compared through an exploratory analysis of data carried out in Chapter 4. The analysis highlights the main differences observed in the characteristics of the data from each provider. Hence, the analysis compares the evolution of terrorism over time, according to each database. The spatial component of the data is further analysed globally, through the global index *Moran's I*, and locally, through the local index *Getis and Ord* $G^*_{\mathbf{s}_i}$. The scale on which cluster processes occur is identified by an analysis of point pattern based on the pair correlation function (*pcf*).

The selection of the database on terrorism and the covariates are described in Chapter 5. It further outlines the discretisation of the domain based on a triangular mesh and provides the specification of the Bernoulli and Poisson models, used to model the probability and the number of lethal attacks, respectively. The results of the models are interpreted within the context of the literature and maps illustrate the spatial dynamics of the studied processes. Furthermore, it includes a prior sensitivity analysis used to assess the robustness to changes in priors related to the Matérn covariance function, which is complemented by an assessment of the robustness to changes in the mesh size. The chapter concludes with an assessment of the validity of the Bernoulli and Poisson models.

Chapter 6 uses the results obtained in Chapter 5 to build a method used to detect hotspots, escalation and diffusion processes. The performance in the detection of these processes is assessed through a comparison of the results from the model with observed data. Moreover, the roles of the stability of states and local economy in the spatio-temporal processes of lethal terrorism are investigated and measured to assess the validity of several theories, from both global (comparison with world's average) and local (comparison among contiguous areas) perspectives. This thesis is concluded by Chapter 7, which recalls the objectives and provides a summary by chapter of the methodology and the points addressed in this work. Furthermore, it describes the main shortcomings and potential improvements and policy implications. `R` code (Appendices A, B, C, and D) and data are provided in the electronic version of this thesis only.

Chapter 2

Fundamental Issues of Terrorism

2.1 The Nature of Terrorism

As emphasised by English (2010, pp. 24-25), any serious work on terrorism needs to acknowledge both the ambiguity of the concept of terrorism and the absence of a consensual definition. Hence, studies based on vague and/or inconsistent interpretation(s) of the term do not allow generating solid knowledge required for policy recommendations. For the sake of transparency, scholars should (at least) explicitly state how terrorism is understood and how their interpretation relates to and/or diverges from other equally reasonable points of view. Accordingly, Section 2.1.1 highlights the current epistemic debates and provides the distinction between terrorism, crime and other forms of violence. More importantly, it clarifies the position adopted in this study and its impact on the generalisation of its results.

Equally problematic, the causes of terrorism are not well understood and have been the subject of an intense debate. Complex interactions of multidimensional factors are involved Richardson (2006, pp. 92-93), which makes causal inference especially challenging. Since this study aims to systematically assess the effects of major local drivers of terrorism in the entire world, a rigorous approach to the identification of potential explanatory factors is all the more important. For this purpose, Section 2.1.3 puts particular emphasis on providing a comprehensive review of the main potential drivers of terrorism identified by the literature.

Despite the seemingly irrational behaviour resulting from the extraordinary violence occasionally used by terrorist groups, Section 2.2.3 shows that the literature suggests that terrorists tend to behave rationally. Terrorism is not the output of "psychopathy" or "mental illness" (Silke, 2004). Moreover, the spatial patterns of terrorist events identified by the literature reviewed in Section 2.2 suggest that there is a rationale behind the choice of the locations of terrorist events. More precisely, Section 2.2.2 brings evidence that terrorist attacks often exhibit spatial dependence and are prone to form *hotspots* in both space and time, from which terrorism might spread towards neighbouring areas, or in contrast, reduce its activity, as described in Section 2.2.3. The study of the patterns and diffusion processes of terrorism has benefited from considerable advances made in related fields of research. Section 2.3 explores the main relevant findings from studies in conflict, crime and relate them with those in terrorism. Moreover, it identifies the drivers that lead terrorists to select their targets, thus explaining the observed patterns. Finally, Section 2.3 reviews the literature gap and summarises the contributions of this present research.

2.1.1 The Definitions of Terrorism

The concept of terrorism is intrinsically ambiguous, still being debated today (Beck and Miner, 2013; Gibbs, 1989). Hoffman (2006, p. 40) attempted to formulate a concise definition of terrorism: "[...] the deliberate creation and exploitation of fear through violence or the threat of violence in the pursuit of political change". Despite its remarkable conciseness, Hoffman (2006, p. 23) acknowledged the incompleteness of his definition: subjectivity is almost inevitable when groups or individuals are defined as terrorists. Inevitably, the definition of terrorism implies a judgement (O'Brien, 1983, p. 91).

While acknowledging that a unique and satisfactory definition could not be reached, Schmid and Jongman (1988, p. 5-6) opted for a systematic investigation of key elements which may occur in each possible form of terrorism. Based on 109 identified definitions of terrorism, the authors reported 22 main characteristics of terrorism, and classified them according to their frequency of occurrence. The first five elements in declining orders are: "violence-force"; "political"; "fear-terror emphasized"; "threat"; and "psychological effects and anticipated reactions". They noted that scholars used on average 8 among the 22 characteristics in order to define terrorism. Despite their considerable effort to systematically examine recurrent elements in all possible definitions, the authors acknowledged that these 22 characteristics are certainly not exhaustive.

Terrorism shares similarities and differences with criminality. Both terrorism and crime are disproportionately perpetrated by young males, and threat societies (LaFree and Dugan, 2004, pp. 54-56). Furthermore, most countries prosecute terrorist acts under criminal laws (Crelinsten, 2009, p. 52) without defining a particular crime category of terrorist acts (LaFree and Dugan, 2004, p. 57). In contrast to criminals that usually try to avoid exposure (LaFree and Dugan, 2004, p. 59), terrorists aim to generate fear on a target audience beyond the act itself, the word 'terror' coming from the Latin word *terrere*, which signifies to frighten, to scare (Weimann, 2012, p. 182). While criminal acts are principally founded on personal motivations, terrorist acts are essentially designed to achieve political goals (Hoffman, 2006, p. 36). Moreover, terrorists are usually organised through a group structure, while such organisational structure is less common in crime, except in a few cases including organised criminality (LaFree and Dugan, 2004, p. 60).

Further distinction of terrorism can be drawn by comparing the nationality and/or the origin of the perpetrators, victims and the target country. Most scholars argued that transnational terrorism differs from domestic terrorism (Gaibulloev and Sandler, 2008; Jongman, 1992; Sambanis, 2008). Enders et al. (2011) suggested the following distinction: "[d]omestic terrorism is homegrown in which the venue, target, and perpetrators are all from the same country", whereas "[t]hrough its victims, targets, supporters, or perpetrators, transnational terrorism concerns more than a single country". In contrast, Sánchez-Cuenca and De la Calle (2009) mentioned that the distinction is doubtful, as it is not theoretically founded. Indeed, since the 1970s almost all terrorist acts have resonated across national borders through the international diffusion of media or external funding from foreign countries (Kegley, 1990, p. 5). In line with Kegley and Sánchez-Cuenca and De la Calle, this study, based on relatively recent terrorist events (2002-2013), does not distinguish domestic from transnational attacks.

As an alternative, Hoffman (2006, pp. 34-37) suggested identifying negative definitions of terrorism. By distinguishing terrorism from other forms of political violence, such as guerilla and insurgency, a convergence towards a sensible definition of terrorism may occur. The author remarked that guerilla and insurgency may employ similar tactics; however, terrorists are usually constituted of a relative smaller number of combatants, and are typically logistically limited. He added that ordinarily, terrorists do not aim to control territory and are not consisted of uniformed armed entities.

Equally problematic is the search of a consensual legal definition of terrorism. Each country uses a peculiar legal definition(s) of terrorism, according to its own legal framework. At the EU level however, the Council adopted a Common Position, which defined terrorist acts as intentional actions which "may seriously damage a country or an international organisation" (The Council of the European Union, 2001). Under the 2002 Framework Decision on Terrorism, a definition has been established and shared by all Member States (Argomaniz, 2011, p. 85).

2.1.2 Terrorism from the Databases' perspective

This notwithstanding, an in-depth examination of the definitions provided by each state in the world or by each scholar in terrorism would not be relevant. This study is based on data, which have been gathered by a specific database provider through its own approach to identify observations as terrorist events. Therefore, I focus on the definitions provided by the main providers of geolocalised data on worldwide terrorism.

Most practitioners (Clauset et al., 2007; Drakos and Gofas, 2006b; Enders et al., 2011) and terrorism database providers (CEACS, 2013; Engene, 2007; LaFree and Dugan, 2007; Mickolus, 2003) exclude *state terrorism* — terrorism perpetrated by state, in contrast to terrorism perpetrated by non-state actors — from their definition of terrorism, consistent with Hoffman (2006, p. 40). Amnesty International (2013) provides annual reports on human rights violations and abuses perpetrated by states, which include state terrorism. Terrorist events, however, are not geolocalised. Dugan and Chenoweth (2013) created a database for terrorist actions perpetrated by states; it is limited by the data — Israel and Canada only — and lacks geolocalisation. Raleigh and Dowd (2015) annually update the *Armed Conflict Location and Event Data Project* (ACLED), which includes violence perpetrated by state actors against civilians. However, it covers Asia and Africa only. There is currently no available geolocalised data on state terrorism with a worldwide coverage. State terrorism is *de facto* excluded from this research work.

Nevertheless, it is worth remembering that state terrorism has produced numerous victims across the world, as reported by Chomsky (2002, pp. 119-133). In the 18th century, the concept of terrorism originally referred to state terrorism and was defined as "violent acts of governments designed to ensure popular submission" (Chomsky, 2002, pp. vii). For example, the Stalin, Hitler, and Pol Pot totalitarian regimes committed mass violence, including state terrorism (Gibbs, 1989). The US have been responsible for massive attacks deliberately targeting civilians, such as: the two atomic bombs launched on Hiroshima and Nagasaki in August 1945, the "Operation Just Cause", led by the US former president George H. W. Bush, which killed several thousands of people in Panama in 1989, or the use of terror by US military in Nicaragua; the latter has been condemned by the International Court of Justice in 1986 (International Court of Justice, 1986; Junkerman, 2002; Chomsky, 2003, pp. 16-

18). Other cases include Russia's attacks in Chechnya in the 1990s, the Israeli invasion of Lebanon in 1982, the use of terror in the 1990s by Turkey in the Kurdish areas, crimes committed by the Colombian army against trade unionist and journalists during the same period (Chomsky, 2003, pp. 52, 62-67), or the use of terror during the military regime *Proceso de Reorganización Nacional* in Argentina from 1976 to 1983 (Pion-Berlin and Lopez, 1991).

The *International Terrorism: Attributes of Terrorist Events* (ITERATE) defines terrorism as: "the use, or threat of use, of anxiety-inducing, extra-normal violence for political purposes by any individual or group, whether acting for or in opposition to established governmental authority, when such action is intended to influence the attitudes and behaviour of a target group wider than the immediate victims". The scope of ITERATE is drastically restricted since it includes only transnational terrorism; domestic terrorism is excluded (Mickolus, 2003, p. 2). The *RAND Database of Worldwide Terrorism Incidents* (RDWTI) defines terrorism as: "violence calculated to create an atmosphere of fear and alarm to coerce others into actions they would not otherwise undertake, or refrain from actions they desired to take. Acts of terrorism are generally directed against civilian targets. The motives of all terrorists are political, and terrorist actions are generally carried out in a way that will achieve maximum publicity". Furthermore, RDWTI distinguishes transnational from domestic terrorism (RAND, 2011).

Apart from databases including exclusively terrorism events, global databases which incorporate other forms of violence and social acts have been developed as well. This is the case of the *Global Database of Events, Language, and Tone* (GDELT), a recently established freely accessible database. GDELT events are coded entirely by machines and the database encompasses more than 200 million of events from 1979 to 2013 (Leetaru and Schrodt, 2013). The process of gathering and analysing data in GDELT contrasts with all databases previously identified. There is no formal definition of terrorism provided by the data provider. The classification of a terrorist event is based on the combination of words related to the act of terrorism, which comes from media's coverage. The automatized classification follows Conflict and Mediation Event Observations (CAMEO) rules, which organize events based on specific criteria (GDELT, 2013). Because of the difficulty to distinguish terrorism from other social acts, it remains problematic to assess its relevance with regard to terrorist events. In Chapter 4, I highlight the differences in spatial patterns of terrorist events provided by GDELT with other databases.

The *Global Terrorism Database* (GTD) provides several definitions of terrorism according to the time period relative to data collection. As regards data collected between 1970 and 1997, terrorism was defined as: "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation. For events collected after 1997, the definition includes at least two of the following three criteria: (i) "the act must be aimed at attaining a political, economic, religious, or social goal"; (ii) "there must be evidence of an intention to coerce, intimidate, or convey some other message to a larger audience (or audiences) than the immediate victims"; (iii) "the action must be outside the context of legitimate warfare activities" (GTD, 2014) (see working definition in Section 5.1.1).

2.1.3 Explaining Terrorism

In 1981, the eminent scholar Martha Crenshaw published an article under the evocative title "The Causes of Terrorism", which aimed at finding a common pattern of causation, based on a comparison of different cases of terrorism. Since then, a voluminous literature devoted to identifying the causes of terrorism has been produced. Before getting into the substance of these investigations, it is worth noticing that for centuries, the concept of *causality* has been debated in various academic disciplines. An in-depth review of the debates would go beyond the scope of this work. Rather, I suggest pointing out the understanding and interpretation of causality, which applies in the context of the quantitative literature on terrorism and related fields of study. As such, David Hume's claim that causes always precede the consequence has been central in time series analysis in particular. Hence, in the case of causality between one or multiple variables, it can be derived that information on past values may be used to forecast future values. In 1969, Clive W.J. Granger formalized one possible form of causality, called "Granger causality" as the following (Kirchgässner et al., 2013, pp. 95-96):

$$\sigma^2(Y_{t+1}|I_t) < \sigma^2(Y_{t+1}|I_t - \bar{X}_t) \tag{2.1}$$

where X_t, Y_t are two different time series, $\bar{X}_t$ is the set of all current and past values of X, I_t represents the total information available at time t, including X and Y, and σ^2 is the variance of the forecast error. Hence, X is defined as (simply) Granger causal to Y if and only if the variance of the forecast error of predicting Y_{t+1} is smaller when using previous and current values of X. To put it another way, a random variable X is causal to Y if X

improves the estimation of Y. In total, Granger characterised 8 forms of causality between two time series, including *instantaneous causality*, which contrast with *lagged causality*; the former implies that the cause occurs simultaneously with its effect, the latter implies a lag time between the cause and its effect (Kirchgässner et al., 2013, pp. 97-98).

Richardson (2006, pp. 59-61) reminds us that it would be futile to provide a unique and simple explanation of the causes of terrorism. The causes of terrorism are multidimensional and operate at the *individual*, *group*, *sub-national*, *national*, and *transnational* levels (Brynjar and Skjølberg, 2000; Richardson, 2006, p. 60). At the *individual* and *group* levels, Crenshaw (1983, p. 29) argued that ideology may significantly affect the choice of target in particular. Wilkinson (1990, p. 141) noticed that the role of ideologies and beliefs has often been neglected although both factors are crucial to generating terrorism[1].

Research has focused on the investigation of social, political and economic factors mostly at national level. Lai (2007) provided evidence that states may reduce terrorism by controlling their territory, strengthening their capacity, and increasing operating expenses of terrorist groups. From a thorough review realised at national level, Gassebner and Luechinger (2011) identified potential variables which may positively influence the occurrence of terrorism in specific country and year such as: population, military expenditures, involvement in wars, foreign portfolio investment, and political proximity to the United States and other explanatory factors which may negatively influence the occurrence of terrorism: economic freedom, physical integrity, rights, and law and order.

The effects of economic conditions on terrorism have been commonly measured through indices of economic development, poverty and income distribution at national level (Drakos and Gofas, 2006b). With regard to poverty, Benmelech et al. (2012) provided empirical evidence that high rate of unemployment offers more well-educated and efficient individuals to terrorist organisations. However, the authors acknowledged that the external validity of their findings is highly questionable, since it is based on a case study (157 suicide terrorists that occurred from 2000 to 2006 in Palestine). In contrast, more comprehensive studies did not find any significant linear relationship between poverty and terrorism at national

[1] Note that this thesis does not explicitly consider the causes of terrorism at individual or sub-levels (e.g. causes at neuronal or gene level). Since the scope of this study aims at making inference based on terrorism data geolocalised in the entire world from 2002 to 2013, it would be unrealistic to attempt including the individual characteristics of each potential perpetrator and victims. Moreover, for most attacks, information on the perpetrator is not available, which makes the inclusion of individual-level factors practically infeasible.

level (Abadie, 2006; Drakos and Gofas, 2006b; Krueger and Laitin, 2008; Krueger and Maleckova, 2003; Piazza, 2006).

With regard to economic development, Gassebner and Luechinger (2011) did not find any significant linear relation between per capita GDP and the occurrence of terrorism. Enders and Hoover (2012) found that per capita GDP may not have a linear but a strong and significant positive non-linear relationship with domestic terrorism and a weak but positive significant effect on transnational terrorism. Note that average levels of economic development do not necessarily correlate with poverty. High levels of economic development might hide poverty rooted in social and economic inequalities, which cannot be properly approximated by per capita GDP for example. In this present work, I focus on economic development proxied by local measures of luminosity (for further details, see Chapters 5 and 6).

Religion, globalisation issues, and mass media are potential causes of terrorism that often operate at *transnational* level. Although religion may be a causal factor of terrorism in some cases, it has never been a unique cause of terrorism; rather, it interacts with other factors to contribute to terrorism (Richardson, 2006, pp. 92-93). Based on an extensive review of 60 quantitative studies, Hsiang et al. (2013) highlighted the significant role of climate events on the increase of interpersonal violence in a broad sense, including terrorism. Berrebi and Ostwald (2011) provided evidence that natural disasters create vulnerabilities, which may be exploited by terrorist groups. They estimated the effect of natural disasters on terrorist incidence, deaths and the number of wounded from terrorism for numerous countries in specific years.

Cronin (2003) suggested that globalisation may engender two antagonistic effects on terrorism. Globalised terrorism may generate long-term destabilisation, but conversely, may strengthen cooperation mechanisms among states, which are required to combat international terrorism. Equally ambiguous, media coverage of terrorism is a “double-edged sword”. Terrorists benefit from publicity and the attention that they desire, however it may also be damaging for them (Hoffman, 2006, p. 194). As stated by Rapoport (1996): “The relationship between publicity and terror is indeed paradoxical and complicated. Publicity focuses attention on a group, strengthening its morale and helping to attract recruits and sympathizers. But publicity is pernicious to the terrorist group too. It helps an outraged public to mobilize its vast resources and produces information that the public needs to pierce the veil of secrecy all terrorist groups require”.

Independently of the scale of the analysis, Crenshaw (1981) distinguished *preconditions* from *precipitants*. Preconditions are factors which influence the occurrence of terrorism on a long term. Furthermore, preconditions are labelled *permissive* if they offer openings to terrorist acts. Precipitants refer to direct causes of terrorism, as they operate just before the act of terrorism. Ross (1993) identified three permissive causes of terrorism: location, type of political system, and degree of modernisation. In addition, he pointed out seven precipitants: social, cultural, and historical enabling factors, organisational split and development, presence of other forms of political conflict, support, counterterrorist organisation failure, accessibility of weapons and materials, and grievances.

Sandler (2013) admitted that the literature did not converge to a consensual explanation of the causes of terrorism. Moreover, different sets of causal factors might be involved depending on the scale at which terrorism is observed. The detection of factors involved in local patterns of terrorism cannot be observed from a national level of analysis; rather it requires setting a sub-national unit of observation (Buhaug and Rød, 2006). Considerable effort has been made to carefully gather, identify and select potential causal factors from a wide range of possible variables highlighted in the literature, conditional upon the availability of data. In addition, this study applies formal tests to the selection of covariates, by weighting their explanatory power with the cost of increasing the model's complexity. More details are provided in Chapter 4.

2.2 Properties of Terrorism Data

2.2.1 Temporal and Spatial Autocorrelation

Working with spatial or spatio-temporal variables brings about several issues which need to be addressed. Similar to epidemic disease, earthquake and crime, characteristics of terrorism are often positively *autocorrelated* in time and in space (Braithwaite and Li, 2007; Heyman and Mickolus, 1980; Midlarsky et al., 1980; Neumayer and Plümper, 2010; Nunn, 2007; Piegorsch et al., 2007). *Autocorrelation* occurs when similarity is observed between close values of a random variable in time (temporal autocorrelation) or in space (spatial autocorrelation). Autocorrelation violates the independence assumption of most standard statistical procedures and specific statistical methods are required (Griffith and Layne, 1999, pp. 3-4).

The analysis of terrorism in time started with Jenkins and Johnson (1975), who provided a chronology of 507 international terrorist events worldwide from 1968 to 1974. Later, Mickolus et al. extended the list of events and published a series of reference books, which were regularly updated within the period 1980 - 2006 (Mickolus, 1980, 1993; Mickolus et al., 1988, 1989; Mickolus and Simmons, 1997, 2002, 2006). These reference works contained the date of all recorded worldwide terrorist events, which allowed the analysis of terrorism in time. In 1980, Gleason described the occurrence of terrorist events based on the data collected by Jenkins and Johnson (1975). Gleason suggested that terrorism events are unlikely to be independent; they exhibit positive autocorrelation in time. In other words, the occurrence of one terrorist attack makes the occurrence of subsequent attacks more likely. This observed increase in the number of events in time is often referred to a *reinforcement* process. A decrease in the number of events is often called *negative* reinforcement or *reversal* reinforcement in time (Hamilton and Hamilton, 1983; Most and Starr, 1980).

Based on data from 16 countries, which include more than 40 terrorist events per country from 1968 to 1978, Hamilton and Hamilton (1983) observed that in most countries, terrorism decreases exponentially after reaching a peak in frequency of attacks. Later, Holden (1986) observed that the occurrence of successful aircraft hijackings in the US increases the likelihood of further attempts of skyjacking inside and outside the US. Weimann and Brosius (1988) found that the variance of the number of terrorist events takes the form of a one-month cycle. Enders and Sandler (1999) suggested that successful incidents may generate additional events until the authority would find an effective counterterrorism tool. The observed reinforcement effect may have generated cycles of attacks and counter-attacks. Enders and Sandler (2000) found that terrorist attacks associated with casualties may present temporal cycles and the associated periodicity may be estimated.

Enders and Sandler (2005) studied radical changes (also called *structural breaks*) in trends of time-series data on terrorism. They did not find evidence that the occurrence of terrorist events changed significantly since 9/11. After 9/11, however, terrorist attacks seemed to become less logistically sophisticated, which could be attributed to a better protection of strategic targets. They found that the first break may be associated with the rise of fundamental terrorism since 1979, which corresponds to the Iranian revolution and the start of the Soviet war in Afghanistan (Cronin, 2003). The second break corresponds to the end of the Cold war. According to Blum et al. (2005), the end of the Cold war prompted many states to stop their support to terrorist groups, which resulted in a decrease of terrorism activity.

Scholars attempted to forecast terrorism through statistical models. Bogen and Jones (2006) forecast the evolution of the maximum number of deaths per event until 2080. Porter and White (2012) estimated the probability of future attacks based on the daily number of terrorist attacks in Indonesia from 1994 to 2007. Raghavan et al. (2013) predicted terrorist attacks perpetrated by the *Fuerzas Armadas Revolucionarias de Colombia* (FARC) in Indonesia/Timor-Leste and Shining Path (also called *Sendero Luminoso*) in Peru.

Furthermore, terrorism often exhibits autocorrelation in space, as observed in many other social phenomena. "Most social science variables tend to be moderately positively spatially autocorrelated because of the way phenomena are geographically organized" (Griffith, 2003, p. 5). Spatial autocorrelation has been famously expressed by Tobler in the so-called *first law of geography*: "[e]verything is related to everything else, but near things are more related than distant things" (1970). Longley et al. (2001, p. 101) stated that "spatial autocorrelation is determined both by similarities in position, and by similarities in attributes".

In the presence of spatial autocorrelation, the main assumptions of standard statistical models — residuals are independent and identically distributed (i.i.d.) — are violated. Hence, errors of the Ordinary Least Square (OLS) regression become spatially correlated. Consequently, OLS estimators are inefficient (but not biased), inference based on standard Student's t-tests and measures of fit (Pearson's r) are biased, and the validity of tests for heteroskedasticity (Anselin and Griffith, 1988) and tests for structural stability of the regression coefficients are affected as well (Anselin, 1990). As a remedy, scholars have developed methods: (i) to remove the nuisance generated by spatial autocorrelation; (ii) to model spatial autocorrelation through a *spatial structure*, which may provide a better understanding of the investigated phenomenon through revealing spatial dependences (see Section 3.2). Furthermore, spatial autocorrelation can be measured through various indicators. The next sections describe the manifestation of autocorrelation processes in space and time, which include *hotspot* and *diffusion* processes.

2.2.2 Hotspot

Concept of Hotspot

Positive *temporal autocorrelation* might lead to clusters in time, which are characterised by a high number of consecutive events close in time[2], so that initial events may activate further

[2] These two terms (cluster and hotspot) will be used interchangeably throughout this work.

events close in space and/or time to the initial triggering events (D'Orsogna and Perc, 2014). Analogously, positive *spatial autocorrelation* might generate *spatial* clusters, commonly understood as "location, or small area within an identifiable boundary, with a concentration of [...] events" (Anselin et al., 2000), or equally, as "[...] an area that has a greater than average number of [...] events [...]". Law et al. (2014) defined crime hotspots as: "areas that have high probabilities of the area-specific differential crime trend (δ_i) being greater than zero, and cold spots as areas that have high probabilities of area-specific differential crime trend (δ_i) being less than zero". Similarly, I suggest a definition used to detect hotspots (see Equation 6.1).

Moreover, one may classify hotspots into two categories. *Growing* hotspots are characterised by a significant and continuous increase in terrorist activity over time, which corresponds to an *escalation* process. Hence, I suggest a definition of escalation (see Equation 6.2). Growing hotspots may require particular attention with regard to their potential to intensify and spread further. Conversely, a significant and continuous decrease in terrorist activity over time refers to a *de-escalation* process (Anselin et al., 2000; Eck et al., 2005; Zammit-Mangion et al., 2012). Finally, *one-off* hotspots are hotspots that do not persist in time, which are also known as *ephemeral* hotspots (Zammit-Mangion et al., 2012).

Hotspots have been observed in phenomena closely related to terrorism, such as conflict, war, and crime: advances in crime studies brought important methodological and theoretical contribution. In particular, the increase of spatio-temporal crime data, computing power, and the development of GIS techniques allowed significant progress in the analysis and modelling of crime in space and/or time (Anselin et al., 2000; Grubesic and Mack, 2008; Ratcliffe, 2004; Ye and Wu, 2011). A vast literature has been devoted to the study of hotspots of crime (Ackerman and Murray, 2004; Chainey et al., 2008; Eck et al., 2000; Evans, 2001; Grubesic and Mack, 2008; Roncek, 2000; Weisburd et al., 2004), including: violent crime (Nelson et al., 2001), homicide (LaFree, 2005; Ye and Wu, 2011), assault/predatory (Loftin, 1986; Sherman et al., 1989), arson (Rogerson and Sun, 2001), drug (Weisburd and Green, 1995), property crime/burglary (Hakim and Shachmurove, 1996; Johnson, 2008; Ratcliffe, 2006; Rey et al., 2012), shooting (Ratcliffe and Rengert, 2008), or vehicle crime (Nelson et al., 2001; Ratcliffe, 2002).

Similarly, it appears that terrorism also exhibits hotspots in time (Enders and Sandler, 2000; Gleason, 1980; Hamilton and Hamilton, 1983; Porter and White, 2012; Raghavan et al., 2013; Weimann and Brosius, 1988), and in space or space-time (Braithwaite and Li, 2007; LaFree et al., 2012, 2010, 2009; Midlarsky et al., 1980; Nacos, 2010; Neumayer and Plümper, 2010; Nunn, 2007; Piegorsch et al., 2007; Steen et al., 2006). Whereas Steen et al. (2006) proposed that international terrorists can strike everywhere, most empirical studies tend to reject such assumptions. Evidence suggests that the distribution of terrorist targets is not completely random in space (Behlendorf et al., 2012), terrorist groups act with purpose (LaFree et al., 2012), and consequently, terrorist events are usually clustered rather than presenting completely random patterns (Siebeneck et al., 2009). Thus, it corroborates the findings of Hoffman (2006, p. 73): "[...] all terrorist groups have one trait in common: they do not commit actions randomly or senselessly".[3]

Mueller (2005) claimed that: "[i]nstead of maintaining that terrorists might strike anywhere at any time, and thereby stoking the fear of random violence, it might make sense to suggest that only certain (relatively small) areas are primarily at risk". From a modelling perspective, it implies two important points. First, there is rationale behind terrorism patterns and its causes could be investigated, hopefully identified, and eventually modelled. Therefore, understanding the processes that lead to patterns of terrorism is crucial in order to build sensible models. Second, explaining terrorism attacks in space and time through current stochastic modelling approaches looks promising, since the uncertainty might be accurately quantified, as exemplified by the work of Zammit-Mangion et al. (2013, 2012).

However, the causes of terrorism and the processes that lead to the formation of hotspots of terrorism still elude scholars. To date, no general theoretical laws of terrorism have been unanimously agreed (Crenshaw, 2014). Although the root causes of terrorism appear to be hardly generalizable, the rationale behind terrorist targeting appears consistent over all terrorist groups. Patterns have often been revealed through advances in data analysis and modelling techniques made in related fields of study, such as crime and conflict studies, which provide valuable theoretical and empirical knowledge in the understanding of terrorism's hotspots. While by no means exhaustive, the next Section discusses the major theories developed in the field of terrorism and related fields of study, aiming at explaining the causes of hotspots in terrorism.

[3] Note that most scholars agree upon the somewhat surprising normality of terrorists and their *rational* behaviour at group level (Crenshaw, 1990; Enders and Su, 2007; Hoffman, 2006; Nemeth, 2010; Richardson, 2006). At the individual level however, the rationality of terrorism is more controversial, especially in the case of suicide terrorism. For further discussion on suicide terrorism, see Hoffman, 2006, pp. 132-133, Pape, 2006, or Richardson, 2006, pp. 135-155. Note also that the concept of rationality has eluded scholars for centuries and an in-depth analysis would go far beyond the scope of this work. For further reading on the philosophical understanding of rationality, see e.g. Elster (2009).

Causes of Hotspots

Within the numerous fields of research devoted to the study of violence in general, scholars in crime studies pioneered theoretical work on the causes of hotspots. Some theories put emphasis on the *conditions* for the meeting between an offender and its target. Directly derived from ecological theories, the *foraging theory* predicts that offenders move in space and time towards potential victims to generate crime opportunity (Cohen and Felson, 1979). The *routine activity theory*, a dominant theory in socio-spatial criminology, posits that most crimes require a synchronous meeting of interdependent actors in a particular space and time (Cohen and Felson, 1979). Hence, crime may occur if the following elements intersect in a specific time and location: “a motivated offender, a suitable target and the absence of a capable guardian” (Ratcliffe, 2002). The “guardian” acts as crime inhibitor since it reduces the opportunity for crime (Ratcliffe, 2006). If one of these three elements (offender, target, absence of guardian) is missing, a crime cannot occur (Cohen and Felson, 1979).

Similar to routine activity theory, *time geography* emphasises the role of space and time, in that criminal activities happen at the crossroads of life paths of various actors (Hägerstraand, 1970). Here, time plays a crucial role as it produces constraints on travel distance of both offender and its target(s), which restricts the area of a potential meeting, as suggested by the *crime potential theory* (Gorman et al., 2013). Thus, if the offender reduces its travel speed, less territory is covered in one time unit and consequently, the area at risk is reduced and can be quantified (Miller, 2005).

Other theories put emphasis on the *characteristics* of the location, which is associated with crime. The *repeat victimization theory* explains that some locations are repeatedly targeted by criminals because they provide good opportunities that can be further exploited by the offenders (Eck et al., 2005). In this framework, the theory assumes that knowledge gained from a crime perpetrated in one location encourages the offenders to target the same location again instead of identifying new locations, which adds additional costs for the offenders (Johnson et al., 2007). As a complementary explanation, the *social disorganisation theory* posits that “neighborhoods with greater residential instability, lower socioeconomic status, and more ethnic heterogeneity are more likely to experience disorder” (Shaw and McKay, 1942) (as cited in Steenbeek and Hipp, 2011). In empirical studies, the measure of social disorganisation has been usually assessed through variables which inform on ethnic diversity, poverty, family problems, and population turnover at the neighbourhood level (Andresen, 2006).

Boost theory assumes that initial criminal events are expected to increase or "boost" the probability of other events to occurring at the same location. Hence, the theory posits that the same offender benefits from returning to a known target. As a complementary theory, *crime generator theory* states that the spatial location itself may favour criminal activities and/or attract different type of offenders which would lead to different types of crime. Brantingham and Brantingham (1995) characterised "crime generator" as a location that attracts people without being subject to any particular crime (e.g. shopping districts, sports stadium, travel nodes). In contrast, "crime attractors" may generate particular types of crime due to the activity carried out in the area (e.g. prostitution zone, drug markets). Equally, *flag theory* predicts that attractive and accessible residences considered as "crime magnets" (the equivalent of "crime attractor") exhibit a similar signal (or "flag"), which attracts different types of offenders (Pease et al., 1998, p. 8).

In line with social disorganisation theory, scholars in terrorism put emphasis on identifying the characteristics of locations that can be potentially targeted by terrorist attacks. The selection of targets may be driven by their level of vulnerability, as noted by Drake (1998, p. 337). By way of illustration, Sandler and Arce (2003) described the target choice of several fanatical terrorist groups based on two options: targeting a business, or alternatively, targeting tourists. The study concluded that terrorists preferably target tourists, considered more vulnerable than business. Since tourists are spatially diffused, it is more difficult to protect them. Santifort et al. (2013) provided evidence through a Bayesian approach that both domestic and transnational terrorists prioritise secure attacks on defenceless targets, such as bombing private events. Legal frameworks and security forces may also influence target selection; the security environment would influence the strategy and all activities of the group (Drake, 1998, p. 142). Krieger and Meierrieks (2011) showed that terrorists may preferably target countries characterised by: high density of population, strong economy, politically open but encountering instability within the country, and political proximity to the US. Nunn (2007); Piegorsch et al. (2007); Savitch and Ardashev (2001) found that terrorist attacks are directed mainly against urban areas rather than rural areas, which tends to generate hotspots. In contrast to rural areas, cities may contain rich, symbolic, and vulnerable targets (e.g. highly populated areas). Moreover, cities may provide more mobility, anonymity, audiences and larger recruitment pools, in comparison to rural areas (Crenshaw, 1990, p. 115).

More systematically, one may categorise potential targets according to their nature. Physical targets represent any physical objects or persons: a consulate, a military officer, or a politician, for example. A psychological target may be one or several individuals: a witness of a terrorist event, a radio audience, a family, or government representatives, for example. Usually, terrorists attack one or multiple physical targets which will affect one or multiple psychological targets (Drake, 1998, p. 8). Equally, targets can be classified according to their purpose. Drake (1998, p. 9) identified four classes: symbolic, functional, logistical, and expressive targets[4]. Terrorists attack symbolic targets, primarily to hurt the psychological target and expecting a reaction from it. Moreover, attacking symbolic targets attracts attention of the public (Crelinsten, 2009, p. 1-2) and satisfies their desire for renown and reaction (Richardson, 2006, p. 131). Terrorists may also attack functional targets, defined as a threat to them. Functional targets could be a traitor, a police officer, or a politician. Finally, terrorists may attack logistical targets in order to optimise their resources. Logistical targets provide money, weapons and other useful elements to terrorists.

Moreover, scholars have investigated the motivations of terrorists beyond target selection. A desire of revenge (Richardson, 2006, p. 131), indignation, or anger may push individuals or groups to employ terrorism as an expressive target (Drake, 1998, pp. 9-15). In addition, the ideology of a terrorist group may orientate its choices. Ideology may set the limit within which a group selects its targets (Drake, 1998, p. 16). Since terrorists aim to attain political change (Hoffman, 2006, p. 40), an efficient strategy needs to be defined in order to achieve it (Drake, 1998, p. 35). Therefore, one aspect of the strategy of terrorist groups is to maximise the impact on the psychological target (Drake, 1998, p. 53). Numerous factors involved in the process of target selection are determined by human, material, financial resources, and the leadership of the terrorist group (Drake, 1998, pp. 73-97).

Some authors found evidence that terrorists select targets in order to maximise the *expected utility*. Based on network analysis, Jordan (2008) suggested that the underground stations bombed on July 7, 2005 in London were not targeted randomly but selected in order to maximise the damage to the underground infrastructure. Sandler and Lapan (1988) suggested a model which depicted the relationship between domestic and transnational terrorist attacks and prevention costs associated to different types of target. Through the use of a bargaining model, Nemeth (2010) found that government attributes, public support, and

[4] It should be noted that the four classes in Drake's classification are not discrete since the choice of targets is generally determined by multiple causes. Moreover, the analysis of terrorist targets does not necessarily shed light on the real intention of the terrorists (Drake, 1998, p. 9)

group-environment factors influence terrorist targeting in the US. Drake (1998, pp. 144-146) mentioned that terrorists do not only need the support of their members, but they also require funding from sponsor, and public opinion support. Elbakidze and Jin (2012) found that victims from countries which highly contribute to the United Nations tend to be more frequently attacked by terrorists. Their study covered all countries in the world from 1980 to 2000.

Terrorist targets do not appear to be selected randomly but rather as the result of a rational selection process. Specificities of the location along with the goals and motivations of the terrorist groups influence the choice of targets. As a result, locations more frequently targeted by terrorism may generate hotspots in space, time, or space and time. Furthermore, terrorism might spread from one location to another (Morgenstern et al., 2013), which is the topic of the next Section.

2.2.3 Diffusion

Concept of Diffusion

In contrast to hotspots that are typically described as static objects, *diffusion* processes are inherently dynamic. The etymology of diffusion comes from the Latin word *diffundere*, which signifies to pour out or shedding forth (Oxford English Dictionary Online, 2014). By analogy with the chemical process of diffusion in which a flow is generated from areas of higher concentration to areas of lower concentration (coldspots) (Encyclopaedia Britannica Online, 2014), terrorism violence may spread from areas of higher terrorist activity (hotspots) to areas of lower terrorist activity. Diffusion implicitly refers to an increase in the occurrence of events in a specific location. Negative diffusion is also called *dissipation*. This study suggests a definition of diffusion (Equation 6.4), which is further discussed in Section 6.2.1.

Initial studies of diffusion processes are far-removed from terrorism research. As a pioneer, Thomas Graham, a Scottish chemist, carried out the first systematic study of the diffusion processes in gases from 1828 to 1833 (Philibert, 2005). In the following centuries, diffusion processes were investigated in many other fields, including epidemiology (Jacquez, 1996; Mantel, 1967) and seismology (Mohler et al., 2011), whose development inspired further investigations in social science.

Early studies in war suggested that war tends to spread as a contagious diffusion over time through contact among actors of war (Davis et al., 1978). Later, it has been shown that conflicts tend to be locally clustered (Buhaug and Gleditsch, 2008), and might spread towards neighbourhood regions, as exemplified by violence in Chechnya, which diffused into neighbouring republics (O'Loughlin and Witmer, 2011). However, diffusion of war towards neighbouring countries remains rare (more than 97% of 3,746 nation-years did not observe war diffusion), although it appeared relatively common in international war (39%) (Collier and Hoeffler, 2004).

Diffusion processes in terrorism have been traditionally analysed in time only (Enders and Sandler, 2000; Gleason, 1980; Hamilton and Hamilton, 1983; Porter and White, 2012; Raghavan et al., 2013; Weimann and Brosius, 1988). Few studies have investigated the processes in space, or in both space and time (Gao et al., 2013; Lewis et al., 2012; Linke et al., 2012; Medina et al., 2011; Most and Starr, 1980; Schutte and Weidmann, 2011). Forsberg (2014) considers diffusion as *unobservable* because it is usually derived from measures of correlation or proximity in space and/or time between events, which can be subject to misestimation, as reported by Black (2013). However, recent approaches further improved the accuracy of the measure of diffusion process and in particular its uncertainty (Zammit-Mangion et al., 2013, 2012). Different nomenclatures have been used to describe diffusion according to several sub-categories of the process (Forsberg, 2014).

Cohen and Tita (1999) distinguish diffusion from *contagion*: the former refers to "the general process of movement", whereas the latter refers to "one mechanism for achieving that movement", which assumes that other non-contagious mechanisms may be involved in diffusion processes as well. The distinction between contagious and non-contagious processes proved to be challenging (Forsberg, 2014) and refers to Galton's problem. In 1899, Edward Tylor presented a paper in which he assimilated an observed correlation between marriage laws and descent patterns as causal. Galton questioned the independence assumption on which Tylor's conclusions were drawn and argued that a contact between tribes cannot be excluded (which is a violation of independence assumption) and rather than being functional, the relationship between marriage laws and descent patterns may come from a common cause (Ross and Homer, 1976).

By analogy with disease's spread, "*[c]ontagious diffusion* depends on direct contact and is the classic way of spreading disease" (italics in original) (Cohen and Tita, 1999). In this case, spatial proximity is a necessary condition. Moreover, contagious diffusion involves two entities: the event(s) which may have potentially triggered the process of diffusion and spatial, temporal or spatio-temporal location(s) that have been affected by the presence of triggering event(s) (Forsberg, 2014). In contrast, "[...] *hierarchical diffusion* does not require direct contact but, rather, occurs through spontaneous *innovation* or *imitation*" (italics in original) (Cohen and Tita, 1999). Moreover, in a hierarchical diffusion, the interactions among events within a hotspot could be generated by a common factor (Heyman and Mickolus, 1980; Midlarsky et al., 1980; Neumayer and Plümper, 2010). Thus, political conflict may diffuse transnationally independently from any deliberate transnational actions (Hill and Rothchild, 1986).

Based on their study of civil war, Schutte and Weidmann (2011) mentioned an additional type of diffusion, where the spatial location changes and relocates beyond the close neighbourhood: they referred to as *relocation* diffusion. Since relocation diffusion can theoretically occur at any distance from the source of conflict, its identification remains challenging. Locations that encounter relocation diffusion might be confused with hotspots that would occur independently from the source of conflict. Therefore, I exclude the investigation of relocation diffusion in this study.

Causes of Diffusion

Since the end of the 1970s, scholars have attempted to better understand the spread of terrorism in space and in space and time. Needless to say, terrorism would neither occur nor spread without the existence of terrorist groups, which pursue goals based on a particular ideology. Moreover, the occurrence of terrorist acts alone cannot explain its spread. Other causal mechanisms should be involved as well (Cliff and First, 2013). It appears that the spread of terrorism may be caused by spatial proximity, which refers to *contagious* factors: (Forsberg, 2014; Gao et al., 2013; Heyman and Mickolus, 1980; LaFree et al., 2009; Midlarsky et al., 1980; Neumayer and Plümper, 2010), and non-contagious factors: modern mass media (Wilkinson, 1979, p. 103; Mazur, 1982; Crenshaw, 1991; Martin, 1990, pp. 161-162; Weimann, 2008; Nacos, 2010), financial support (Heyman and Mickolus, 1980), transnational collaboration (Heyman and Mickolus, 1980; Crenshaw, 1983, p. 17; Brynjar and Skjølberg, 2000), freedom of movement and transportation (Heyman and Mickolus, 1980; Wilkinson, 1979, p. 189), state capacity (Helman and Ratner, 1992; Gros, 1996; Zartman, 1995, pp. 1-13; Braithwaite and Li, 2007), state policy (Murphy, 2003, p. 49), population density (Crenshaw, 1990, p. 115; Savitch and Ardashev, 2001), and the dynamics of the terrorist group (e.g. tactical change) (LaFree et al., 2012; Raghavan et al., 2013).

Seminal studies emerged two years after pioneering research on spatial processes of war carried out by Davis et al. (1978). Midlarsky et al. (1980) proposed the *hierarchical diffusion theory*, which states that countries with high diplomatic status ("strong" states) appear to reproduce the patterns of terrorism of low diplomatic status countries ("weak" states). Consistent with the hierarchical diffusion theory, *failed state theory* (Gros, 1996; Helman and Ratner, 1992), also called *state collapse theory* (Zartman, 1995, pp. 1-13), states that violent conflict and terrorism spread from weak countries to neighbourhood countries essentially because of 'the weak states' inability to control bordering areas[5]. An incapacity to stop terrorism's spread could be explained as a result of a lack of financial, human resources, and infrastructure to control territory (Takeyh and Gvosdev, 2002; Salehyan, 2006, p. 70).

The *theory of radicalization* argues that countries that encounter repeated violence may become a "breeding ground for terrorism" (Conteh-Morgan, 2004, p. 256) (as cited in Cliff and First, 2013). Piazza (2008) brought evidence that failed states appear more prone to transnational terrorism. Hence, the spread of terrorism within and across national borders is facilitated through: state policy (Murphy, 2003, p. 49), freedom of movement and transportation (Heyman and Mickolus, 1980; Wilkinson, 1979, p. 189), transnational collaboration (Heyman and Mickolus, 1980; Crenshaw, 1983, p. 17), and exchange of information (Brynjar and Skjølberg, 2000) for example.

More generally, Huntington (1993) suggested that the diffusion of conflicts could be explained by "civilisation rallying effect". In other words, if groups from a particular civilisation[6] but from different countries target a country from another civilisation, it will influence other groups to pursue attacks on targets from the same country and civilisation previously attacked. However, based on *COSIMO 1* conflict data[7], Chiozza (2002) did not find empirical evidence that conflicts are more frequent between states from different civilisations, not even in the post-Cold War period. Hence, Neumayer and Plümper (2010) suggested that terrorist groups that share similar ideologies compete for scarce financial, logistic and

---

[5] The concept of failed states is known to be ambiguous. However, common features of failed states can be identified such as "their inability or unwillingness to protect their citizens from violence and perhaps even destruction" or "their tendency to regard themselves as beyond the reach of domestic and international law, and hence free to carry out aggression and violence" (Chomsky, 2006, pp. 1-2). [6] Huntington (1993) defines civilisation as "[...] the highest cultural grouping of people and the broadest level of cultural identity people have short of that which distinguishes humans from other species". [7] Chiozza (2002) used the first version of data, *COSIMO 1*. Since then it has been completely restructured to *COSIMO 2*, a relational data base. It has recently been renamed to *CONIS* (Conflict Information System). More details are provided on: http://www.hiik.de/en/kosimo/kosimo1.html.

recruitment support. Therefore, this competition may increase their propensity to attack enemies from other "civilisations"[8].

Except few studies, including Picard (1986), who claimed that there is no "credible supportive evidence" that mass media are a factor of contagion; most studies found evidence that the media play an important role in spreading terrorism through demonstration and imitation processes (Brosius and Weimann, 1991; Brynjar and Skjølberg, 2000; Enders et al., 1992). The coverage of successful terrorist operations provide crucial information for other terrorist groups to perpetrate further attacks (Holden, 1986). By acknowledging that "[t]errorism is a theatre" (Jenkins, 1975, p. 16) (as cited in Hoffman, 2006, p. 174), scholars argued that terrorists use media as an instrument to spread fear and terror (Jenkins, 1983; Hoffman, 2006, p. 174; Weimann, 2005).

Empirical studies showed that both contagious and non-contagious processes might operate simultaneously. Based on the analysis of terrorism that occurred in three states dyads (Lebanon-Israel, Columbia-Peru, and India-Pakistan), Cliff and First (2013) found that both non-contagious and contagion diffusion across borders can occur. Moreover, the authors found evidence of diffusion as a tactic used by terrorists. LaFree et al. (2012) compared terrorist events perpetrated by the Spanish terrorist group ETA from 1970 to 1978 with events from 1979 to 2007 in different French and Spanish regions. They identified and quantified the differences in spatial patterns of terrorist attacks between these time periods. They showed that contagious diffusion and hierarchical diffusion appear related to the strategies adopted by ETA, *control* versus *attrition* attack, respectively.

2.3 Literature Gap

There is a rich empirical and theoretical literature that has actively investigated the causes, clustering and diffusion processes in terrorism and in related fields of study. Development in mathematics, statistical modelling, GIS, and the increase of available data and computational power have lead to a better understanding of the spatial dynamics of terrorism and related phenomena. In conflict and in crime research, scholars have successfully reduced the gap between theoretical and empirical findings through local-scale analysis rather than the commonly-used country-year level of analysis studies (Anselin et al., 2000; Grubesic

[8] Although Neumayer and Plümper (2010) observed spatial dependencies for specific civilisation dyads with regard to terrorist attacks that occurred in the post-Cold War, contagion effect attributed to Huntington's theory (Section 2.2.3) does not significantly affect the pattern of international terrorism.

and Mack, 2008; Raleigh et al., 2010b). To achieve this, scholars have used data aggregated into sub-national geographical areas, such as regular grid, also called *grid-data* (Buhaug and Gleditsch, 2008; Buhaug and Rød, 2006; O'Loughlin and Witmer, 2011; O'Loughlin et al., 2010; Raleigh and Hegre, 2009; Raleigh and Urdal, 2007; Rezendes and O'Sullivan, 1986; Weidmann, 2013), or disaggregated data which usually consist of geolocalised points in space or space and time (Lewis et al., 2012; Mohler, 2014; Mohler et al., 2011; O'Loughlin et al., 2010; Raleigh et al., 2010a; Rey et al., 2012; Zammit-Mangion et al., 2013, 2012).

In the field of terrorism, however, most research has focused on national level of analysis, whilst ignoring sub-national dynamics of terrorism (Braithwaite and Li, 2007). Moreover, scholarly work that investigated sub-national level unit of analysis processes have used most exclusively descriptive or confirmatory analysis (Berrebi and Lakdawalla, 2007; Brown et al., 2004; Gao et al., 2013; LaFree et al., 2010, 2009; Lakdawalla and Zanjani, 2005; Nunn, 2007; Öcal and Yildirim, 2010; Piegorsch et al., 2007) rather than using advanced approaches able to accurately model spatial dynamics and the uncertainty that comes with high-resolution spatio-temporal data. Except some rare examples such as the work of Mohler (2013) and Nemeth et al. (2014), it appears that terrorism research has not made the most of disaggregated data and recent development in statistical modelling.

Consequently, theories explaining the spatial dynamics of terrorism have not been systematically assessed on a fine-grained scale, which could have been beneficial to academic and practitioners alike. So far, most studies in terrorism have quantitatively assessed the theories of contagion by focusing exclusively on processes in time, neglecting the spatial dimension of the phenomenon (Barros, 2003; Bilal et al., 2012; Brandt and Sandler, 2012; Enders and Sandler, 1993, 1999, 2000, 2005; Enders et al., 2011; Gleason, 1980; Hamilton and Hamilton, 1983; Holden, 1986; Mohler, 2013; Porter and White, 2012; Raghavan et al., 2013; Suleman, 2012; Weimann and Brosius, 1988). Other studies have focused on the spatial dimension only, failing to consider the spatial dynamics of the process of contagion (Braithwaite and Li, 2007; Brown et al., 2004; Nunn, 2007; Piegorsch et al., 2007; Savitch and Ardashev, 2001).

Moreover, most studies that have considered both space and time dimensions have been carried out at national or wider level of analysis (Enders and Sandler, 2006; Gao et al., 2013; LaFree et al., 2010; Midlarsky et al., 1980; Neumayer and Plümper, 2010), or at regional level of analysis but in restricted study area (Behlendorf et al., 2012; LaFree et al., 2012; Medina et al., 2011; Nunn, 2007; Öcal and Yildirim, 2010; Piegorsch et al., 2007;

Siebeneck et al., 2009). The validity of inference has been limited and the causes of spatial heterogeneity within countries have not been systematically investigated yet. So far, the literature has failed to identify and explain clustering and diffusion processes of terrorism at sub-national levels and across the world.

2.4 Conclusion

Zammit-Mangion et al. (2012) used sophisticated Bayesian spatio-temporal techniques to accurately model conflict intensity in Iraq. Based on the SPDE approach (Section 3.3), the authors took into account the uncertainty in diffusion and escalation processes, which brought valuable knowledge in the understanding of the mechanisms behind the conflict. Similar approaches have not yet been applied to the field of terrorism. As a result, sub-national processes, such as escalation or diffusion, have not been accurately modelled yet. This present research work addresses the literature gap by suggesting the use of state-of-the art modelling techniques that explicitly consider spatio-temporal interactions of terrorism and covariates at high-resolution and on worldwide scope. Thus, it provides a rigorous framework to the assessment of specific theories and findings that have been reviewed in this Chapter.

Moreover, the use of computationally efficient model fitting approaches such as INLA (Section 3.3), which can fit complex spatio-temporal models on a sphere[9] within a reasonable time frame, have not been deployed in the field of conflict, crime, or terrorism studies. Thus, by combining geolocalised worldwide data on terrorism (based on GTD (START, 2014)), techniques which model fine-scale spatial dynamics processes, and computationally efficient model fitting techniques, this research work contributes to significant advances with regard to the understanding of the dynamics of terrorism in space and time, whose causes remain important and unresolved issues in terrorism studies.

This work puts particular emphasis on the quantitative assessment of several theories, which predict the causes of clustering[10]. According to the *terrorism geography theory* —

[9] The SPDE approach allows the construction of high-resolution meshes, which discretise the entire surface of the earth at any desired level of resolution. Moreover, INLA allows us to be efficient on the meshes. More detail is provided in Section 5.2.1. [10] Note that some theories will not be assessed since data required to support their claims are not available. For example, it is highly problematic to quantitatively assess the *theory of radicalization* even though one might use a proxy for estimating radicalization through the measure of the degree of "radical beliefs" (Borum, 2011). Satisfactory proxies for radical beliefs are not available for the entire world and at sub-national level of analysis.

analogous to the *crime generator theory* —, locations described as "terrorism generator" or "terrorism attractor" should exhibit similar characteristics (Brantingham and Brantingham, 1995). In the same vein, Eck et al.'s *theory of repeat victimization* predicts that terrorists should find good opportunity to repeat attacks in the same locations, generating hotspots, as often observed in crime data (Johnson et al., 2008; Pease et al., 1998).

Hence, this study will assess the validity of the theories on a local scale through the investigation of potential effects of economic, and political factors involved in clustering processes of terrorism. First, one might reasonably expect a positive and significant relationship between population density and the lethality of terrorism. Savitch and Ardashev (2001) found that terrorist attacks are directed mainly against urban areas rather than rural areas. In contrast to rural areas, cities may expose rich, symbolic, and vulnerable targets. Moreover, cities may provide more mobility, anonymity, audiences and larger pool of recruitment in comparison to rural areas (Crenshaw, 1990, p. 115).

Second, despite economic factors appearing not to play a role at national level (Abadie, 2006; Drakos and Gofas, 2006a; Gassebner and Luechinger, 2011; Krueger and Laitin, 2008; Krueger and Maleckova, 2003; Piazza, 2006), within-country variation cannot be excluded and its analysis might reveal effects that cannot be captured from a national-level perspective. Richer areas tend to count more attractive and symbolic targets (Savitch and Ardashev, 2001). Since terrorists favour symbolic targets to attract attention of the public (Crelinsten, 2009, p. 1-2) and satisfy their desire of renown and reaction (Richardson, 2006, p. 131), it is reasonable to expect a positive and significant relationship between economic indicators and terrorism on a local scale.

In addition, the causes of the diffusion processes will be investigated through the assessment of the *failed states theory* (Gros, 1996; Helman and Ratner, 1992). Failed state theory (also-called *state collapse theory* (Zartman, 1995, pp. 1-13)) predicts that terrorism spreads from weak countries to neighbourhood countries essentially because of the incapacity of weak states to control bordering areas. Moreover, the spread is expected to be mainly located in failed states, less capable of preventing terrorism (Gros, 1996; Helman and Ratner, 1992).

Besides its methodological and theoretical contributions, this study may be valuable for policy makers since it provides a decision-support tool which can be exploited to better allocate resources —financial, material, and personnel (Porter and White, 2012) — in locations facing a high risk of lethal terrorism (hotspots) within countries or administrative divisions.

The output of the models provides quantitative measures of the probability of lethal terrorism and the expected number of lethal attacks along with the uncertainty of the measures at high spatial resolution and across the world. However, since input data used in the models are provided at city-level, the applicability of the models are restricted in use that require higher spatial accuracy (e.g. allocation of resources to counter terrorism at street-level). Second, counterterrorism measures can be accurately assessed (Perl, 2007) through accurate measures of the risk of lethal terrorism at different time intervals (i.e. before and after the implementation of counterterrorism measures in a given area). Third, the models are flexible since they can be easily updated with new data and applied in any desired study area and time framework.

In summary, the cutting edge spatio-temporal models applied in this study will provide a conceptually simple and mathematically rigorous way to combine data and models to quantify uncertainty and learn about model parameters. More precisely, the models accurately capture uncertainty related to a complex spatio-temporal phenomenon such as terrorism on a fine-scale and in the entire world. Moreover, it allows for a reduction of the uncertainty based on large dataset ($>$ 35,000 events) that covers terrorist events over 12 years across the world. Moreover, the methodological advances offer a rigorous framework to assess the theories which seek to increase our understanding of the spatial dynamics of terrorism, which will benefit academics and practitioners involved in the field of terrorism.

Chapter 3

Modelling Temporal and Spatio-Temporal Processes

In the previous Chapter, I introduced the concepts related to the pattern and diffusion processes of terrorism. This Chapter aims to review the main stochastic models used to model terrorism in time, space, and space and time. Note that algorithmic techniques such as supervised classification and unsupervised clustering methods, event sequences analysis, and network analysis, which have been also applied in terrorism studies (Schrodt et al., 2013), are not considered here. Algorithmic techniques are no less valuable than statistical methods to predicting social phenomena in space and time, as illustrated by recent promising works (Stanton et al., 2015; Subrahmanian et al., 2013; Thuraisingham, 2004). However, I exclude them from this review, since they are less suitable than statistical models to explain the relationship between covariates and lethal terrorism, and hence, less able to answer the research questions (Section 1.1).

This Chapter does not intend to cover the whole literature on time, spatial, and spatio-temporal modelling. Rather, its aims are: (i) to provide a brief overview of the modelling techniques used so far in terrorism and closely related fields; (ii) to identify the methodological gap; (iii) to describe a suitable approach to filling the gap. First, this review allows the non-specialist reader to familiarize gradually with fundamental concepts, which are shared among all the models. Second, it helps to understand the purpose and the limits of simple models and help appreciate the improvement provided by more sophisticated models.

Gradually increasing complexity, I start describing the simplest autoregressive time series models (Section 3.1), moving to the most complex spatio-temporal models (Section 3.2), while illustrating this review through applications in terrorism. Furthermore, I compare the main three approaches used to modelling spatial and spatio-temporal data, which include: geostatistical, lattice, and point processes models. Section 3.3 concludes the Chapter by describing a recent efficient approach, which will be used in this study to model lethal terrorism in both space and time (Chapter 5).

3.1 Time-Series Models

Time series analysis could be defined as the science of modelling dynamics. In this framework, the phenomenon under investigation is represented by a time series, defined as a sequence of values of a random variable Y_t typically taken at successive equally spaced discrete intervals over time (Harvey, 1993, p.1). In the Western world, the early development of time series analysis seems to go back to the 17^{th} century with the analysis of games of chance by Blaise Pascal and Pierre de Fermat, and profit analysis from John Graunt and other merchants (Klein, 1997). Since then, it has developed at a frightening rate. The models have been adjusted in order to improve their level of specificity, also improving their sophistication. For the sake of concision, I essentially focus this review on models which have been applied in terrorism[1].

3.1.1 AR Models

Certain phenomena exhibit *temporal dependencies*: present values depend on past values. One of these is temperature. It is likely that the average temperature on day t, say 23 °C, is similar to the average temperature on day $t-1$, which was perhaps 19 °C. The simplest — but not necessarily simplistic — way to represent temporal dependencies is the *first-order autoregressive* model (AR(1)), which expresses a simple departure from independence in time series (Cressie and Wikle, 2011, p. 19). It may be formulated as a first-order difference equation (Hamilton, 1994, p. 1):

$$Y_t = c + \rho Y_{t-1} + \varepsilon_t, \tag{3.1}$$

[1] For a thorough review of most temporal models, see e.g. Harvey (1993).

where c is a constant, the dependent variable[2] Y_t depends on time and is related only to its previous value Y_{t-1} through a linear operator ρ and an error term ε_t. The model assumes a *white noise* sequence ε_t, which consists of serially uncorrelated random variables with constant mean and finite variance (Hamilton, 1994, p. 7)[3]. Note that without ε_t, the process would be fully determined by the function ρ, which characterises a deterministic process in contrast to a stochastic process; the latter always includes at least one stochastic element. Often, the exact mechanism behind the process is not known, which calls for the use of a stochastic element in the model. Moreover, if $\rho = 1$, the process describes a particular case called *random walk* (Cressie and Wikle, 2011, p. 85).

Hence, one can build an AR(p) process which extends the temporal dependence, so that Y_t depends on p previous values, from $t-1$ to $t-p$, rather than depending exclusively on one previous state, as in Equation 3.1. Moreover, AR(p) can be further generalised to multivariate data, within a *p-order* vector autoregressive VAR(p) model. VAR(p) models have been used by scholars in terrorism studies in order to take into account simultaneous and lagged relationships among multiple variables (Barros, 2003; Brandt and Sandler, 2012; Enders and Sandler, 1993, 2000). Hence, the vector $\mathbf{Y_t}$ includes several dependent variables, including for example: Y_{1t}, which counts the number of terrorist attacks, and Y_{2t} which sums up the number of fatalities. In this representation of a VAR(p) model, $\mathbf{c}$ and $\boldsymbol{\varepsilon}$ are vectors of constant and white noise, respectively, and $\mathbf{P}_i$ with $i = \{1, \dots, p\}$ is a time-invariant matrix of autoregressive coefficients (Hamilton, 1994, p. 291).

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{P}_1 \mathbf{Y}_{t-1} + \cdots + \mathbf{P}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t \tag{3.2}$$

3.1.2 ARMA Models

A dynamic system may be constructed to form an autoregressive moving average (ARMA) model. ARMA(p,q) processes are composed by p autoregressive and q moving average parameters: $\boldsymbol{\rho} = \{\rho_1, \dots, \rho_p\}$ and $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_q\}$, respectively. The process can be for-

[2] The dependent variable Y_t (also called *response* variable, or *regressor*) could be *continuous* or *discrete*. A continuous variable can take all possible values in a given interval. In contrast, discrete variables can only take specific values within a given interval. The number of fatalities of a terrorist attack in a country for each year t is an example of a discrete variable Y_t. [3] Throughout this thesis, the symbol ε_t will always denote a white noise, and unless specified otherwise, it will have a mean of zero and finite variance σ^2.

mulated as (Hamilton, 1994, p. 59):

$$Y_t = c + \rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \cdots + \phi_q \varepsilon_{t-q}, \tag{3.3}$$

where c is a constant and $\boldsymbol{\varepsilon} = \{\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q}\}$ is white noise. An ARMA process may be suitable to describe and forecast the dynamics of an individual time series, characterised by a *weakly stationary* process in time. In the time dimension, a process is weakly stationary if neither the mean nor the *autocovariance* — the covariance of the process with itself at pairs of time points — of the process depend on time. The parameters of ARMA processes: c, $\boldsymbol{\rho}$, $\boldsymbol{\phi}$, and the variance of $\boldsymbol{\varepsilon}$ can be estimated through the likelihood function (Hamilton, 1994, p. 117), Bayesian analysis (Hamilton, 1994, p. 351), or through generalized method of moments (GMM) to name but three approaches (Hamilton, 1994, p. 409).

For example, Weimann and Brosius (1988) analysed the number of terrorist events Y_t aggregated into $t = \{1, \ldots, 228\}$ months from 1968 to 1986. A preliminary analysis based on the *autocorrelation function* (ACF)[4] indicated that the autocorrelation of Y_t does not extend beyond one month, which suggested the use of a first-order moving average MA(1) model — a special case of an ARMA model, also denoted ARMA$(0,1)$, which represents a MA(1) without autoregressive (AR) terms — defined as: $Y_t = c + \varepsilon_t + \phi\varepsilon_{t-1}$, with expected value $\mathbb{E}(Y_t) = c$. The authors found that Y_t does not occur purely randomly in time (ϕ significantly different from zero).

3.1.3 ARIMA Models

Non-stationary processes in time occur when the mean is not constant over time. These processes can be modelled through an integrated autoregressive moving average (ARIMA) , which has been used for example by Enders and Sandler (1991, 1996) to investigate the impact of terrorism on tourism and foreign investment and Faber et al. (1984) to modelling war outbreaks.

ARIMA(p,d,q) models are constructed on ARMA(p,q) models with an additional parameter d. Let us define $\Delta Y_t = Y_t - Y_{t-1}$ as the first difference ($d = 1$). A time series follows an ARIMA(p,d,q) process if the d^{th} difference $\Delta^d Y_t$ is a stationary ARMA(p,q) process. If

[4] The ACF of a time series Y_t is the covariance of Y_t and subsequent values of Y_{t-k} divided by the variance of Y_t. The division by the variance of Y_t generates a standardised covariance. Thus, for each time lag $k = 1, 2, \ldots$, the ACF provides a value between -1 and 1. Note that for $k = 0$, the ACF equals 1 since data are perfectly correlated with themselves at an identical temporal location by definition (Mitchell and Moore, 2002).

we consider the first difference ($d = 1$) only, an ARIMA($p, 1, q$) process can be written as the following (Cryer and Chan, 2008, p. 92):

$$\Delta Y_t = \rho_1 \Delta Y_{t-1} + \cdots + \rho_p \Delta Y_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \cdots + \phi_q \varepsilon_{t-q} \tag{3.4}$$

From an ARIMA(p, d, q) process, it is possible to generate "sub-models" that do not contain any autoregressive terms, such as an integrated moving-average (IMA (d, q)) model. Similarly, models without moving average terms may take the form of an autoregressive integrated model (ARI (p, d)) (Cryer and Chan, 2008, p. 93). The choice of the model depends on the assumptions made on the distribution of the data, which may vary according to the investigated phenomenon.

3.1.4 TAR Models

Drastic changes observed in time series data represent a specific case of non-linearity, which cannot be well represented through AR, ARMA, or ARIMA models. Abrupt changes in the activity of worldwide terrorism over time may be caused by political transitions (e.g. dissolution of the Soviet Union in 1991) and invasion (e.g. US-led coalition invasion of Afghanistan in 2001 and Iraq in 2003), to name but two. The threshold autoregressive (TAR) models have been developed to identify drastic changes (usually called *breaks*) in time series data, which split the data into different *regimes*. In its simplest version, TAR is constituted of two sub models, which are active or inactive based on a threshold parameter l (Cryer and Chan, 2008, p. 395):

$$Y_t = \begin{cases} \rho_{1,0} + \rho_{1,1} Y_{t-1} + \varepsilon_{1t}, & \text{if } Y_{t-1} \leq l \\ \rho_{2,0} + \rho_{2,1} Y_{t-1} + \varepsilon_{2t}, & \text{if } Y_{t-1} > l \end{cases} \tag{3.5}$$

where the ρ's are autoregressive parameters, l is the threshold parameter, and $\varepsilon_{1t}, \varepsilon_{2t}$ are white noise in regime type 1 and 2, respectively. In Equation 3.5, the first index of the ρ's represents the type of regime (1 or 2), while the second index distinguishes the ρ's within each regime's type. The model runs for two different regimes according to the threshold l. Note that TAR models may be extended to higher dimensional orders and to more than two regimes (Cryer and Chan, 2008, p. 399).

The use of TAR brought key insight into the analysis of drastic change in terrorism over time. Enders and Sandler (2002) detected cycles of transnational terrorism in the post-

Cold War. Later, Enders and Sandler (2005) identified two structural breaks in transnational terrorism from 1968 to 2000: the rise of Islamic fundamentalism (mid-1970s) and the end of the Cold war (1990s).

3.1.5 ARCH Models

So far, we have examined time series models which assume i.i.d. variance within the model (or within each sub-model in the case of TAR models). As a further improvement, the *p-order* autoregressive conditional heteroskedasticity (ARCH) model relaxes this assumption. In ARCH(p), the variance is not constant over time (heteroskedasticity) and it varies conditionally on p previous time periods. Hence, the square of the error term ε_t^2 follows a *p-order* autoregressive process AR(p):

$$\varepsilon_t^2 = c + \rho_1 \varepsilon_{t-1}^2 + ... + \rho_p \varepsilon_{t-p}^2 + \varepsilon_t, \tag{3.6}$$

where c is a constant, $\boldsymbol{\rho} = \{\rho_1, \ldots, \rho_p\}$ are the autoregressive parameters of the model, and ε_t is Gaussian white noise (Hamilton, 1994, p. 657). An ARCH(p) model can be extended to a more generalised model called generalized autoregressive conditional heteroskedasticity (GARCH) model, introduced by Bollerslev (1986). While ε_t^2 is modelled as an AR(p) process in ARCH(p), it is modelled as an ARMA(p,q) process in GARCH(p,q). By way of illustration, both ARCH(p) and GARCH(p,q) have been used to understand the effect of terrorism on the volatility of stock exchange markets (Bashir et al., 2013; Bilal et al., 2012; Suleman, 2012). More sophisticated forms of GARCH(p,q) exist, including non-linear relations and multivariate variable data (Hamilton, 1994, pp. 665-671).

3.1.6 Count data Models

The autoregressive models discussed so far are designed to describe or predict the values of single or multiple *continuous* dependent variable(s) that can take any value in $\mathbb{R}$ over time. However, dependent variables can also be *discrete*. In other terms, they are allowed to take only particular values, usually in $\mathbb{Z}$. One particular case is the so-called *count* variable. The values of count variables are the result of counting the number of outcomes in a system, usually over time. Count variables take non-negative integer values in $\mathbb{N}$. Count data models have been applied in numerous works in terrorism (Berrebi and Ostwald, 2011; Brandt and Sandler, 2012; Elbakidze and Jin, 2012; Enders and Hoover, 2012; Gleason, 1980; Hamilton and Hamilton, 1983). The homogeneous Poisson process model has been used as a benchmark model for count time series data modelling. The *probability mass function* (p.m.f.)[5] of a homogeneous Poisson distribution in time can be expressed as the following (Tijms, 2003, p. 3):

$$P(y_t) = \frac{(\lambda t)^{y_t}}{y_t!} e^{(-\lambda t)}, \qquad y_t = 0, 1, \ldots \tag{3.7}$$

where the random process Y_t can take any discrete value $y_t \in \mathbb{N}$, which counts the occurrence of events at time t[6]. The parameter λ is always positive and often called *rate* in one dimensional (e.g. time processes) and can be interpreted as the average number of points per unit of time. The main properties of the Poisson distribution are the following: (i) $Y_0 = 0$; (ii) independent increments (i.i.d. events); (iii) the number of events in any interval of length t is a Poisson random variable with expected value $\mathbb{E}(Y_t)$ and variance $\mathbb{V}(Y_t)$ equal to λt (Tijms, 2003, pp. 3-18).

If the underlying process is characterised by an autoregressive behaviour, the independence assumption is violated, and the homogeneous Poisson model would not be suitable since it might generate biased and inefficient estimates. As an improvement, Brandt and Williams (2001) proposed the Poisson autoregressive (PAR) model, which includes autoregressive terms. Hence, an autoregressive temporal component is added to Equation 3.7, and a PAR(p) model can be expressed as:

$$P(y_t) = \frac{\lambda_t^{y_t} e^{(-\lambda_t)}}{y_t!}, \qquad y_t = 0, 1, \ldots \tag{3.8}$$

where $\lambda_t = \sum_{i=1}^{p} \rho_i y_{t-i} + c$ represents the conditional expectation of an AR(p) process based on previous realisations $y_0, y_1, \ldots, y_{t-1}$, with p autoregressive parameters $\boldsymbol{\rho} = \{\rho_1, \ldots, \rho_p\}$, and a constant c. Although other autoregressive forms of Poisson models such as Poisson exponential weighted moving average (PEWMA) model can accommodate persistent time series processes, they may not be suitable to modelling cyclical or short-memory processes, the latter showing a mean-reverse effect (Fokianos, 2012, pp. 315-320).

[5] The probability mass function (p.m.f.) gives the probability that a discrete random variable takes one particular value from a finite or countably infinite possible outcomes. For continuous variables, the probability density function (p.d.f.) describes the probability that a continuous random variable takes its value within a specific interval that contains an infinite number of possible values. [6] Note that in count data models, $P(y_t)$ is used as a simplified notation to express $P(Y_t = y_t)$, which translates into "the probability that the random variable Y_t takes the value y_t".

Terrorism scholars interested in counting the number of deaths from terrorist attacks over time are often confronted with data including a large number of non-lethal events, and therefore, a high number of zeroes in the sample (*zero-inflation*). Often, terrorists favour publicity rather than fatalities (Martin, 1990, p. 159), which may result in a higher than expected number of non-lethal attacks (Drakos and Gofas, 2006a). Data with excess of zeroes cannot be adequately modelled by standard Poisson distributions or other distribution functions in the exponential family (Anselin and Griffith, 1988; Anselin, 1990; Yang et al., 2012, p. 6).

Zero-inflated models (ZIM) can be used to remedy the issues. ZIM models are a class of regression models that separate processes into two groups: (i) only zero values; (ii) a combination of non-zeros and few zeros (Hall, 2000). Moreover, in contrast to simple regression techniques, ZIM models accommodate several violations of assumptions such as: non-normal residuals and a skewed distribution. In this case, zero-inflated Poisson (ZIP) models, which are a sub-class of ZIM models, can be suitable to model count data which exhibit excess of zeros (Drakos and Gofas, 2006b). Hence, the p.m.f. of a ZIP model can be expressed as the following (Lambert, 1992):

$$P(y_t) = \begin{cases} \omega_t + (1-\omega_t)e^{-\lambda_t}, & \text{if } y_t = 0, \\ \frac{(1-\omega_t)e^{-\lambda_t}\lambda_t^{y_t}}{y_t!}, & \text{if } y_t = 1, 2, \ldots, \end{cases} \tag{3.9}$$

where y_t is a nonnegative integer since it represents count data, λ_t is the intensity parameter of the baseline Poisson distribution, and zero-inflation parameter $0 \leq \omega_t \leq 1$ represents the probability of extra zeroes. The probability that Y_t takes the value zero follows the top right hand of Equation 3.9. The probability that Y_t takes a positive integer value follows the bottom right hand of Equation 3.9. This model assumes i.i.d. observations.

The presence of outliers (e.g. mass-casualty attacks) and the excess of zeros are often both present in terrorism data. As a result, the observed variance (sample variance) is greater than the expected variance (population variance), which refers to *overdispersion*. ZIP models can be generalized to zero-inflated negative binomial (ZINB) models to take zero-inflation and overdispersion into account simultaneously. In the field of terrorism studies, both ZIP, ZINB and other forms of negative binomial models have been used to model the occurrence and fatalities of terrorist attacks (Burgoon, 2006; Drakos and Gofas, 2006b; Dreher and Fischer, 2010; Krueger and Laitin, 2008). The p.m.f. of the ZINB model is given by (Mwalili et al., 2008):

$$P(y_t) = \begin{cases} \omega_t + (1-\omega_t)(1+\frac{\lambda_t}{\alpha})^{-\alpha}, & \text{if } y_t = 0, \\ (1-\omega_t)\frac{\Gamma(y_t+\alpha)}{y_t!\Gamma(\alpha)}(1+\frac{\lambda_t}{\alpha})^{-\alpha}(1+\frac{\alpha}{\lambda_t})^{-y_t}, & \text{if } y_t = 1,2,\ldots, \end{cases} \tag{3.10}$$

where $\alpha \geq 0$ is a shape parameter which quantifies the amount of overdispersion. ZINB models are relatively flexible and particular cases can be derived. Note that the ZINB distribution approaches: (i) the ZIP distribution (see Equation 3.9) as $\alpha \to \infty$ and (ii) the negative binomial distribution as $\omega_t \to 0$. If both $1/\alpha$ and $\omega_t \approx 0$, the ZINB distribution reduces to the Poisson distribution (see Equation 3.7).

One limitation with zero-inflated models is that the zero observations come from two different data-generating processes, which might not necessarily reflect the true underlying mechanisms. In contrast to zero-inflated models, hurdle models (also called *two-part* models) assume that all zero observations have been generated by the same process, while different mechanisms generated non-zero observations. For example, hurdle models have been used to identify the relationship between temporal patterns of terrorism and political characteristics of countries (Neumayer and Plümper, 2010). Hurdle models include two components: (i) zero counts (e.g. days without any terrorist attack), which are usually in excess; (ii) non-zero counts (e.g. days with at least one terrorist attack). The truncated count component can be based on a Poisson or on a negative binomial distribution. If the truncated Poisson distribution is used, then hurdle models can be defined as follows (Mullahy, 1986):

$$P(y_t) = \begin{cases} \pi_t, & \text{if } y_t = 0, \\ \frac{(1-\pi_t)e^{-\lambda}\lambda^{y_t}}{(1-e^{-\lambda})y_t!}, & \text{if } y_t = 1,2,\ldots, \end{cases} \tag{3.11}$$

where the probability to observe zero observations $\pi_t = \omega_t + (1-\omega_t)e^{-\lambda_t}$, with $0 \leq \pi_t \leq 1$, is a reparametrisation of the ZIP distribution (Equation 3.9). Hence, the probability of clearing the *hurdle* (i.e. the probability to generate a non-zero count) is given by $1-\pi_t$ (Ridout et al., 1998).

More complex models have been developed (not detailed here), such as: the *self-exciting hurdle model* (SEHM). This model combines the properties of the hurdle model, together with temporal dependencies among events, materialised into a *self-exciting* component. Porter and White (2012) used a SEHM to predict the probability of future attacks based on the daily number of terrorist attacks in Indonesia from 1994 to 2007.

3.2 Spatial and Spatio-Temporal Models

So far, I have discussed the basic concepts of modelling through the examination of the main time series models, which have been commonly applied in terrorism and related fields of research. The fact remains that terrorist events are inherently *spatial* and are likely to exhibit dependences in the spatial dimension. Recall that the i.i.d. assumption does not hold in space (Cressie and Wikle, 2011, p. 4). These non-linearity characteristics must be taken into account, which require *non-classical* statistical methods (Griffith and Layne, 1999) described in further detail in this Section.

Depending on the type of spatial data and level of spatial aggregation, spatial models are commonly classified into three groups: *point processes* (Figure 3.1), *lattice* (Figure 3.2) and *geostatistical* (Figure 3.3). Spatial point processes models (Section 3.2.1) describe finite discrete random processes, in which their elements (i.e. *points*) can be located everywhere in a given domain. Lattice approaches (Section 3.2.3) are used to model continuous or discrete data, where reference locations are defined on discrete spatial features such as *pixels* within a grid feature or irregular polygons such as administrative regions for example (Cressie and Wikle, 2011, p. 167). Finally, geostatistical models (Section 3.2.2) cover random processes continuous in space whose parameters are defined through first and second moments (Cressie and Wikle, 2011, p. 124).

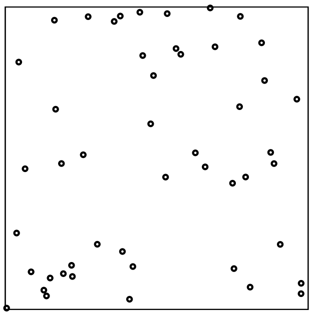


Fig. 3.1 Realisation of a point process

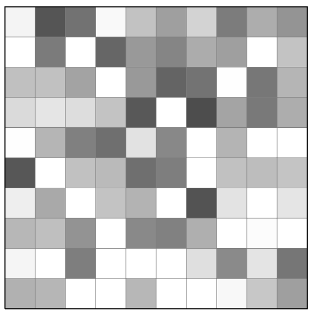


Fig. 3.2 Realisation of a lattice process

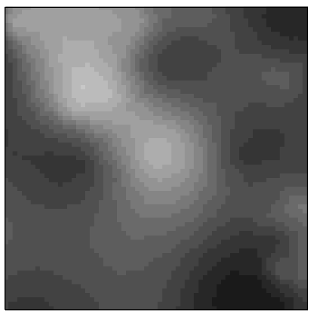


Fig. 3.3 Realisation of a geostatistical process

3.2.1 Point Processes Models

Burglary (Mohler et al., 2011), homicide in a city (Mohler, 2014), or insurgencies in a country (Zammit-Mangion et al., 2012) are few among many examples of phenomena randomly distributed in space which have been elegantly modelled through a point process. A point process is defined as a "stochastic process whose realisations consist of countable set of points" (Diggle, 2014, p. 199). Furthermore, a *point pattern* represents "a collection of points in some area or set and is typically interpreted as a sample from (or realisation of) a point process" (Illian et al., 2008, p. 23). In space, the set of points that characterises a spatial pattern and locations refer to events (Diggle, 2014, p. 1).

Let us denote a point process through a random counting measure N operating on sets. For a given subset $\mathscr{B} \subseteq \mathbb{R}^d$ (with $d = 2$ for two-dimensional processes), a point process can be defined as follows:

Definition 3.1. (**Point process** (Illian et al., 2008, p. 24)) *A point process $X(\mathscr{B})$ is the random number of points in $\mathscr{B}$, i.e. the set $\mathscr{B}$ is assigned the number $X(\mathscr{B})$. It is assumed that $X(\mathscr{B})$ is 'locally finite' ($X(\mathscr{B}) < \infty$) for all bounded sets $\mathscr{B}$. Furthermore, for disjoint subsets $\mathscr{B}_1$ and $\mathscr{B}_2$, and similarly for countably many sets, $X(\mathscr{B})$, which is considered as a function of $\mathscr{B}$, is additive: $X(\mathscr{B}_1 \cup \mathscr{B}_2) = X(\mathscr{B}_1) + X(\mathscr{B}_1)$.*

Hence, one may estimate the probability P that a point process X is in the set $\mathscr{A}$, where $\mathscr{A}$ consists in a set of point patterns with particular characteristics. For example, $\mathscr{A}$ could represent the set of all point patterns without points in the set $\mathscr{B}$. Then, $P(X \in \mathscr{A}) = P(X(\mathscr{B}) = 0)$. Furthermore, similar to classical statistics, the expectation of the number of points X is often of interest. It can be formulated as $\mathbb{E}(X(\mathscr{B}))$. The mean number of points also refer to the concept of *intensity* (Section 3.2.1), often denoted λ (Illian et al., 2008, pp. 27-28)).

The homogeneous and inhomogeneous Poisson Models

There are numerous examples in real situations where points (also called events) appear "haphazardly" distributed, such as the emission of radioactive particles in time (one dimension), the location of trees in a forest (two dimensions), or the spatial distribution of stars observed in a portion of the sky (three dimensions). These phenomena might be modelled by the spatial *homogeneous* Poisson process (SHPP) (Kingman, 1992, pp. 1-3). The number of events is assumed random but finite, events are assumed independent, and the intensity of the process λ is constant over space and equal to its variance. Moreover, λ can be estimated by the observed number of points divided by the study area (Baddeley, 2010, p. 347). This particular form of spatial randomness specific to SHPP is called *complete spatial randomness* (CSR). Hence, *clustering* and its antagonist, *inhibition* phenomena can be assessed as a deviation from CSR (Isham, 2010, p. 289). A point process X that follows a SHPP can be expressed as (Cressie and Wikle, 2011, p. 205):

$$X(\mathscr{B}) \sim \text{Poisson}(\lambda|\mathscr{B}|), \qquad \mathscr{B} \subset \mathscr{D}, \tag{3.12}$$

where λ is a non-negative constant intensity, $|\mathscr{B}|$ is the *cardinality* of the set $\mathscr{B}$, which is a measure of the number of elements in the set $\mathscr{B}$. The spatial domain $\mathscr{D}$ is assumed bounded and $X(\mathscr{B})$ is finite for all subsets $\mathscr{B} \subset \mathscr{D}$. It should be noted that SHPP is also called *spatial whiteness*, for its reference to its randomness properties in space (Zammit-Mangion et al., 2013, p. 17). In most real cases, λ cannot be assumed constant throughout space. Its variability can be described by a spatial *inhomogeneous* (also called *nonhomogeneous*) Poisson process (SIPP) (Cressie and Wikle, 2011, p. 207):

$$X(\mathscr{B}) \sim \text{Poisson}\left(\int_{\mathscr{B}} \lambda(\mathbf{s})\, d\mathbf{s}\right), \qquad \mathscr{B} \subset \mathscr{D}. \tag{3.13}$$

The SIPP is determined by the integral of its intensity function $\lambda(\mathbf{s})$, which may vary depending on its spatial location $\mathbf{s}$ in $\mathscr{B}$.[7] As in the SHPP, SIPP models assume a finite number of independent events (Isham, 2010, p. 290).

Cox Process Models

In contrast to CSR, points may exhibit clusters at different scales, which suggest a positive association between points. In this case, clustered point processes can be modelled through an adaptation of the inhomogeneous Poisson process which allows overdispersion in space since the intensity becomes a random process, which characterises a Cox Process (Isham, 2010, p. 291). The *Log-Gaussian Cox process* (LGCP), a subclass of Cox processes, assumes that the logarithm of the intensity is a GRF (Section 1.3.3) (Baddeley, 2010, pp. 355-356)[8].

Gibbs Process Models

In addition to random or clusters patterns, points may exhibit regularities. *Regular* patterns suggest negative association between points, which may be the result of an *inhibition*

[7] Note that the spatial location ($\mathbf{s}$) is written in bold since one assumes that $\mathbf{s}$ lies within a spatial domain of dimension d, with $d \geq 2$. [8] Other processes (not detailed here) can be derived from Cox processes such as the Poisson-Gamma, Shot-Noise Cox, Poisson-Gamma and some Neyman-Scott processes (including Neyman-Scott special cases, such as Thomas and Matérn Cluster processes) (Cressie and Wikle, 2011, p. 208).

process. In other words, points seem to repel each other. The Gibbs point process with n number of points can be used to model such processes.

$$f(\mathbf{x}) = \frac{\exp\left(-\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \iota\left(\left\|x_i - x_j\right\|\right)\right)}{C_n}, \tag{3.14}$$

where $\mathbf{x} = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ is a point pattern, composed of n points observed in locations within a spatial domain $\mathscr{D}$. The Gibbs point process is defined by the *location density functions* $f_n(x_1, \ldots, x_n)$, which represent a family of multivariate density functions. The term C_n is a normalising constant, which ensures that the Equation 3.14 is a probability density. Gibbs point processes arise in physics, where the so-called *pair potential function* ι may be interpreted as the "potential energy" among points (x_i, x_j), which is expressed as a function of the inter-point distance $\left\|x_i - x_j\right\|$. A repulsive pattern is observed if $\iota(\left\|x_i - x_j\right\|) > 0$. Conversely, points attract each other if $\iota(\left\|x_i - x_j\right\|) < 0$ (Illian et al., 2008, Chap.3).

3.2.2 Geostatistical Models

Continuous spatial processes have been traditionally investigated through methods initially developed in the field of *geostatistics*. Krige (1951), Matérn (1960) and Matheron (1963) pioneered the development of models that take into account spatial dependence. *Geostatistical* models use measures from a finite number of observations $\mathbf{y}(\mathbf{s})$ in n spatial locations $\mathbf{s} = \{\mathbf{s}_1, \ldots \mathbf{s}_n\}$ in a given spatial domain $\mathscr{D}$ of dimension d ($\mathscr{D} \subseteq \mathbb{R}^d$), with d usually equal to 2 or 3. The main aim is usually to predict values at any location within $\mathscr{D}$.

This approach assumes the existence of an underlying random field (RF), most often specified as a Gaussian random field (GRF), a specific class of stochastic processes, which is characterised by Gaussian properties (Section 1.3.3).

Definition 3.2. (**Gaussian random field** (Abrahamsen, 1997, p. 7)) *A Gaussian random field (GRF) is a random field where all the finite-dimensional distributions, $F_{\theta_1,\ldots,\theta_k}$, are multivariate normal distributions for any choice of k and $\{\theta_1, \ldots, \theta_k\}$.*

Note that the GRF is entirely defined by its *mean* and *covariance* function[9]. According to Tobler's law (Section 2.2.1), one generally assumes that locations close to each other are

[9] While it is common to represent the GRF through its covariance function, it can also be specified in the frequency domain (Bochner, 1955) or through a *(semi)variogram* (Matheron, 1971).

more similar, therefore possessing a larger covariance compared to more distant locations. *Spatially independent* observations exhibit a covariance equal or close to zero. A GRF is *stationary* in space if its mean is constant over space and if its covariance depends only on the vector difference between two points of the process (Diggle, 2007, p. 47). Along with its computational and analytical advantages, GRFs offer a high degree of flexibility, which usually allows for a good fit with data (Diggle, 2007, p. 46). Furthermore, GRFs are central to the geostatistical Gaussian process, which can be defined as the following:

Definition 3.3. (**Geostatistical Gaussian process** (Cressie and Wikle, 2011, p. 125)) *Let $Y(\mathbf{s})$ represents a geostatistical Gaussian process in n locations $\mathbf{s} = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ in a spatial domain $\mathscr{D} \subseteq \mathbb{R}^d$, with d usually equal to 2 or 3, $\mathbf{z}(\mathbf{s})$ is a p-dimensional vector of covariates, $\boldsymbol{\beta}$ is a p-dimensional fixed effect (with $p < n$), and $\xi(\boldsymbol{s})$ is a spatial random effect (GRF), defined as a mean-zero Gaussian process. A geostatistical Gaussian process may be defined as*:

$$Y(\mathbf{s}) = \mathbf{z}(\mathbf{s})'\boldsymbol{\beta} + \xi(\mathbf{s}), \qquad \mathbf{s} \subseteq \mathscr{D}. \tag{3.15}$$

More sophisticated Gaussian models have been developed, such as the (Gaussian) spatial moving average (SMA) model, which can be expressed as (Cressie and Wikle, 2011, p. 153):

$$Y(\mathbf{s}) = \int k(\mathbf{s}, \mathbf{u}) W(d\mathbf{u}), \qquad \mathbf{s} \subseteq \mathscr{D}, \tag{3.16}$$

where $k(\cdot,\cdot)$ is a kernel function, $\mathbf{u}$ is a grid of spatial locations on $\mathscr{D}$, $W(\cdot)$ is a Gaussian process with independent increments and the mean $\mathbb{E}(\{W(d\mathbf{u})\}^2) = d\mathbf{u}$ (or equivalently, a d-dimensional *Brownian motion*). The similarity with a moving average (MA) time series model is that both MA and SMA are linear combination of i.i.d. random variables (Cressie and Wikle, 2011, p. 154). Note that SMA can be also applied to multivariate spatial processes and adjusted to accommodate discrete processes (e.g. lattice processes) in turn.

In parallel, Casetti (1972) introduced an expansion of the linear regression model adapted to processes wherein the relationship between variables varies in space, which refers to spatially *non-stationary* processes. Since then, his model has been further developed through the incorporation of a spatial structure (Brunsdon et al., 1998; Fotheringham et al., 2002) and the possibility of using flexible bandwidths[10] (Yang et al., 2012). Hence, a simple *geographically-weighted regression* (GWR) model may be written as the following (Öcal and Yildirim, 2010):

$$Y_i = \alpha_{i0} + \sum_{k=1}^{p} \beta_{ik} z_{ik} + \varepsilon_i, \qquad i = 1, ..., n, \tag{3.17}$$

where Y_i is the dependent variable at location i, z_{ik} is the k^{th} independent explanatory variable at location i, and the i.i.d. error term $\varepsilon_i \sim N(0, \sigma^2)$. There is a total of p variables $(k = 1, ..., p)$ and n locations $(i = 1, ..., n)$. In contrast to the standard ordinary least squares (OLS) regression method, GWR allows for a variation of the intercept term α_{i0} and each regression coefficients β_{ik}, according to each location i. The coefficients are estimated through a kernel weighted regression function. Weights ($\mathbf{w}$) of the kernel function are usually assigned to observations as a decay function, which depends on the distance to the regression point at location i. Usually, $\mathbf{w}$ is a Gaussian function of distance between points at locations i and j and expressed as: $w_{ij} = e^{(\eta d_{ij}^2)}, \quad i = 1, ..., n$. The distance decay parameter η is a positive number. For a given distance d_{ij} between locations i and j, if η is large, it implies a small corresponding weight w_{ij}, in line with Tobler's first law (Öcal and Yildirim, 2010).

3.2.3 Lattice Models

In contrast to geostatistical methods which consider spatially continuous processes observed in specific measurement points, spatial processes can be analysed through discrete grids that cover a spatial domain $\mathscr{D} \subseteq \mathbb{R}^d$ (Bolin, 2012, p. 10). Data are modelled on a lattice, where the measurement points (also called *reference locations*) usually correspond with the centres of the grid-cells. The discrete grid can be either regular (e.g. a grid of identical squares in $\mathbb{R}^2$) or irregular (e.g.regions defined by administrative boundaries in $\mathbb{R}^2$) (Besag, 1974).

A first-order spatial dependence of a lattice process can be modelled through a *Markov random field* (MRF). MRFs assume that the process Y is defined on a d-dimensional lattice $\mathscr{D} \equiv \{\mathbf{s}_1, ..., \mathbf{s}_n\} \subseteq \mathbb{R}^d$. The conditional distribution of Y at the i^{th} location, conditioned on $\mathbf{Y}$ at all other locations ($-i$), depends only on its neighbouring values. MRFs are used to model

[10] A kernel function with a given bandwidth describes the geographical weights which decline according to the distance from a regression point. The bandwidth corresponds to the rate of decay and can be fixed or variable, according to the model used to fit the observations (Yang et al., 2012).

lattice data, where $\mathbf{s}$ are reference locations. As an example, the *conditional autoregressive* (CAR) model is a *Gaussian Markov Random Field* (GMRF) defined on a lattice $\mathscr{D}$[11].

Definition 3.4. (**Conditional autoregressive process** (Cressie and Wikle, 2011, pp. 171-175)) *Assume a sequence of conditional Gaussian distributions* $\{[Y(\mathbf{s}_i)|\mathbf{Y}_{-i}]: \quad i = 1,...,n\}$. *The conditional autoregressive process* $\mathbf{Y}_{-i} \equiv (Y(\mathbf{s}_1),...,Y(\mathbf{s}_{i-1}),Y(\mathbf{s}_{i+1}),...,Y(\mathbf{s}_n))'$ *and* $\mathscr{N}(\mathbf{s}_i)$ *is a set of pre-specified locations that define the neighbourhood of* $\mathbf{s}_i$. *The expectation of* Y *at location* $\mathbf{s}_i$ *is given by*:

$$\mathbb{E}(Y(\mathbf{s}_i)|\mathbf{Y}_{-i}) = \sum_{\mathbf{s}_j \in \mathscr{N}(\mathbf{s}_j)} c_{ij} Y(\mathbf{s}_i), \tag{3.18}$$

In some cases, the value at a specific location might be influenced by a larger order neighbourhood. For this purpose, a *simultaneous autoregressive* (SAR) model can be suitable. Indeed, SAR considers simultaneously specified spatial processes on an irregular lattice $\mathscr{D} \equiv \{\mathbf{s}_1,...,\mathbf{s}_n\}$ and can be expressed as (Cressie and Wikle, 2011, p. 198):

$$Y(\mathbf{s}_i) = \mu(\mathbf{s}_i) + \sum_{j=1}^{n} \beta_{ij}\left(Y\left(\mathbf{s}_j\right) - \mu\left(\mathbf{s}_j\right)\right) + \varepsilon(\mathbf{s}_i), \tag{3.19}$$

where a random spatial noise $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, for $i = 1,...,n$, the coefficient for a same location $\beta_{ii} = 0$; the mean of $\mathbf{Y}$ is the mean at all locations $\boldsymbol{\mu} \equiv (\mu(\mathbf{s}_1),\ldots,\mu(\mathbf{s}_n))'$, and the variance of the spatial noise $\boldsymbol{\Sigma} \equiv \operatorname{diag}\left(\sigma_1^2,\ldots,\sigma_n^2\right)$.

3.2.4 Spatio-Temporal Models

While time can be viewed as an additional dimension to the usual spatial dimensions (e.g. latitude, longitude), time is nevertheless inherently different from space; there is not any (a priori) privileged direction in space, while time ineluctably moves forward. This fundamental distinction needs to be taken into account when space-time processes are under investigation (Gneiting and Guttorp, 2010, p. 427). Similarly to spatial data, spatio-temporal data can be investigated through point process, lattice, or geostatistical approaches. Space and time can be either considered continuous or discrete, according to the nature of the data and the purpose of the research. Thus, three types of spatio-temporal processes can be identified: continuous (space and time are continuous), spatially discrete (space is discrete),

[11] More detail on the GMRF and its applications are provided in Section 3.3.

or temporally discrete (time is discrete) (Diggle, 2014, p. 201). A continuous space-time process Y can be defined as the following:

$$Y(\mathbf{s},t) : (\mathbf{s},t) \in \mathscr{D}, \tag{3.20}$$

where $Y(\mathbf{s},t)$ varies in space and time, with $(\mathbf{s},t) \in \mathscr{D} \subseteq \mathbb{R}^d \times \mathbb{R}$ (Gneiting and Guttorp, 2010, p. 427). A specific case of the continuous space-time process (Equation 3.20) is the Gaussian spatio-temporal process, which is characterised by its mean and covariance. The estimation of the spatio-temporal covariance structure could be computationally challenging and may often require to be assumed decomposable as a product of a pure temporal and a pure spatial (*separable*) covariance function. Chapter 5 will highlight an application of a Gaussian spatio-temporal process assuming a *separable* covariance function to modelling the spatial dynamics of lethal terrorism.

Point processes can be also extended to spatio-temporal domains. They could be considered as a sequence of events that occur in a chronological order $\{(\mathbf{s}_i,t_i) : i = 1,\ldots,n\}$, where $\mathbf{s}$ denotes the spatial location, t denotes time and n is the number of events that occur within a space-time domain $\mathscr{D} \subseteq \mathbb{R}^d \times [0,T]$, where T denotes the largest time in the subset (Diggle and Gabriel, 2010, p. 449)[12]. One example is the spatio-temporal inhomogeneous Poisson process, which takes the same formulation which describes the SIPP (see Equation 3.13), except that time represents an additional dimension. Let define the spatio-temporal region $AT = ST$, with the spatial region S and time interval $T[0,T]$. Hence, $X(AT)$ represents the number of events in AT in an inhomogeneous spatio-temporal Poisson process of intensity $\lambda(\mathbf{s},t)$. Thus, $X(AT) \sim$ Poisson distribution with its mean equals to:

$$\int_0^T \int_A \lambda(\mathbf{s},t)\, d\mathbf{s}dt. \tag{3.21}$$

The integration of the intensity is realised over space A and time T. Moreover, $X(AT)$ represents an independent random sample with a p.d.f. proportional to the intensity $\lambda(\mathbf{s},t)$ (Diggle, 2014, p. 224).

Similar to Poisson processes, Cox processes can be adapted in a space-time framework, where the intensity process $\Lambda(\mathbf{s},t)$ depends on both location $\mathbf{s}$ and time t. Its covariance is, however, stationary if parametrised as $\Lambda(\mathbf{s},t) = \lambda(\mathbf{s},t)R(\mathbf{s},t)$, with the unconditional in-

[12] One usually assumes that point processes generate events in a continuous space and time although other settings can be considered as well: (i) both space and time are discrete; (ii) only space is discrete; (iii) only time is discrete (Diggle and Gabriel, 2010, p. 449).

tensity of the point process $\lambda(\mathbf{s},t)$, and $R(\mathbf{s},t)$ is a stationary process which constrains the point process to be "intensity-reweighted" stationary[13]. The spatio-temporal LGCP process, $\log(\Lambda(\mathbf{s},t))$ is a Gaussian process and defined by its mean and covariance structure (Diggle, 2014, pp. 225-226). In a hierarchical model (HM) framework (Section 1.2.3), the *data model* of the log-Gaussian Cox process is an inhomogeneous Poisson process and the *process model* is a Gaussian process (Cressie and Wikle, 2011, p. 350). Recently, space-time point processes have successfully used in modelling space-time processes, such as crime (Lewis et al., 2012; Rodrigues et al., 2010), conflict (Lewis et al., 2012) and insurgency (Cseke et al., 2015; Zammit-Mangion et al., 2012).

3.3 Modelling Terrorism with SPDE Models

As previously discussed in Section 1.3.1, since the early work of Midlarsky et al. (1980), scholars have attempted to identify and explain the patterns of terrorism in time, space, and space and time. Most research focused on temporal processes through the application of a wide range of auto-regressive models, including VAR (Barros, 2003; Enders and Sandler, 2000), ARMA (Weimann and Brosius, 1988), TAR (Enders and Sandler, 2002, 2005), ARCH/GARCH (Bashir et al., 2013; Bilal et al., 2012). Findings from research investigating spatial and space-time processes of terrorism have been mainly generated by studies carried out at global scale (Braithwaite and Li, 2007) or focused on specific areas, including e.g. the US (Nunn, 2007; Piegorsch et al., 2007), Israel (Berrebi and Lakdawalla, 2007), Iraq (Medina et al., 2011; Siebeneck et al., 2009), or Turkey (Öcal and Yildirim, 2010).

Zammit-Mangion et al. (2013) successfully employed Stochastic Partial Differential Equation (SPDE) models to explain and forecast on a local scale the spatial dynamics of conflicts in Afghanistan. As briefly introduced in Section 1.3.2, SPDE models elegantly incorporate complex space-time structure and dependencies present in the data and allow for non-linearities often observed in spatio-temporal processes. Despite the availability of disaggregated data on terrorism, such flexible models have not been applied to modelling terrorism in space and time. Scholars have failed to provide a rigorous and systematic ap-

[13] In space and time, point processes are defined as *stationary* if the intensity λ of a space-time process is constant and independent of the spatial location and time (Diggle, 2014, p. 202). However, most phenomena are not stationary in space and time, especially in the presence of spatial autocorrelation and/or temporal autocorrelation (Section 2.2.1). Time and space dependencies can be taken into account through conditional probability distributions within dynamical statistical models (Cressie and Wikle, 2011, p. 297).

proach to explain and predict the local space-time patterns of terrorism and this PhD thesis aims at filling this literature gap.

The following Sections provide further detail on concepts related with the SPDE approach, such as the GRF and GMRF, which have been briefly introduced in previous Sections (3.2.2 and 3.2.3). They allow better appraisal of the added-value of the SPDE approach, used as a realistic framework to modelling spatio-temporal processes with complex dependence structures. Section 3.3.1 provides a formal description of the GMRF and further detail on the concept of *spatial contiguity*. Section 3.3.2 draws attention to the main issues of inference methods based on *Markov Chain Monte Carlo* (MCMC) methods. Furthermore, it highlights the advantages of using INLA, which provides fast approximation of the posterior marginals for the class of *latent* Gaussian models (LGM) (Rue et al., 2009). Section 3.3.3 describes in more detail the SPDE models and the main approximation methods used to find a *finite* representation of a solution to the linear SPDE.

3.3.1 GMRF: An Approximation of Continuous Gaussian Fields

As briefly mentioned in Section 3.2, GRFs are continuous and the estimation of their parameters is considered as "big n" problem. Computation is challenging ($\mathscr{O}(n^3)$) since it requires computing the determinant and inverse of dense covariance matrices. Several methods of representations have been developed, such as: spectral representations based on Fourier transform (Fuentes, 2007; Whittle, 1954), likelihood approximation methods (Eidsvik et al., 2014; Stein et al., 2013, 2004), or low-rank methods (using exact computations on a simplified model) (Banerjee et al., 2008; Eidsvik et al., 2012; Furrer et al., 2012; Kaufman et al., 2008)[14].

As an alternative, GMRFs, which are characterised by Markov properties (Equation 3.22) provide an efficient approach to approximate GRFs (Rue et al., 2009). Besag (1974, 1975) pioneered the extension of Markov chains to spatial processes. The author specified a GMRF through the so-called *full conditionals* $\{\pi(\xi_i|\boldsymbol{\xi}_{-i})\}$, which describe the probability density $\pi(\cdot)$ for a random vector $\boldsymbol{\xi}$, which takes a specific value in a given location i (e.g. with i within a spatial domain $\mathscr{D}$) given its values in other locations ($-i$) of the domain. Since the full conditional of spatial random variables can be more easily specified, it also allows a time reduction in its computation (Rue and Held, 2010, p. 175). Even with a

[14] For a review of the current methods to represent GRFs and the advantages of the GMRF approach, see e.g. Bolin and Lindgren (2011).

relative limited neighbourhood, the homogeneous GMRFs efficiently approximated GFs on a lattice. Because of their Markov property, GMRFs can be rapidly and exactly simulated. Moreover, their likelihood can be easily and rapidly computed in comparison with GFs (Rue and Held, 2010, pp. 175-178).

The simplest form of the GMRF is the AR(1) process, which takes us back to the first model introduced in this Chapter (Equation 3.1). More generally, GMRF can be defined in a d-dimensional space as follows:

Definition 3.5. (**Gaussian Markov random field** (Rue and Held, 2005, p. 21)) *Let the points* $\mathbf{s}_j$ *(*$\{j \in \mathscr{N}_i\}$*) be the neighbours* $\mathscr{N}_i$ *to a point* $\mathbf{s}_i$*, with* $\{\mathbf{s}_1, \ldots, \mathbf{s}_i, \ldots, \mathbf{s}_n\} \in \mathscr{D}$*, with domain* $\mathscr{D} \subseteq \mathbb{R}^d$*. A random field* $\boldsymbol{\xi} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *that satisfies the following Markov property is a GMRF*:

$$p\left(\xi_i \,|\, \left\{\xi_j : j \neq i\right\}\right) = p\left(\xi_i \,|\, \left\{\xi_j : j \in \mathscr{N}_i\right\}\right) \quad \forall i. \tag{3.22}$$

Note that different types of neighbourhoods can be defined, according to the assumed extent of the role of the neighbourhood in the process under investigation. Often, one assumes *first-order* neighbourhood, in line with the Markov assumption. In lattice data, the typology of *first-order* contiguity has been inspired by the moves of pieces on a chess board. First-order neighbours may consist in units that share a common edge (Figure 3.4: *rook*), a common vertex (Figure 3.5: *bishop*), or units that share a vertex or an edge (Figure 3.6: *queen*) (Darmofal, 2015, p. 15).

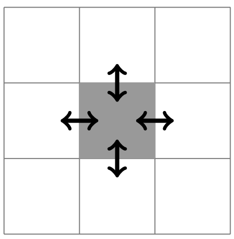


Fig. 3.4 Rook contiguity

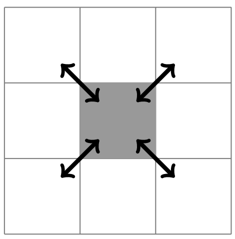


Fig. 3.5 Bishop contiguity

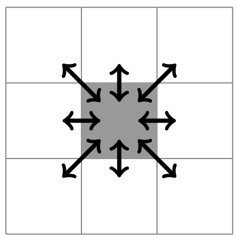


Fig. 3.6 Queen contiguity

Hence, Darmofal (2015, p. 15) indicates that contiguity can be extended to *k-nearest* neighbours, defined through distance measures (*distance band* or *distance-decay*), or defined through *non-spatial* dependence (e.g. based on economic relationship (Beck et al., 2006)). Also, if and only if i and j are not neighbours, it naturally follows that ξ_i and ξ_j are conditionally independent ($\xi_i \perp \xi_j | \boldsymbol{\xi}_{-\{i,j\}}$), or equally, the corresponding element in the precision matrix $\boldsymbol{\Sigma}^{-1}_{i,j} = \mathbf{Q}_{i,j}$ is zero. As a result, the elements in the precision matrix are not zero only when i and j are neighbours, therefore $\mathbf{Q}$ is sparse. Under mild conditions, the Cholesky factorisation $\mathbf{Q} = \mathbf{L}\mathbf{L}^\intercal$ is sparse as well[15]. Assuming a close neighbourhood (i.e. first-order contiguity), the computational cost of factorizing $\mathbf{Q}$ into $\mathbf{L}\mathbf{L}^\intercal$ is reduced from $\mathscr{O}(n^3)$ to $\mathscr{O}(n^{3/2})$ for a two-dimensional GMRF (Lindgren and Rue, 2013; Rue et al., 2009).

3.3.2 Towards a Fast and Accurate Bayesian Inference

Since uncertainty is present in data, processes, or parameters, uncertainty appears in conclusions as well. Hence, *statistical inference* also simply called *inference* can be defined as "drawing of conclusions in the presence of uncertainty". Note that inference may refer to both the *estimation* of unknown parameters or the *prediction* of unknown quantities in the future (Cressie and Wikle, 2011, p. 17). For the sake of clarity, I restrict the usage of the term *forecasting* to prediction in the time dimension only.

In the analysis of spatio-temporal processes, inference is usually highly complex and computationally demanding, which calls for the use of advanced statistical methods. Through the construction of HM, Bayesian models bring a convenient and rigorous mathematical framework, which allows quantifying uncertainty of complex processes. However, Bayesian inference of spatio-temporal models is computationally challenging (Zammit-Mangion et al., 2013, pp.8-16). "Bayesian modelling quantifies uncertainty in predictions by treating all relevant variables as random variables and modelling dependencies as conditional distributions" (Zammit-Mangion et al., 2013, p.17). As a matter of fact, Bayesian inference was limited before the development of computers able to execute computationally intensive sampling methods of estimation for inference purposes (Congdon, 2007, pp. 1-2). Indeed, the estimation of posterior probability distributions related to complex models requires efficient numerical methods of approximation associated with high computational power (Stevens, 2009).

Bayesian inference has been commonly performed through MCMC methods. The contribution of MCMC to the field of statistics is immense and its usage has spread widely through numerous fields of research since its inception. I provide below a brief overview of the main aspect of the MCMC approach. A deep and thorough review would go beyond the scope of this work (for an introduction to MCMC, see e.g. Gilks et al. (1996) or Dellaportas

[15] For a positive semi-definite matrix (i.e. a Hermitian matrix all of whose eigenvalues are non negative) $\mathbf{Q}$, the Cholesky factor is the unique lower triangular matrix $\mathbf{L}$ with strictly positive diagonal elements that satisfy $\mathbf{Q} = \mathbf{L}\mathbf{L}^\intercal$ (Bolin, 2012, p. 12).

and Roberts (2003), and further reading, see e.g. Brooks et al. (2011); Hammersley and Handscomb (1964)). Subsequently, I describe a faster and accurate alternative method, the so-called INLA approach (Section 3.3.2).

Markov chain Monte Carlo

MCMC algorithms are used in most statistical fields, including Bayesian spatial modelling inference. A *Markov chain* is a stochastic process with the Markov property (Equation 3.22). Consider a sequence of random variables (e.g. the parameters to be estimated in a given model) $\boldsymbol{\theta} = \{\theta^{(0)}, \theta^{(1)}, \ldots, \theta^{(i)}, \ldots, \theta^{(n)}\}$, defined on the state-space S. The idea behind MCMC is to simulate a Markov chain by generating a new state of the chain $\theta^{(i+1)}$, which depends only on the previous state $\theta^{(i)}$, through a transition kernel $\mathscr{K}$, which uniquely describes the changes in the state of the Markov chain:

$$\theta^{(n+1)} \sim \mathscr{K}(\theta^{(n)}, \theta^{(n+1)}). \tag{3.23}$$

Hence, each element of the sequence depends only upon the last. For example, $\theta^{(2)}$ depends exclusively on $\theta^{(1)}$. Assuming that the chain is *aperiodic* (i.e. the exploration of the state space by the chain is not cyclic), *irreducible* (i.e. the probability of reaching any location in the state space S is greater than zero independently of the choice of the initial value $\theta^{(0)}$) and *recurrent* (i.e. the expected number of visits to a given state $S_i \subseteq S$ is infinite); the distribution eventually converges after a number of steps to an equilibrium (*stationary* distribution) (Blangiardo and Cameletti, 2015, pp. 90-91). The states are updated in a relatively simple manner and MCMC can accommodate complex posterior distributions with a high number of parameters (e.g. in HMs). The first MCMC method was provided by Metropolis et al. (1953), which has been later generalised by Hastings (1970) and became known as the *Metropolis-Hastings* algorithm. Since then, numerous MCMC methods have been developed, which include, in particular, the Gibbs sampler (Geman and Geman, 1984)[16].

Since the 1980s, MCMC has been extensively used as a very flexible and universal approach. The performance of MCMC algorithms is however drastically reduced in the presence of spatial dependence structures, which are often complex (Dellaportas and Roberts, 2003, pp. 1-2). In order to reduce errors, a large number of iterations is often required which

[16] A detailed description of the currently available MCMC methods would go far beyond the scope of this present Chapter. For further details on the Gibbs sampler, see e.g. Casella and George (1992). The Metropolis-Hastings algorithm is described in e.g. Chib and Greenberg (1995).

leads to high computational time (Eidsvik et al., 2012). This limits the level of complexity of the models that may realistically be fitted. As a remedy, alternative techniques have been developed, including INLA, used for Bayesian inference for models characterised by a *latent* Gaussian field (Section 3.3.2).

The Integrated Nested Laplace Approximation

INLA was developed as a computationally efficient alternative to MCMC for Bayesian inference (Lindgren and Rue, 2013; Rue et al., 2009). INLA avoids highly computationally demanding MCMC runs by approximating the distributions and speeding up parameter estimation substantially, making the fitting of complex spatial dependence structures practically feasible. INLA has proved to be more accurate and faster than MCMC approaches (Rue et al., 2009).

In combination with the SPDE approach (Section 3.3.3), INLA can handle stationary and non-stationary spatial and spatio-temporal models, as reflected in its successful applications in a wide range of research fields, including epidemiology, risk assessment, and geostatistics (Lindgren and Rue, 2013). For example, Cameletti et al. (2013b) used the combination of the SPDE approach and INLA to model and forecast an atmospheric pollutant in the Piemonte region, Italy.

In order to fully appreciate the added value brought by INLA, it is worthwhile to dwell on the concept of the *latent* Gaussian model (LGM). Assume a classical multiple linear regression model with: intercept α, linear coefficient β_j of covariates z_j to be estimated such that the mean $\boldsymbol{\mu}$ of the vector of the observations $\mathbf{y}$ is given by:

$$\mu_i = \mathbb{E}(Y_i) = \alpha + \sum_{j=1}^{n_\beta} \beta_j z_{ji}, \quad i = 1, \ldots, n. \tag{3.24}$$

In a Generalized Linear Model (GLM) framework, the mean $\boldsymbol{\mu}$ is linked to the linear predictor η_i with the link function $g(\cdot)$, so that:

$$\eta_i = g(\mu_i) = \alpha + \sum_{j=1}^{n_\beta} \beta_j z_{ji}, \quad i = 1, \ldots, n. \tag{3.25}$$

In a Generalized Additive Model (GAM) framework, the mean $\boldsymbol{\mu}$ is linked to the linear predictor η_i, the link function $g(\cdot)$, non-linear smooth effects $f_k(\cdot)$ of covariates c_k, so that:

$$\eta_i = g(\mu_i) = \alpha + \sum_{k=1}^{n_f} f_k(c_{ki}), \quad i = 1, \ldots, n. \tag{3.26}$$

The unknown function $f_k(\cdot)$ can be of various forms, which provides a tool flexible enough to include more specific variables, such as spatially indexed covariates. If we combine GLM and GAM (Equations 3.25 and 3.26), and add i.i.d. random effects ε_i, we obtain the following equation:

$$\eta_i = g(\mu_i) = \alpha + \sum_{j=1}^{n_\beta} \beta_j z_{ji} + \sum_{k=1}^{n_f} f_k(c_{ki}) + \varepsilon_i, \quad i = 1, \ldots, n. \tag{3.27}$$

All parameters (which are considered as random variables in a Bayesian framework) present in Equation 3.27 define the so-called latent field $\boldsymbol{\xi} = \{\alpha, \boldsymbol{\beta}, f_k(\cdot), \boldsymbol{\eta}\}$. Moreover, $\boldsymbol{\xi}$ is a Gaussian latent field if Gaussian priors are assigned to all elements of $\boldsymbol{\xi}$. The class of model using Gaussian latent fields is LGM. This includes a wide range of models: regression models (e.g. GLM, GAM), dynamic models (e.g. autoregressive time series), spatial and space-time models (e.g. Besag-York-Mollié, continuous index Gaussian models) (Rue et al., 2009).

Numerical methods are used to estimate the *posterior marginal distributions* (also called *marginals*[17]) $\pi(\xi_i|\mathbf{y})$ of each element of the Gaussian latent field $\boldsymbol{\xi}$ and the hyperparameters $\boldsymbol{\varpi}$[18]. Hence, summary statistics (e.g. the mean, standard deviation, and credible intervals) can be computed (Sørbye, 2014).

Based on numerical integration, INLA uses three steps to approximate the posterior distribution of the latent Gaussian field $\pi(\xi_i|\mathbf{y})$, with $, i = 1, \ldots, d$ elements. The first step uses Laplace approximation to approximate the posterior marginal of the hyperparameters

[17] The *marginal* distribution of a subset of a collection of random variables is the probability distribution of the variables contained in the subset. It gives the probabilities of various values of the variables in the subset without reference to the values of the other variables. This contrasts with the *conditional* distribution, which gives the probabilities contingent upon the values of the other variables (Gordon, 2015). [18] Recall that in Bayesian statistics, a hyperparameter refers to a parameter of a prior distribution (Section 1.2.3). Furthermore, note that the hyperparameters are not necessarily Gaussian.

$\pi(\boldsymbol{\varpi}|\mathbf{y})$:

$$\tilde{\pi}(\boldsymbol{\varpi}|\mathbf{y}) \propto \left.\frac{\pi(\boldsymbol{\xi},\boldsymbol{\varpi},\mathbf{y})}{\tilde{\pi}_G(\boldsymbol{\xi}|\boldsymbol{\varpi},\mathbf{y})}\right|_{\xi=\xi^*(\boldsymbol{\varpi})} \tag{3.28}$$

where $\tilde{\pi}_G(\boldsymbol{\xi}|\boldsymbol{\varpi},\mathbf{y})$ is the Gaussian approximation to the full conditional distribution of $\boldsymbol{\xi}$, and $\boldsymbol{\xi}^*(\boldsymbol{\varpi})$ is the mode of the full conditional distribution for $\boldsymbol{\xi}$ for a given $\boldsymbol{\varpi}$[19]. The normalising constant (see denominator in the right-hand side of Equation 1.1) is not known, which justifies the proportionality symbol $\propto$ used in Equation 3.28 (Rue et al., 2009). This first step integrates out the uncertainty in the approximation of the posterior marginal of ξ_i with respect to $\boldsymbol{\psi}$.

The second step computes the Laplace approximation (as one among other possible approximations, e.g. the Gaussian or simplified Laplace) of the posterior marginal $\tilde{\pi}_{LA}(\xi_i|\boldsymbol{\varpi},\mathbf{y})$, for selected values of $\boldsymbol{\varpi}$. The Laplace approximation is expressed as:

$$\tilde{\pi}_{LA}(\xi_i|\boldsymbol{\varpi},\mathbf{y}) \propto \left.\frac{\pi(\boldsymbol{\xi},\boldsymbol{\varpi},\mathbf{y})}{\tilde{\pi}_{GG}(\boldsymbol{\xi}_{-i}|\xi_i,\boldsymbol{\varpi},\mathbf{y})}\right|_{\boldsymbol{\xi}_{-i}=\boldsymbol{\xi}^*_{-i}(\xi_i,\boldsymbol{\varpi})} \tag{3.29}$$

Here, $\tilde{\pi}_{GG}$ is the Gaussian approximation to $\boldsymbol{\xi}_{-i}|\xi_i,\boldsymbol{\varpi},\mathbf{y}$ and $\boldsymbol{\xi}^*_{-i}(\xi_i,\boldsymbol{\varpi})$ is the modal configuration, and $-i$ designate all components except i. Note that $\tilde{\pi}_{GG}$ (Equation 3.29) is not identical to $\tilde{\pi}_G$ (Equation 3.28). According to Equation 3.29, $\tilde{\pi}_{GG}$ is simulated for each value of ξ_i and $\boldsymbol{\varpi}$, since its precision matrix depends on ξ_i and $\boldsymbol{\varpi}$. In order to reduce the computational time to compute $\tilde{\pi}_{GG}(\boldsymbol{\xi}_{-i}|\xi_i,\boldsymbol{\varpi},\mathbf{y})$ and make the approximation feasible, Rue et al. approximate the modal configuration, as the following:

$$\boldsymbol{\xi}^*_{-i}(\xi_i,\boldsymbol{\varpi}) \approx \mathop{\mathbb{E}}_{\tilde{\pi}_G}(\boldsymbol{\xi}_{-i}|\xi_i). \tag{3.30}$$

Here, the Gaussian approximation $\tilde{\pi}_G(\boldsymbol{\xi}|\boldsymbol{\varpi},\mathbf{y})$ (Equation 3.28) is used to evaluate Equation 3.29. Moreover, Rue et al. save computational time by applying Tobler's law (Section 2.2.1) in a n-dimensional space problem, i.e. only the ξ_j in the neighbourhood $\mathcal{N}_i$ of ξ_i have an effect on the marginal of ξ_i. Through their approach, Rue et al. provide fast and accurate approximations of posterior marginals in a wide range of latent Gaussian models, including spatial models. This thesis employs the R implementation of INLA (`R-INLA`)

[19] As noticed by Rasmussen in the discussion of Rue et al.'s paper, the mode may not be typical of non-Gaussian or non-symmetric posteriors. distribution. Therefore, the use of Laplace approximation in non-Gaussian latent field might be problematic.

and further demonstrates the accuracy and rapidity of the inference on complex space-time models using the Stochastic Partial Differential Equation (SPDE) approach provided by Lindgren et al. (2011).

3.3.3 Approximation of the SPDE Approach

In 2011, Lindgren et al. proposed an approach based on SPDE models, which provides a spatially consistent method for efficiently approximating continuous GF using the sparse properties of the GMRF (Bolin, 2012, p. 11). The authors successfully demonstrated that their approach provides accurate inference while reducing the computation time by the square-root compared to a standard approximation of GFs in a two-dimensional problem. As briefly mentioned in Section 1.3.3, SPDE models are not bound by the use of regular grid-cells and can define GRFs on complex spatial objects such as manifolds (e.g. sphere) (Lindgren et al., 2011), which have a direct application in modelling the spatial dynamics of lethal terrorism worldwide (see Chapter 5). Moreover, the approach can handle non-stationary spatial models as well (not discussed here, see Bolin and Lindgren (2011) for further detail) and explanatory variables can be easily included in the dependence structure, which leads to non-stationary second-order behaviour (Ingebrigtsen et al., 2014). The great flexibility of SPDE models allows for modelling data with highly complex spatio-temporal dependence structures[20].

The linear fractional SPDE — considered in the SPDE approach (Lindgren et al., 2011) — can be formulated as:

$$(\kappa^2 - \Delta)^{\alpha/2}\tau(\xi(s)) = \varepsilon(s), \quad s \in \mathscr{D}\,, \tag{3.31}$$

with the Laplacian $\Delta = \sum_{i=1}^{d} \partial^2/\partial \xi_i^{\,2}$, smoothness parameter $\alpha = \nu + 1$ (for two-dimensional processes), variance parameter τ, scale parameter $\kappa > 0$, domain $\mathscr{D}$, and Gaussian spatial white noise $\varepsilon(s)$. The stationary solution of Equation (3.31) is the GF ($\xi(s)$) with Matérn covariance function:

$$Cov(\xi(\mathbf{s}_i), \xi(\mathbf{s}_j)) = \sigma^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\kappa \left\|\mathbf{s}_i - \mathbf{s}_j\right\|\right)^{\nu} K_\nu\left(\kappa \left\|\mathbf{s}_i - \mathbf{s}_j\right\|\right), \tag{3.32}$$

[20] For a review of second-order non-stationarity models for univariate geostatistical data, see e.g. Fouedjio (2016).

where $\|\mathbf{s}_i - \mathbf{s}_j\|$ is the Euclidean distance between two locations, σ^2 is the marginal variance, $\Gamma(\cdot)$ is the Gamma function[21] and K_ν is the modified Bessel function of the second kind and order $\nu > 0$. Usually, the smoothness parameter is fixed due to poor identifiability, namely $\nu = \alpha - d/2 = 1$, with $d = 2$ for a two-dimensional process and $\alpha = 2$ (`R-INLA` default value). For $d = 2$, the marginal variance $\sigma^2 = \Gamma(\nu)/\Gamma(\alpha)4\pi\kappa^{2\nu}\tau^2$. The distance from which the spatial correlation becomes negligible (for $\nu > 0.5$) is given by the range r, which can be empirically derived from the scale parameter $r = \sqrt{8\nu}/\kappa$ to be estimated (Lindgren et al., 2011).

As mathematical objects, SPDEs and their solutions are of *infinite* dimension. The approximation method suggested by Lindgren et al. (2011) is based on the fact that a solution of the linear SPDE (Equation 3.31) is a Gaussian field (GF), defined by its mean and Matérn covariance (Equation 3.32), as initially noted by Whittle (1963). A *finite* representation of a solution to the linear SPDE can be obtained through a GMRF representation of the Matérn field. Lindgren et al. (2011) suggest the use of a finite element approximation of the solution to the SPDE, using basis functions defined on a triangulation (Constrained Refined Delaunay Triangulation, also called "mesh") that covers the region of interest or domain $\mathscr{D}$. The mesh allows the construction of smaller triangles in specific regions of interest in $\mathscr{D}$, whilst using larger triangles in other areas of $\mathscr{D}$, which reduces the number of vertices, and hence, the computational time required to fit the model.

The triangulation of $\mathscr{D}$ most strongly implies the use of the standard finite element basis of piecewise linear function as method of approximation[22], which reduces the spatial dimensionality by decomposing the Matérn field in a series of basis functions. First, the inner product is defined as follows:

$$\langle f, g \rangle = \int f(\mathbf{s})g(\mathbf{s})d\mathbf{s} \tag{3.33}$$

where the integral is taken over the region of interest. Hence, the procedure requires the stochastic weak formulation of the SPDE (Equation 3.34):

$$\left\{ \left\langle \phi_j, (\kappa^2 - \Delta)^{\alpha/2}\xi \right\rangle, j = 1, \ldots, n \right\} \stackrel{d}{=} \{\langle \phi_j, \varepsilon \rangle \; j = 1, \ldots, n\} \tag{3.34}$$

[21] The Gamma function is defined as $\int_0^\infty \exp(-b)b^{a-1}db$. Note that if a is a positive integer, the Gamma function can be viewed as an extension of the factorial function, with its argument shifted down by 1: $\Gamma(a) = (a-1)!$ (Davis, 1959). [22] Other more complex basis functions might be used. However, Bolin (2012, p. 23) showed that no gain is obtained unless the range of the covariance is very large.

to hold for every appropriate set of *testing functions* $\{\phi_j(\mathbf{s}),\, j = 1, \ldots, n\}$, where $\stackrel{d}{=}$ indicates equality in distribution. Second, a finite element representation of the solution to the SPDE is constructed on a mesh with n the number of vertices, so that:

$$\xi(\mathbf{s}) = \sum_{k=1}^{n} \psi_k(\mathbf{s}) w_k \tag{3.35}$$

for some chosen basis functions ψ_k and Gaussian distributed weights w_k. The chosen ψ_k are piecewise linear in each triangle, with $\psi_k = 1$ at vertex k and zero at all other vertices. The final step of the approximation consists in finding the distribution of w_k that satisfies Equation 3.34 for only a specific set of *testing functions*, the choice of which, controls the characteristics of the approximation. Lindgren et al. use $\phi_k = (\kappa^2 - \Delta)^{1/2}\psi_k$ for $\alpha = 1$ (*least squares* approximation) and $\phi_k = \psi_k$ for $\alpha = 2$ (*Galerkin* approximation)[23]. Define the $n \times n$ matrices $\mathbf{C}$, $\mathbf{G}$, and $\mathbf{K}$ with entries $C_{ij} = \langle \psi_i, \psi_j \rangle$, $G_{ij} = \langle \nabla\psi_i, \nabla\psi_j \rangle$, and $K_{ij}(\kappa^2) = \kappa^2 C_{ij} + G_{ij}$. Hence, the precision matrix $\mathbf{Q}_\alpha(\kappa^2)$ for the Gaussian weights $\mathbf{w}$ (defined in Equation 3.35) is defined as:

$$\mathbf{Q}_1(\kappa^2) = \mathbf{K}(\kappa^2) \tag{3.36}$$

$$\mathbf{Q}_2(\kappa^2) = \mathbf{K}(\kappa^2)\mathbf{C}^{-1}\mathbf{K}(\kappa^2) \tag{3.37}$$

It results that $\mathbf{C}$ and $\mathbf{G}$ are easy to compute, since non-zero elements are present only for pairs of basis functions which share common triangles and their values do not depend on κ^2. Despite that $\mathbf{C}^{-1}$ is dense — and therefore $\mathbf{Q}$ is dense — , the authors show that it can be replaced by the sparse diagonal matrix $\tilde{\mathbf{C}}^{-1}$, with entries $\langle \psi_i, 1 \rangle$. The authors therefore provide an explicit link between latent Gaussian fields and Gaussian Markov random fields using an approximate weak solution of the corresponding SPDE. Moreover, the SPDE approach can be combined with efficient inference methods such as the INLA approach (Section 3.3.2), which requires the latent field to be a GMRF (Lindgren et al., 2011).

---

[23] For the sake of simplicity, I present only the results for $\alpha = 1, 2$. A slightly modified approach may be used for $\alpha \geq 3$ as well. However, non-integer values of α cannot be used. For further details, see Lindgren et al. (2011).

3.4 Conclusion

This chapter highlighted the development of temporal, spatial, and spatio-temporal models that have been used to explain terrorism and related phenomena. First, it allowed the reader to gradually become familiar with the concepts related to statistical modelling. Second, it described the main shortcomings present in each method, which provided elements to better appreciate the improvements made by more complex approaches.

The analysis of terrorism in time has provided valuable insight into the understanding of its patterns and its dynamics. I have introduced several auto-regressive models, such as AR, ARMA, ARIMA, and ARCH. I have also highlighted the usefulness of count data models widely used in the study of terrorism, since the dependent variable is rarely continuous; rather it often represents a count of e.g. the number of terrorist events, fatalities, etc.

Inherently, temporal analysis does not take into account an important aspect of most (if not all) social phenomena, including terrorism: the *spatial* dimension. Since terrorist events occur in space, spatial dependence is likely and needs to be taken into account for proper inference. I have therefore briefly discussed different approaches that have been used to model spatial processes, including point process, geostatistical, and lattice approaches.

Geostatistical models are suitable to model spatially continuous processes. The Bayesian framework — especially deployed through hierarchical models —, has proved to be a suitable approach to tackle issues due to these complex dependence structures. However, computational issues often stem from the complexity of the spatial dependence structure present in the data, which requires in particular computing the inverse and determinant of large dense matrices. Traditional MCMC approaches are usually too slow to fit complex models, which calls for the use or more efficient approaches, such as INLA. I have concluded the Chapter by introducing the SPDE approach, developed by Rue et al. (2009) and Lindgren et al. (2011), which provides an efficient alternative to MCMC.

By combining the SPDE approach with INLA, accurate and fast inference has been made possible for a wide class of models, including those applied in this present study. Although the availability of accurate and fast inference techniques are a necessary condition, it does not suffice to build models able to achieve the objectives of this present study. Extracting knowledge of the data is a crucial pre-requisite, which should not be discounted. Therefore, I devote the next Chapter to the exploration of data on terrorism and covariates, prior to implementing the SPDE approach and analysing the results described in further details in Chapters 5 and 6.

Chapter 4

Exploratory Data Analysis of Terrorism

It is well recognised that in-depth exploration of data is a crucial step in statistical analysis (Tukey, 1962, 1980). In particular, exploratory data analysis permits the proper assessment of the data required to answer the research questions of this study. This present work calls for the use of high-resolution data on terrorist events over several years, as well as key driving factors of terrorism. The suitability of the data cannot be guaranteed before having carried out meticulous quality control.

In addition to revealing potential issues in the data, the exploration of data might help finding patterns, which allows a better understanding of the underlying phenomenon and provides crucial knowledge used in the modelling process. Therefore, the main aim of this present Chapter is: (i) to assess the quality of data; (ii) to select the most suitable data which will be used in the modelling process; (iii) to capture the trends and patterns of terrorism in space and in time.

This Chapter is divided into five Sections. Section 4.1 introduces the currently available databases on worldwide geolocalised terrorism. It describes the issues specific to terrorism data and compares the temporal and spatial patterns of terrorist events according to their data source. Section 4.2 explores the temporal patterns of terrorism according to different databases. In Section 4.3, the *spatial structure* of terrorist events is analysed through *global* (Moran's I index) and *local* (Getis and Ord $G^*_{\mathbf{s}_i}$ index) measures of spatial autocorrelation. Furthermore, the analysis of the point patterns through the pair correlation function (*pcf*) function reveals clusters present at various scales in the data. Section 4.4 concludes the Chapter by providing a summary of the results of the data exploration.

4.1 Databases on Worldwide Terrorism

4.1.1 Database Providers

Currently, there are three main sources of geolocalised data on terrorism[1] that cover the entire world (our study area): the *Global Terrorism Database* (GTD), the *RAND Database of Worldwide Terrorism Incidents* (RDWTI) , and the *Global Database of Events, Language, and Tone* (GDELT). Although the *International Terrorism: Attributes of Terrorist Events* (ITERATE) has been extensively referred to in terrorism research, terrorist events are not geolocalised (neither city name nor geographic coordinates are provided), which does not allow subnational-level analyses. Moreover, it provides data on transnational terrorism only, excluding domestic terrorism (Enders et al., 2011)[2].

In addition to the geolocalisation of terrorist events, the aforementioned databases provide a different range of complementary variables. Some databases provide the number of injuries along with the number of fatalities. However, the number of injuries will not be incorporated in this study because of its low measurement accuracy, which is mainly due to challenges posed by distinguishing injured from uninjured people, whose reports may vary within and across different countries. In line with De la Calle and Sánchez-Cuenca (2011), I agree that reports on fatalities are usually more precise than those on injuries, and therefore better for comparison purposes.

GTD is a free-access database, which includes more than 125,000 georeferenced terrorist events in 209 countries from 1970 to 2013. GTD provides 132 variables, including the date, the types of attack and target, and the number of fatalities due to terrorist incidents (START, 2014). RDWTI counts approximately 40,129 georeferenced terrorist events from 1969 to 2009 and provides 8 variables (RAND, 2011). GDELT is updated every day and covers more than 200 million events from 1979 to now as well as incorporating data about several types of political violence, including terrorism (Leetaru and Schrodt, 2013).

[1] The analysis carried out in this Chapter and the following ones is based on data on terrorist attacks perpetrated by non-state actors exclusively, which refers to *non-state* terrorism (Section 2.1.1). State terrorism is therefore excluded. In the following lines, the word *terrorism* should always be understood as *non-state terrorism*. [2] ITERATE is not a free-access database (Vinyard Sotfware, 2008) but the members of Duke University (http://library.duke.edu/data/collections/iterate.html) or Cornell University (http://ciser.cornell.edu/ASPs/search_athena.asp?IDTITLE=2340) may access ITERATE free of charge.

4.1.2 Issues of Terrorism Databases

Data on terrorism are subject to several considerable issues. First, the notion of terrorism is inherently subjective and no consensus about its definition has been reached (Section 2.1.1), which implies that data providers use definitions that are inherently subject to debate. Second, gathering data on all possible forms of terrorism is virtually impossible. For example, most databases do not include state terrorism, planned but failed, and/or cancelled attacks (Kyung et al., 2011). Some providers distinguish the types of terrorism (e.g. domestic and international), while others do not. Consequently, terrorism data inevitably stem from a narrow and subjective interpretation of its provider.

Second, measurement errors related to the localisation of the events in space and/or in time could be attributed to the quality of the source of information and the coding method, whose quality differs among the databases. Kyung et al. (2011) added that "[t]he key problem is that humans in covert, especially dedicated terrorist networks, for strategic reasons work to conceal not just their identities and intentions, but also their interactions with others. As a result of these collection issues, the data [on terrorism] contain confounding effects, overlapping explanatory variables, high measurement error, and unmeasured clustering forces".

Third, terrorism data suffer from under-reporting issues. Drakos and Gofas (2006a) and Drakos (2007) provided evidence that databases understate the true number of terrorist incidents, as they include only events available from open sources or reported by the media, as noticed by Sheehan (2012). Additionally, the number of covered events appears to be correlated with country characteristics, such as government control of the media. On a study including data from 153 countries in the period 1985-1998, Drakos (2007) estimated an overall under-reporting bias of 19% [3].

As a general remark, it is worth noticing that terrorism does not necessarily mean death, as one might erroneously believe. From 2002 to 2009, the proportion of reported non-lethal attacks is approximately 45% and 52% in GTD and RDWTI, respectively. During the same period, the average number of deaths (with standard deviation in parentheses) per terrorist event is approximately 3.1 (9.8) in GTD and 1.9 (7.6) in RDWTI. A relatively high standard deviation suggests a high variation of the number of fatalities around the average, illustrated by the presence of peaks due to highly lethal events, also called *mass-casualty* events.

[3] In order to address this issue, the authors proposed the use of a statistical regression to approximate the true number of transnational terrorist events, which is adjusted in function of the freedom of press and polity for each studied country.

For example, in January 2004, 334 people were killed during a hostage taking, which took place in the School of Beslan, North Ossetia, Russia (Tuathail, 2009). Moreover, on March 11, 2004, 191 people were killed in Madrid (GTD and RDWTI recorded the same values) by an Islamic terrorist group who bombed commuter trains. The event is also known as *M-11* (Los Angeles Times, 2014). The peaks of terrorism lethality in 2006 and 2007 are mainly due to terrorist events that occurred during and after the civil war in Iraq (2006-2007) (Boyle, 2009).

4.1.3 Key Characteristics of Terrorism Databases

Even though terrorism databases ought to provide identical data — given that they collect data on the same phenomenon — important discrepancies occur owing to the databases' peculiarities: distinctive collection methodology, data sources and scope. Through the use of exploratory techniques, variations among the databases can be identified, compared, and analysed. The first operational step is to collect and aggregate data from GTD, GDELT, and RDWTI into a common framework (`R` code in Appendix A). Table 4.1 summarises the main characteristics of the databases[4].

Unlike the other databases, GDELT gathers not only data on terrorism but also on various social events, using a fully automated coding system based on Conflict and Mediation Event Observations (CAMEO)[5], which may lead to a strong geographic bias, as mentioned by Hammond and Weidmann (2014). Hammond and Weidmann suggested waiting for a more reliable coding method before using GDELT data[6]. Terrorist events may appear in various categories and the extraction of data (`R` code in Appendix A.3) does not eliminate *noisy* data (terrorist events which are not accurately geolocalised or misspecified events) or missing some terrorist events.

The number of available variables differs between databases. GTD has the largest number of variables (132), while RDWTI provides only eight variables. Despite its large number of variables, numerous missing values are present in GTD. In contrast to GDELT, GTD and RDWTI provide the number of injuries and fatalities, distinguish domestic from international terrorist events, and use clear definitions of terrorism, from which data is col-

[4] For an in-depth qualitative comparison of several terrorism databases, including GTD and RDWTI, see e.g. Sheehan (2012). [5] For further information on CAMEO, see: http://data.gdeltproject.org/documentation/CAMEO.Manual.1.1b3.pdf. [6] Note that the provider mentions that in the future, the update of data in GDELT will be available on a 15 minute frequency (http://gdeltproject.org).

Table 4.1 Databases on terrorism (GTD, GDELT, and RDWTI): main characteristics

Main Characteristics	GTD	GDELT	RDWTI
Nb. events*	125,087	35,050	40,129
Geolocalised events**	77	100	88
Time min	1970	1979	1968
Time max	2013	now	2009
Update	yearly	daily	n/a
Nb. variables	132	57	8
Nb. killed	yes	no	yes
Nb. injuries	yes	no	yes
Coverage	world	world	world
Coding	human	machine	human
City	yes	yes	yes
Coordinates	yes	yes	no
Level of spatial accuracy	yes	no	no
Definition(s) of terrorism	yes	no	yes
Domestic	yes	yes	yes
Transnational	yes	yes	yes
Distinction	yes	no	yes***

* The calculation is based on events within the study area (the world) in the following periods: 1970 to 2013 (GTD); 1979 to 2011 (GDLET); 1968 to 2009 (RDWTI).

** The percentage of geolocalised events is derived as the following: events with `specificity<3` and events with `specificity=0` and `vicinity=1` whose coordinates are present or found through Google Earth™ (GTD); all events are geolocalised (GDELT); events labelled with a city and country and whose coordinates have been found through Google Earth™ (RDWTI).

*** The distinction between transnational/domestic events is possible only after 1997, when RDWTI started to include domestic events in addition to international events.

lected. GTD reports the highest net number of events (125,087). Moreover, GTD is the only database providing the level of spatial accuracy of each event, through the variables called `specificity` and `vicinity`[7]. This information is crucial for the purpose of this present study, which requires for and foremost spatially accurate data.

4.2 Exploring the Temporal Patterns of Terrorism

4.2.1 Reinforcement Effect

Terrorism data exhibit temporal dependencies, which suggest that values observed at a given time depend on previous values (Enders and Sandler, 1993, 2000; Hamilton and Hamilton, 1983). This is an important information for the analysis of the risk of terrorism. One might know if e.g. peaks of terrorism are followed by further intense terrorist activity or in contrast by a decrease in terrorism, or alternatively, events are i.i.d. and do not influence each other (also-called "truly random" events) in time. As suggested by research work based on different datasets (Enders and Sandler, 1999; Gleason, 1980; Heyman and Mickolus, 1980; Holden, 1986; LaFree et al., 2012; Midlarsky et al., 1980), the so-called *reinforcement* processes (Section 2.2.1) are likely at stake, which materialise into *clusters* in time.

In order to assess any potential deviation from a truly random process, I use the *Bartels Ratio test for randomness*, based on the ranked version of von Neumann's ratio (Bartels, 1982). The samples include one observation per month, which consist in a total of 96 observations (for 8 years) for each database. The results of the test suggest to reject the hypothesis that the number of monthly terrorist events is truly random, displaying different levels of significance: GTD ($p < 0.01$), GDELT ($p < 0.05$), and RDWTI ($p < 0.1$). Likewise, the results of the test suggest to reject the hypothesis that the number of monthly fatalities is truly random: GTD ($p < 0.05$) and RDWTI ($p < 0.01$)[8]. In line with the literature, reinforcement processes in both the monthly number of events and fatalities cannot be excluded.

[7] GTD provides five levels of spatial accuracy through the variable `specificity`: events occurred in city/village/town (`1`); events occurred in city/village/town but no latitude or longitude have been found (the coordinates correspond to the centroid of the smallest subnational administrative region) (`2`). Levels `3`, `4`, and `5` include events which are inaccurately reported in space and are therefore excluded from this analysis. Moreover, the variable `vicinity=1` indicates that the event occurred in the immediate vicinity of the city (GTD, 2014). [8] Note that GDELT is excluded, since it does not provide information about the number of deaths (`R` code in Appendix B.2)

Moreover, one might reasonably expect that the monthly number of events is positively correlated among the databases, which would suggest some forms of consistency among them. The correlation between the number of terrorist events in RDWTI and GDELT is negative and highly significant ($r = -0.46$; $p < 0.01$). On the contrary, the correlation between the observations in GTD and RDWTI is close to 0 and not significant ($r = 0.02$; $p = 0.82$). The absence of positive correlation suggests that the databases use different data collection methods and/or sources of information. Only the pair GDELT-GTD is positively correlated ($r = 0.82$; $p < 0.01$)[9]. A significant positive correlation does not guarantee consistency, however. GDELT uses an automatic procedure to collect data, while data in GTD is manually validated by experts through several processes, including cross-referencing of multiple sources (Section 2.1.2). The monthly number of fatalities is however significant and positively correlated between RDWTI and GTD ($r = 0.75$; $p < 0.001$).

Despite the differences, the databases agree that both the monthly number of events and fatalities are temporally auto-correlated. In the time dimension, terrorism does not appear to be the result of a purely random process, at least on a monthly level of analysis. Hence, patterns of terrorism can be identified and further explored and analysed through the use of complementary time series methods, which is the focus of the next two Sections.

4.2.2 Time-Series Decomposition

GTD, GDELT, and RDWTI provide the day, month, and year of each terrorist event that occurred in the world from 2002 to 2009. For ease of comparison, the total number of events and fatalities have been aggregated on a monthly basis. Although a visual inspection of time series data is always valuable, it cannot replace a statistical analysis of the *trend*, *cyclical*, and *random* components of time series (observations). The trend component informs us on deterministic long-term behaviour in the series; a statistically positive and significant trend term may reveal a persistent upward pattern.

The cyclical component corresponds to a sinusoidal (also called "wavelike") pattern. A significant cyclical component could reveal the presence of high levels of terrorist activity followed by lull periods. The literature provides various explanations about the presence of cycles observed in terrorism data. It appears that the main cause of most observed cycles in terrorism stems from a repetitive sequence of terrorist actions and government's responses.

[9] The correlation analysis uses Pearson's product-moment correlation to measure correlations of the monthly number of events from 2002 to 2009 (`R` code in Appendix B.2).

Following a series of terrorist attacks, public opinion often successfully pushes government to take counterterrorism actions, which temporally reduce the activity of terrorism. In the meantime, terrorists seize the opportunity to recruit new members and organise fresh attacks, which will contribute to perpetuate cycles of violence (Alexander and Pluchinsky, 1992). A better understanding of the cycles helps authorities to anticipate peaks. This allows for better resource allocation: defensive measures can be increased during peaks and reduced during troughs (Sandler and Enders, 2007). All changes in the series that are not determined by the trend and/or cyclical components are captured by the random component (Enders and Sandler, 1999).

Accordingly, I decomposed the time series into three components: (i) *trend*; (ii) *cyclical*; (iii) *random* (also-called *error* component). The *trend* value x_t for a given month t uses a *two-sided* weighted moving average smoother based on values from $t-6$ to $t+6$ and centred on t (PSU, 2014)[10]. Hence the *cyclical* component is computed by subtracting the trend from the time series and the result is averaged for each month and adjusted so that the total average within the study period (2002-2009) is zero[11]. Furthermore, the *random* component is deduced from removing the trend and cyclical components from the original time series. Note that the random component can be further modelled through an ARIMA model (Section 3.1.3).

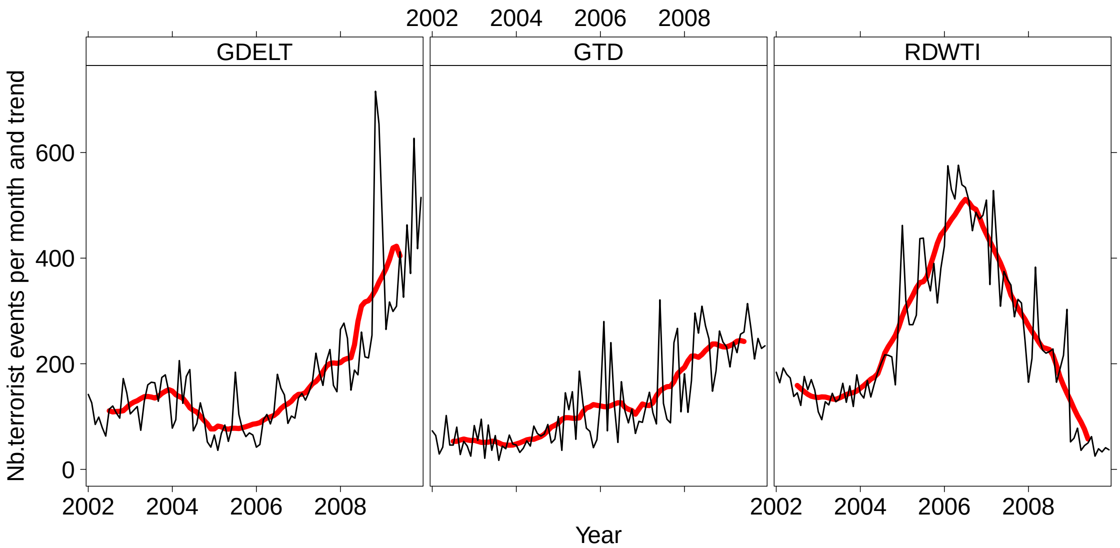
4.2.3 Monthly Number of Terrorist Events

Figure 4.1 illustrates the time series analysis based on the monthly number of terrorist events. In Figure 4.1a, GDELT and GTD suggest a persistent increasing trend (*thick red line*), especially from 2006 to the end of 2009, which contrasts with RDWTI, which indicates a persistent decrease (*thick red line*) for the same period. Note that it remains difficult to ensure that the observed trends (e.g. between 2008 to 2009) in GTD are representative of reality, or rather, artefacts due to inconsistent data collection methods (Jensen, 2013; START, 2014). From 2008, data collection of GTD was done in real time, whereas between 1998 and 2007 the process was done retrospectively. Inevitably, some events that occurred

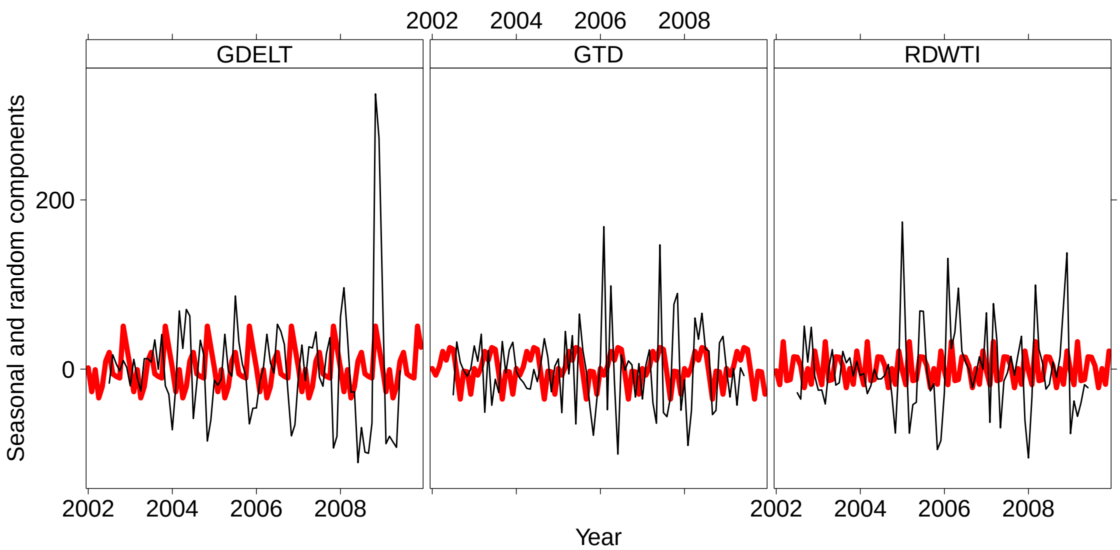
[10] I used the following weights: $x_t = \frac{1}{24}x_{t-6} + \frac{1}{12}x_{t-5} + \frac{1}{12}x_{t-4} + \cdots + \frac{1}{12}x_{t+4} + \frac{1}{12}x_{t+5} + \frac{1}{24}x_{t+6}$. In other words, a weight $\frac{1}{24}$ is applied to values at months $t-6$ and $t+6$ and a weight $\frac{1}{12}$ is applied to between $t-5$ and $t+5$ (except t). This procedure is also known as $2 \times MA$ (Hyndman, 2011). [11] Recall that 2002-2009 is the largest common period in which all databases provide terrorist events since 2002.

between 1998 and 2007 are missing since some media sources have become unavailable (GTD, 2014)[12].

[12] Note that GTD indicates that events from 1998 to 2007 were collected retrospectively starting April 2006 and the provider released the extended version of GTD in March 2009 (LaFree, 2010).



(a) Terrorist events 2002-2009: monthly total number of events and trend component



(b) Terrorist events 2002-2009: random and cyclical components

Fig. 4.1 Worldwide terrorist events 2002-2009, according to GTD, RDWTI, and GDELT. Figure 4.1a: monthly total number of terrorist events (*black line*) and trend component (*thick red line*). Figure 4.1b: random component (*black line*) and cyclical component (*thick red line*).

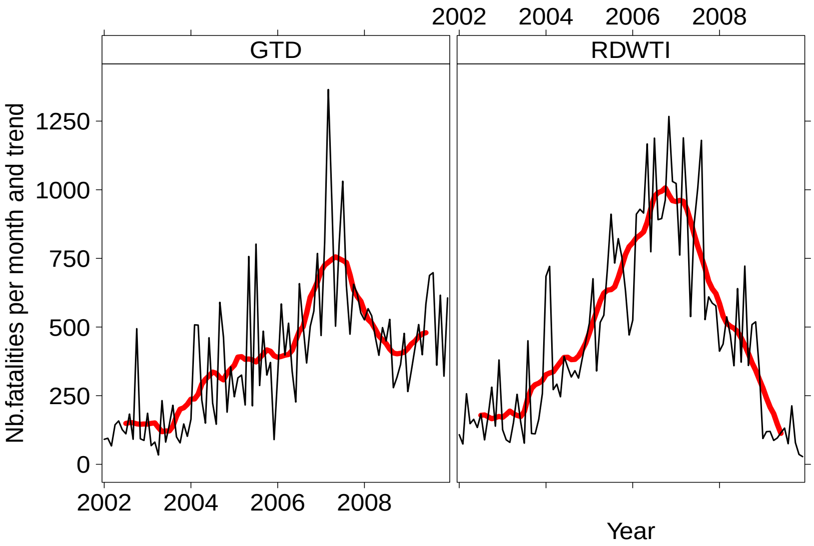
GDELT, GTD and RDWTI suggest the presence of a one-year cycle in the total monthly number of terrorist events, as illustrated by the patterns of the cyclical component (Figure 4.1b, *thick red line*). For each database, the width of the repeated pattern (*period*) corresponds precisely to one year. The *amplitude* (i.e. size of the top (or bottom) half of the wave) of the cyclical component is higher in GDELT compared to the other databases. GTD and RDWTI show similar amplitude values. The patterns of the random component (*thick red line*) are also similar in GTD and RDWTI, in contrast with those observed in GDELT.

4.2.4 Monthly Total Number of Terrorism Fatalities

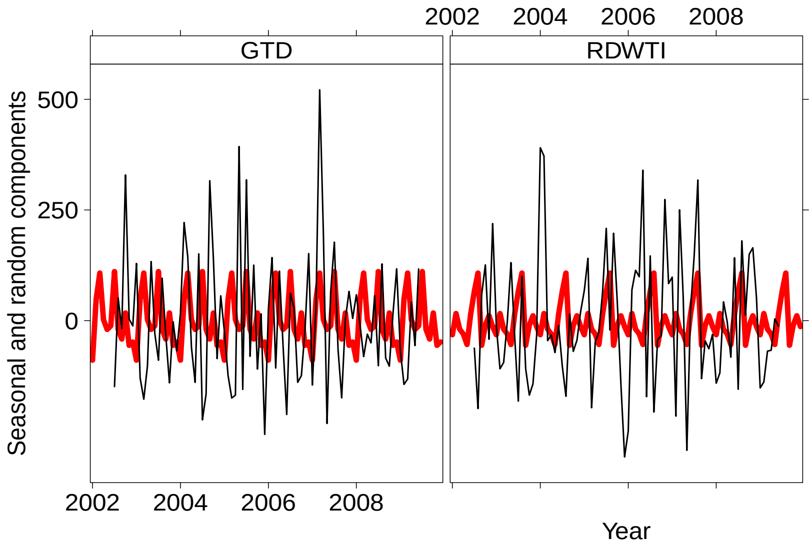
Figure 4.2 illustrates the time series analysis based on the monthly total number of terrorism fatalities. Figure 4.2a reports the observations (*black line*) and the trend component (*thick red line*). Figure 4.2b shows the random (*black line*) and the cyclical components (*thick red line*). Note that GDELT does not provide the number of fatalities, and therefore, does not appear in Figure 4.2.

RDWTI and GTD suggest a persistent decrease in the number of fatalities, especially from 2007 (Figure 4.2a, *thick red line*). The decrease is persistent until 2009 for RDWTI, while GTD suggests a slight increase after 2008. Note that, as previously mentioned with regard to the number of events (Section 4.2.3), one may assume that the slight increase from 2008 to 2009 in GTD is an artefact due to inconsistent data collection methods. From 2007, more events have been collected (GTD, 2014), which might also artificially increase the number of fatalities with regard to previous years. However, this also indicates that the number of events, and indeed the number of fatalities collected prior to 2007 are underestimated in GTD.

Similar to the number of terrorist events, GTD and RDWTI suggest the presence of a one-year cycle in the number of fatalities, as illustrated by the patterns of the cyclical component (Figure 4.2b, *thick red line*). For each database, the period corresponds precisely to one year. The amplitude of the cyclical component is relatively similar between the databases. The patterns of the random component (*thick red line*) are also similar in GTD and RDWTI. Nevertheless a slightly higher amplitude is observed in GTD.



(a) Terrorism fatalities 2002-2009: monthly total number of fatalities and trend component



(b) Terrorism fatalities 2002-2009: random and cyclical components

Fig. 4.2 Worldwide terrorist events 2002-2009, according to GTD and RDWTI. Figure 4.2a: monthly total number of terrorism fatalities (*black line*) and trend (*thick red line*). Figure 4.2b: random component (*black line*) and cyclical component (*thick red line*). Note that GDELT does not provide the number of fatalities, and therefore, does not appear here.

The results of this present work does not support Chalk (1995), who predicted that cycles in terrorism should last approximately between three to five years, given that the time required for the public opinion to pressure the authorities to act accordingly could take several years. Empirical studies have shown that the duration of cycles may depend on various factors, including the type of terrorism (e.g. national, transnational) and the type of attack (e.g. bombing, hijacking). Based on the analysis of transnational terrorist events from ITERATE (1970 to mid-1996), Enders and Sandler (1999) found cycles lasting 3.6 to 18 quarters, according to the type of attacks under investigation. By extending the study period to 2003, Enders and Sandler (2006) found that more logistically complex attacks (e.g. hostage events or hijacking) have longer cycles compared to less sophisticated acts (e.g. threat, hoax) since the action/reaction loop between terrorist groups and governments takes longer to happen[13].

The study of terrorism over time revealed that both the number of terrorist events and fatalities are not the result of a truly random process; rather, events influence each other in time. Both variables are auto-correlated in time and cycles and trends have been identified. Analogously, should we expect that the number of events and fatalities are not independent in space, i.e. auto-correlated in space? In other words, do events influence each other in space? If so, to which extent should one expect events to be influenced each other? On what scale(s) should we expect to identify patterns of terrorism? Section 4.3 will provide these answers.

4.3 Exploring the Spatial Patterns of Terrorism

4.3.1 Terrorist Events in Space

The number of reported events varies according to each database on terrorism. From 2002 to 2009, the total number of reported events is 12,031 and 23,873 in GTD and RDWTI, respectively. Baghdad, Iraq is the most targeted city with 2,237 and 4,083 observations reported by GTD and RDWTI, respectively. Iraq is the most targeted country. GTD reports 4,952 attacks, while RDWTI reports 9,858 attacks. The most widely adopted type of attack is bombing. GTD reports 7,313 `bombing/explosion` events while RDWTI reports 12,250

[13] This present study uses GTD, RDWTI, and GDELT rather than ITERATE, the latter covers a different study period, and additionally, does not differentiate terrorist events according to attack types. Consequently, discrepancies with the results from the aforementioned literature are expected. Given that each database uses different data collection methodologies (see Section 4.1), a proper comparison of the results is therefore difficult.

`explosives` events. The most active group is the Taliban, with a total of 708 and 933 perpetrated attacks, according to GTD and RDWTI, respectively[14].

The databases also report the localisation in space of each terrorist attack that occurred in the world from 2002 to 2009. A quick glance at Figure 4.3 is enough to see that considerable discrepancies in the localisation of terrorist attacks occur within the databases. Despite the differences, GTD and RDWTI share a relatively similar pattern, with a higher number of events in particular regions (e.g. Colombia, East India, Pakistan, or Iraq). By contrast, the events in GDELT exhibit a different pattern. For example, GDELT indicates numerous events which are not pointed out by GTD and RDWTI (e.g. Australia, China). As discussed in Section 2.2.2, the analysis of the observed spatial patterns of terrorism (illustrated in Figure 4.3) could provide valuable insight into the mechanisms that have generated them. More specifically, it might help answering the following questions: are terrorist attacks haphazardly distributed in space — which would provide evidence that terrorists are likely to strike everywhere —, as it has been advocated by Steen et al. (2006)? Or, do terrorists tend to repeatedly target specific locations, leading to the formation of *clusters*, as suggested by empirical findings from Nunn (2007) and Piegorsch et al. (2007) in the US case, or Braithwaite and Li (2007) and Gao et al. (2013) on a global scale? Moreover, do all terrorism databases lead to the same conclusions? In order to answer the questions, I compare the spatial structure of GTD, GDELT, and RDWTI at local and global scales through different techniques described in further details below.

4.3.2 Global Spatial Autocorrelation

A simple visualisation of a spatial pattern does not suffice to identify potential spatial structures in the data. More problematically, humans tend to perceive structures that do not exist within random data, which refers to the concept of *apophenia* (Darmofal, 2015, p. 24). Fortunately, several suitable statistical approaches have been developed to this end. Viewed overall, spatial structures can be identified through the Moran's I (Equation 4.1). The advantage of Moran's I is that it provides information about the spatial structure of spatial patterns, based on a normalised function which can only take values from -1 to 1. Therefore the index, as defined below, is easily interpretable (Zhukov, 2010):

[14] Note that most perpetrators are unknown in GTD (8,182 events) and RDWTI (17,559). Moreover, GDELT does not provide the name of the perpetrators (terrorist groups or individuals).

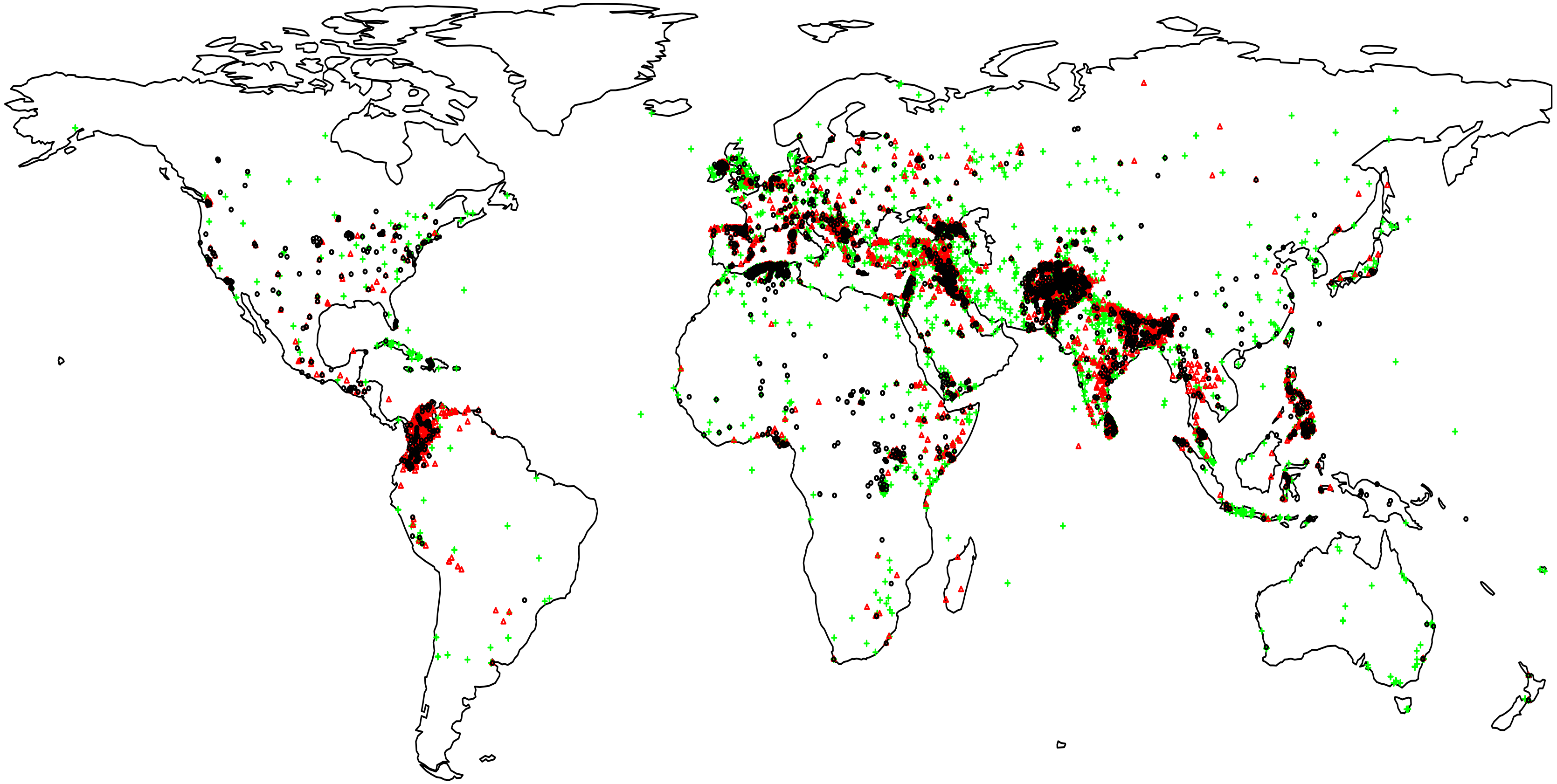


Fig. 4.3 Geolocalisation of worldwide terrorist events perpetrated by *non-state* actors (*state terrorism* is excluded) from 2002 to 2009, according to GTD (*black circle*), GDELT (*green cross*), and RDWTI (*red triangle*) databases. Both *lethal* ($\geq$ 1 death(s)) and *non-lethal* (no death) terrorist attacks are displayed. GTD and RDWTI report high concentration of events in e.g. the Middle East, Pakistan, Colombia, or East India. Discrepancies between GDELT and the other databases can be observed e.g. events in Australia reported by GDELT exclusively.

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{\mathbf{s}_i\mathbf{s}_j}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{\mathbf{s}_i\mathbf{s}_j}(y_{\mathbf{s}_i}-\bar{y})(y_{\mathbf{s}_j}-\bar{y})}{\sum_{i=1}^{n}(y_{\mathbf{s}_i}-\bar{y})^2}. \tag{4.1}$$

In Equation 4.1, the *density* of the events $y_{\mathbf{s}_i}$ could represent the number of attacks that occurred in a given location $\mathbf{s}_i$ in a spatial domain $\mathscr{D} \subseteq \mathbb{R}^d$, with $d = 2$ for two-dimensional phenomena[15]. Moreover, n is the number of different locations (number of observations), $\bar{y}$ is the sample mean, and $w_{\mathbf{s}_i\mathbf{s}_j}$ is an element of a matrix of spatial weights between observations from two distinct locations $\mathbf{s}_i$ and $\mathbf{s}_j$. Following Tobler's law (Section 2.2.1), one expects that close observations are more likely to be similar than those far apart. Hence, it is common to associate a weight $w_{\mathbf{s}_i\mathbf{s}_j}$ to each pair $(y_{\mathbf{s}_i}, y_{\mathbf{s}_j})$ that takes these spatial characteristics into account. The simplest approach is to define $w_{\mathbf{s}_i\mathbf{s}_j} = 1$ for observations in the "close neighbourhood" (three examples of close neighbourhood are illustrated in Figures 3.4, 3.5 and 3.6) and set the weights to 0 otherwise. Furthermore, note that $w_{\mathbf{s}_i\mathbf{s}_i}$ is usually set to 0. More complex approaches can be used such as e.g. inverse distance (see below).

The expected value of Moran's I under the null hypothesis of no spatial autocorrelation is: $-1/(n-1)$, where n is the number of observations. For example, for $n = 3,187$ (which corresponds to the number of GTD observations with a positive density), the values of a homogeneously distributed spatial pattern (i.e. absence of spatial autocorrelation) are $-1/3187 \cong -0.00031$. If the observed I (denoted $\hat{I}$) is significantly greater than the expected value of Moran's I under the null hypothesis, the values of y (e.g. the number of terrorist events) are spatially positively autocorrelated. Conversely, a negative and significant difference between $\hat{I}$ and the expected values under the null hypothesis indicates negative autocorrelation. The absence of spatial autocorrelation cannot be excluded if $\hat{I}$ is not significantly different from the expected values under the null hypothesis (Zhukov, 2010).

Moran's I is highly sensitive to the choice of spatial weights (Zhukov, 2010). Therefore, for each database (GTD, GDELT, and RDWTI), I estimate $\hat{I}$ based on two different approaches used to define the spatial weights (`R` code in Appendix B.3). The comparison of the results allows assessment of the robustness of the spatial weights. Moran's $\hat{I}_1$ uses

[15] Note that locations that have not encountered terrorist attacks are not taken into account.

Table 4.2 World terrorism 2002-2009: Moran's I

Database	n	$\hat{I}_1$	H_{01}	σ_{I1}	$\hat{I}_2$	H_{02}	σ_{I2}
GTD	3,187	0.0028	-0.00032	0.0038	0.0080***	-0.00032	0.0018
GDELT	3,046	-0.0045*	-0.00033	0.00037	0.0024	-0.00033	0.0031
RDWTI	3,654	0.0061	-0.00027	0.0039	0.0061***	-0.00027	0.0014

Two approaches are used to estimate Moran's I: without distance threshold ($\hat{I}_1$); with threshold distance (1,000 km) ($\hat{I}_2$). The estimated values of the mean $\hat{I}_1, \hat{I}_2$ and standard deviation σ_{I1}, σ_{I2} are provided for the first and the second approach, respectively. The Moran's I under the null hypothesis is denoted with index H_{01} and H_{02} with regard to the first and second approach, respectively. A positive spatial autocorrelation is identified if the difference $\hat{I}_1 - H_{01}$ (or $\hat{I}_2 - H_{02}$) is positive and significant. Conversely, a negative spatial autocorrelation is identified if the difference $\hat{I}_1 - H_{01}$ (or $\hat{I}_2 - H_{02}$) is negative and significant.
* $p < 0.1$; *** $p < 0.01$.

an inverse-distance without threshold so that the spatial weight is determined as a function of the inverse distance between observations. This assumes that nearby observations have more influence than far observations on a given observation. Moran's $\hat{I}_2$ is calculated from an inverse-distance with a 1000 km threshold. In the latter case, events beyond the threshold are not considered (`R` code in Appendix B.3).

The expected values under the null hypothesis are reported in Table 4.2 as H_{01} and H_{02} for the first and the second approach, respectively. The number of observations is indicated as n, and the standard deviation of $\hat{I}_1$ and $\hat{I}_2$ are indicated as σ_{I1} and σ_{I2}, respectively. The results (Table 4.2) indicate that the density of terrorist events from GTD and RDWTI does not appear autocorrelated in space on a global scale ($\hat{I}_1$ not significantly different from H_{01}). Surprisingly, the density of events in GDELT appears negatively autocorrelated ($p < 0.1$). This latter result indicates that a *repulsion* process might be at stake, which appears rather an artefact due to methodological issues (Hammond and Weidmann, 2014) instead of the results of a "true" *repulsion* process. Nevertheless, the results from the second approach, which uses a threshold distance (1,000 km), suggests a different pattern. The null hypothesis, which assumes the absence of spatial autocorrelation, is refuted in GTD and RDWTI. In both GTD and RDWTI, positive spatial autocorrelation is suggested, since $\hat{I}_2$ is significantly higher than H_{02} ($p < 0.01$).

Based on the data from GTD and RDWTI, a higher density of attacks in e.g. Baghdad, Iraq contributes to increase the density of attacks within a 1000-km radius, including those perpetrated in e.g. Beirut, Lebanon ($\approx$ 800 km from Baghdad). In contrast, one cannot refute that GDELT events are randomly distributed in space within a 1000-km radius ($p >$

0.1), which may reveal issues in the data collection methodology (e.g. misspecified presence and/or locations of some events)[16].

According to the results of the analysis of the Moran's I based on GTD and RDWTI, terrorism does not seem to strike just anywhere; rather, events appear to influence each other. The analysis based on the Moran's I provides therefore a useful information for a preliminary exploration of the *global* spatial dependence structure in data. The results suggest that a clustering process of terrorism is likely within the study area, which is useful but not sufficient to answer the research questions formulated in this present study (Section 1.1). Without identifying areas of particular clusters, global indices such as the Moran's I cannot capture potential processes that might operate on a *local* scale and generate local spatial patterns. As a further step, the following paragraphs are dedicated to the analysis of the *local* spatial structure of terrorism.

4.3.3 Local Spatial Autocorrelation

It is undeniably of interest to identify space locations of "abnormally" high level of terrorism, also called hotspots (Section 2.2.2). One index of *local* spatial autocorrelation is the Getis and Ord statistic $G^*_{\mathbf{s}_i}$, which provides valuable information on spatial autocorrelation that might occur on a local scale[17]. It is defined as:

$$G^*_{\mathbf{s}_i} = \frac{\sum_j w_{\mathbf{s}_i\mathbf{s}_j} y_j - \sum_j w_{\mathbf{s}_i\mathbf{s}_j}\bar{y}}{\sigma_G \left\{ \left[n\sum_j w^2_{\mathbf{s}_i\mathbf{s}_j} - \left(\sum_j w_{\mathbf{s}_i\mathbf{s}_j} \right)^2 \right] / (n-1) \right\}^{1/2}}, \tag{4.2}$$

where n is the number of observations, y the variable of interest (e.g. density of terrorist attacks), $\bar{y} = \frac{1}{n}\sum_j y_j$ the sample mean, $\sigma_G = (\frac{1}{n}\sum_j y_j^2 - (\bar{y})^2)^{1/2}$ the standard deviation, and $w_{\mathbf{s}_i\mathbf{s}_j}$ is the spatial weight for the pair $(\mathbf{s}_i, \mathbf{s}_j)$. In contrast to *global* indices (e.g. Moran's I), which provide one value to estimate spatial autocorrelation, local measures (e.g. $G^*_{\mathbf{s}_i}$) estimate spatial autocorrelation for each location i (e.g. centroid of grid-cells for lattice data)

[16] As an additional robustness check (not reported in Table 4.2), I used threshold distances of 500 km and 200 km, respectively. Consistently, GTD and RDWTI suggest the presence of positive spatial autocorrelation ($p < 0.01$), while one cannot refute an absence of spatial autocorrelation in GDELT ($p > 0.1$).

[17] Note that the Getis and Ord statistic is often denoted G^*_i in the literature. For consistency with the notation used to denote spatial location in this thesis, the Getis and Ord statistic is denoted $G^*_{\mathbf{s}_i}$.

within the study area. Positive/negative values of $G^*_{\mathbf{s}_i}$ indicate the presence of positive/negative spatial autocorrelation around the location $\mathbf{s}_i$ (Ord and Getis, 1995). Note that $G^*_{\mathbf{s}_i}$ is standardised (asymptotically normal with $G^*_{\mathbf{s}_i} \sim N(0,1)$) (Rogerson, 2015).

In order to compute $G^*_{\mathbf{s}_i}$ for each location $\mathbf{s}_i$, the study area has to be gridded. I used PRIO-GRID, which covers all terrestrial areas of the world at a resolution of 0.5 x 0.5 decimal degrees (Tollefsen et al., 2012). PRIO-GRID provides standardised approach to the analysis of spatial data at the subnational level (Salehyan, 2015) and has therefore been widely used in various fields of research, including conflict and terrorism (see e.g. Buhaug et al. (2009); Buhaug and Rød (2006); Nemeth et al. (2014)). The grid cuts the study area into regular polygons whose size is small enough to reveal the fine-scale spatial structure, while being consistent with the average level of spatial accuracy of terrorism data[18]. Hence, for each PRIO-GRID polygon, I computed the density of terrorist attacks, i.e. the number of terrorist events that occurred between 2002 and 2009 divided by the area of the corresponding PRIO-GRID polygon) for GTD, GDELT, and RDWTI in locations that encountered at least one terrorist attack (`R` code in Appendix B.3).

The number of observations (i.e. the number of PRIO-GRID polygons with at least one terrorist attack) n is the following: 1,320 (GTD), 1,993 (GDELT), and 1,546 (RDWTI). The weights w_{ij} are based on the *distance-based neighbours* (Bivand, 2015), where the nearest neighbours are identified within a radius to be specified. For consistency consideration, an identical radius has been used for all databases. It corresponds to the minimum distance that ensures that all areas have at least one distance-based neighbour, this for all databases. Highly significant high-value clusters ($G^*_{\mathbf{s}_i} > 0$, $p < 0.01$) of terrorism density are illustrated in Figure 4.4. Note that I show only areas which exhibit significant high-value clusters of terrorism's density. It includes mainly areas located in: south-east Europe, north-east Africa, the Middle East and Russia. It is also interesting to note the discrepancies among the databases with regard to location of the clusters, which have been already observed in the analysis of the locations of terrorist events (Figure 4.3). For example, some clusters appear only in one database, e.g. in Iran: GDELT or in Pakistan: GTD.

[18] In the study area, one could assume that events localised by a city name (or by the centroid of the city) are usually localised within a radius $\leq$25 km around the centroid of the city. Indeed, as illustrated in the case of a large city, a disk with a radius of approximatively 22 km from the centre of London until its periphery would almost cover entirely the metropolitan area. Therefore, the uncertainty related to the spatial measurement is suitably taken into account within PRIO-GRID. Note that in order to reduce the complexity of the area, small polygons (e.g. Balearic Islands) have been removed.

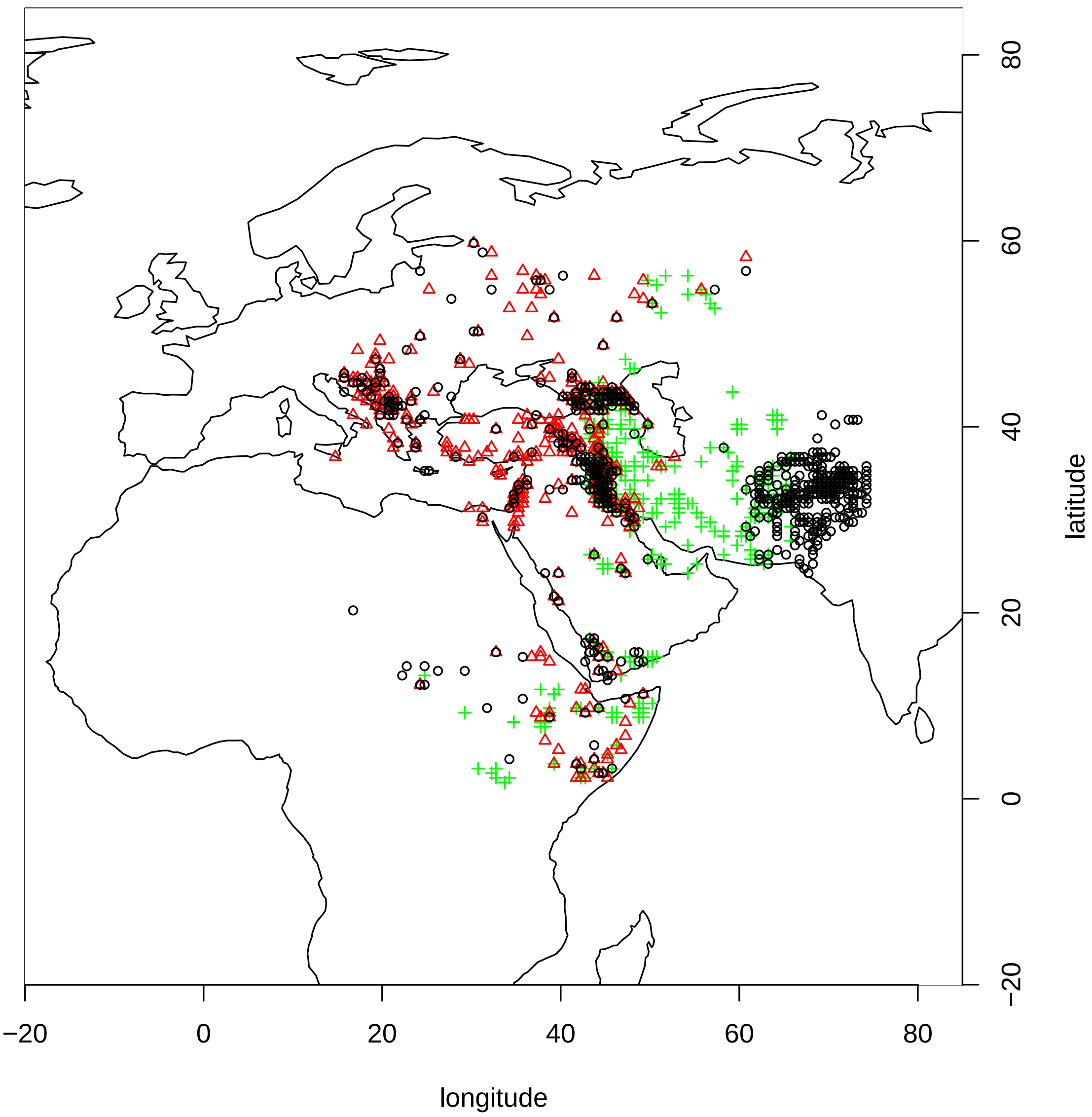


Fig. 4.4 Local hotspots of worldwide terrorism density 2002-2009 identified with the Getis and Ord statistic $G^*_{\mathbf{s}_i}$. Only highly significant high-value clusters are reported ($G^*_{\mathbf{s}_i} > 0$, $p < 0.01$) in GTD (*black circle*), GDELT (*green cross*), and RDWTI (*red triangle*). Note the discrepancies observed among the databases with regard to the location of clusters (e.g. clusters exclusively shown by GTD in Pakistan or by GDELT in Iran).

4.3.4 Point Pattern Analysis

The identification of hotspots through exploratory statistical methods, such as the $G^*_{\mathbf{s}_i}$, is highly dependent on the scale on which spatial autocorrelation is measured. For lattice data, the choice of the grid resolution may have a considerable impact on the detection of patterns, a well-known problem referred to as the modifiable areal unit problem (MAUP) (Holt et al., 1996).

By considering terrorist events as a point pattern (i.e. a realisation of a point process), the analysis of the interactions among the events (i.e. the points in a point pattern) at different distances might reveal clusters, and equally important, the scale on which they take place. The degree of interactions within an observed point pattern can be compared with those from a SHPP. The SHPP is considered as "benchmark model" or model of CSR because the locations of each point are independent from the locations of all the other points. In other words, there is no interaction among points in the SHPP (Equation 3.12). Recall that clustering, or its opposite, *repulsion* processes can be interpreted as a deviation from CSR (Section 3.2.1).

When numerical data are under investigation, one often begins by looking at the sample mean as an indicator of central tendency. The analogue of the mean in the context of point processes is called the *intensity*. The intensity (or *first moment*) is defined as the *expected* number of random points per unit area. The intensity can be constant across space (homogeneous process), as in the SHPP, or may vary from one location to another (inhomogeneous process). The examination of the intensity should be one of the first steps in analysing a point pattern (Baddeley, 2008). Hence, I examine the intensity of the process (estimated by the number of terrorist events per PRIO-GRID unit) of the terrorist events that occurred from 2002 and 2009 for each database (GTD, GDELT, and RDWTI).

There are numerous approaches and associated statistical packages, which provide methods to estimate the intensity of the process (Deng and Wickham, 2011). Here, I used the function `density.ppp` provided by the `R` package `spatstat` to estimate the intensity of the process that generated the point pattern (Diggle, 1985). Numerous functions are available, including various types of Gaussian, Uniform, or Epanechnikov smoothers. The estimation of the intensity function $\hat{\lambda}$ can be expressed as follows:

$$\hat{\lambda}(\mathbf{s}) = \sum_{i=1}^{n} K\big(x_i - \mathbf{s}\big) w_i e\big(\mathbf{s}\big) \quad \mathbf{s} \in \mathscr{D}, \tag{4.3}$$

where $e(\mathbf{s})$ is an edge correction factor, w_i are weights[19], $K(\cdot)$ is used to estimate the intensity of the point pattern $\mathbf{x}$ at location $\mathbf{s}$ in domain $\mathscr{D} \subseteq \mathbb{R}^d$ ($d = 2$ in this present case study). I applied an isotropic Gaussian kernel smoother: $K(s_x, s_y, \sigma_K) = \frac{1}{2\pi\sigma_K^2} \exp\left(-\frac{s_x^2+s_y^2}{2\sigma_K^2}\right)$, with coordinates $(s_x, s_y) \in \mathscr{D}$ and σ_K represents the standard deviation of the isotropic Gaussian kernel smoother. Larger values of σ_K generate smoother estimates. Therefore, the choice of σ_K may have an important impact on the results. For the sake of illustration, I compare the results based on different values of σ_K, namely $\sigma_{K1} = 150$ and $\sigma_{K2} = 300$ (Figure 4.5, *left* and *right* maps, respectively) for GTD, GDELT, and RDWTI. The edge correction factor $e(\mathbf{s})$ is the reciprocal of the kernel mass inside the domain $\mathscr{D}$: $1/e(\mathbf{s}) = \int_{\mathbf{v} \in \mathscr{D}} K(\mathbf{v} - \mathbf{s}) d\mathbf{v}$ (`R` code in Appendix B.4).

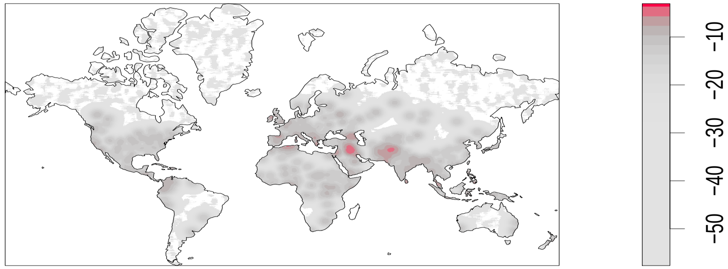
Note that the maps in Figure 4.5 should not be interpreted as a probability density. More precisely, they provide an estimate of the (log) *expected* number of random points per unit area (log (intensity)) based on the point process that could have generated the point pattern (Baddeley and Turner, 2014). As expected, for each database, the estimated (log) intensity varies over space, which suggests that the underlying processes that generated the patterns in GTD, GDELT, and RDWTI are not stationary.

Moreover, it should be further noticed that clusters cannot be properly identified through this approach since their identification is highly dependent on the size of the bandwidth (or σ in the case of the use of an isotropic Gaussian kernel) and the kernel function used. The choice of the kernel function and the bandwidth remains essentially subjective without reliable knowledge on the expected spatial characteristics of the underlying processes.

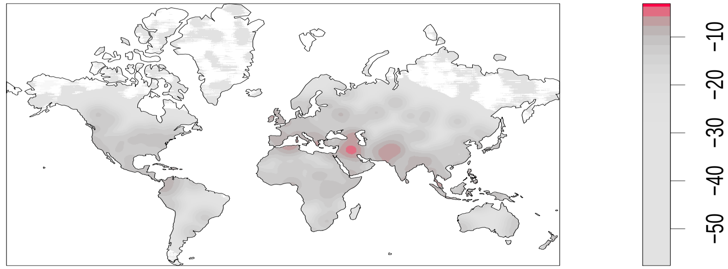
It might be of interest for mayors and local authorities to better understand the extent to which their city might be affected by hotspots of terrorism present in their surroundings and in more remote areas. The pair correlation function (*pcf*), also called *g-function*, provides insights into the *second-order* structure of the spatial patterns of terrorism at various distances. It can determine the scale on which clustering processes take place. The *pcf* is easy to interpret. For a complete spatial randomness process (CSR), $g(r) = 1$ for $r \geq 0$ (*black dashed line* in Figure 4.6). In clustered processes, $g(r) > 1$, while $g(r) < 1$ in regular (also called *dispersed* pattern) processes (Illian et al., 2008).

Thus, Figure 4.6 (*solid lines*) provides the values of the *pcf* estimated at regular intervals of approximately 4-6 km. In total 512 values of *pcf* are computed separately for GDELT (*green line*), GTD (*black line*), and RDWTI (*red line*). First, I set a large distance to examine

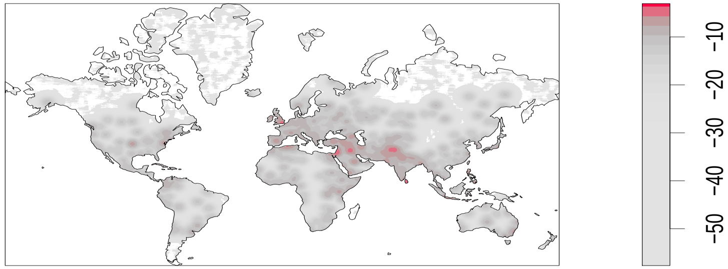
[19] Note that for the sake of simplicity, the weights are set to 1 in this study. However, weights can be defined if desired.



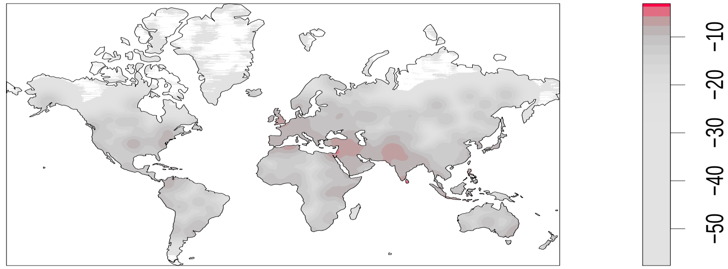
(a) Gaussian kernel density: GTD ($\sigma_1 = 150$)



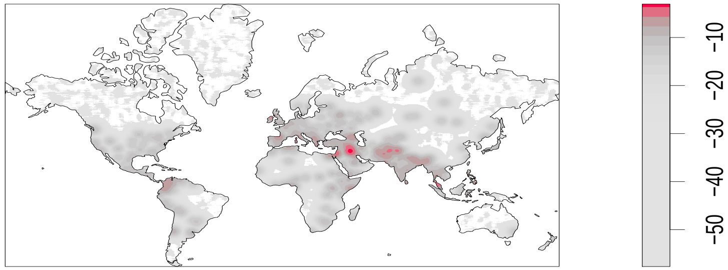
(b) Gaussian kernel density: GTD ($\sigma_2 = 300$)



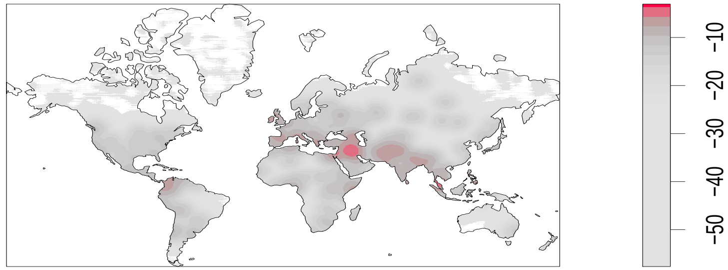
(c) Gaussian kernel density: GDELT ($\sigma_1 = 150$)



(d) Gaussian kernel density: GDELT ($\sigma_2 = 300$)



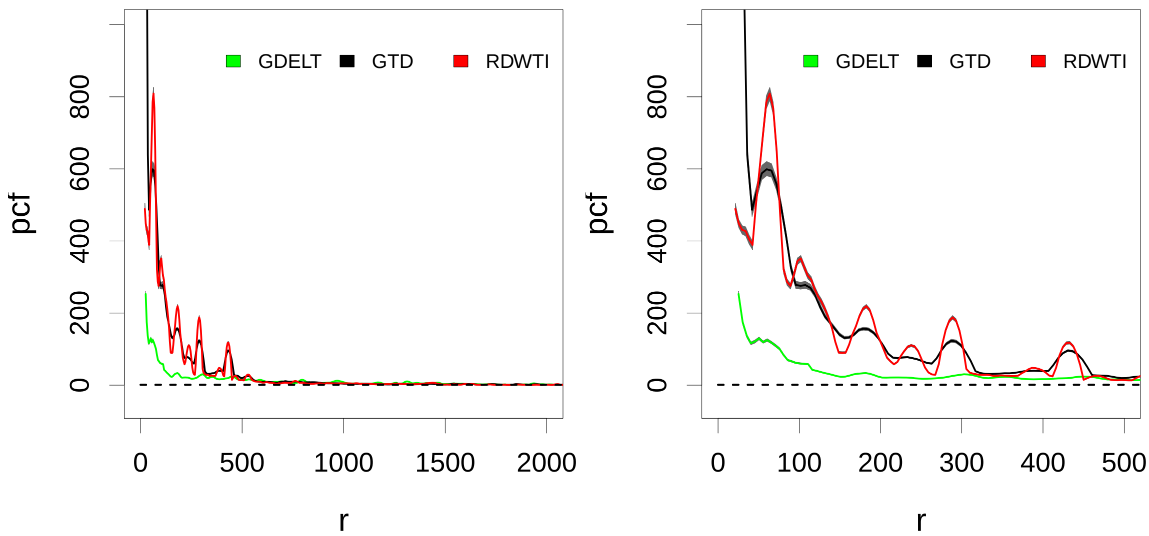
(e) Gaussian kernel density: RDWTI ($\sigma_1 = 150$)



(f) Gaussian kernel density: RDWTI ($\sigma_2 = 300$)

Fig. 4.5 Worldwide terrorist events perpetrated from 2002 to 2009: Gaussian kernel density estimation of the (log) expected intensity of the number of terrorist events in GTD (*top*), GDELT (*centre*), and RDWTI (*bottom*). The unit is the (log) expected number of terrorist events per km^2.

the extent of the clustering process (Figure 4.6a: from 0 to 2,000 km). The clustering processes appear to be drastically reduced after 500 km; beyond this distance, all estimates are close to 1 (CSR). Therefore, I carried out an additional procedure focused on lower distances (Figure 4.6b: from 0 to 500 km), which provides a more detailed view of the scale on which the process that generates clusters take place[20].



(a) *pcf* computed between 0-2,000 km **(b)** *pcf* computed between 0-500 km

Fig. 4.6 Worldwide terrorist events perpetrated from 2002 to 2009: pair correlation function estimated for GTD (*black line*), GDELT (*green line*), and RDWTI (*red line*) and their corresponding 95% credible intervals (*grey area*) between 0 and 2,000 km (Figure 4.6a) and between 0 and 500 km (Figure 4.6b). The theoretical value (*dashed line*) of a CSR process is 1.

In Figure 4.6b, all point patterns show clusters at various distances from 0 to approximately 500 km since the values are well above 1, which is the theoretical value for a *pcf* in CSR. The values of the *pcf* at 0 km (or close to 0 km, until 17 or 24 km depending on the database) are infinite (not showed in Figure 4.6) and decrease drastically until reaching 20-30 km depending on the database. At low distances (50-70 km), the values of the *pcf* in GTD and RDWTI are high, while GDELT indicate lower values. In contrast to GTD and RDWTI, GDELT shows few clusters until approximately 150 km, but beyond this distance

[20] In order to quantify the uncertainty related to the estimation of the *pcf* (Figure 4.6), 95% bootstrap confidence intervals have been generated. 99 simulations using the bootstrap method developed by Loh (2008) have generated the confidence intervals. Moreover, a border correction is applied to adjust for edge effect bias.

its pattern cannot be distinguished from those from a CSR process. GTD and RDWTI show approximately five peaks at similar distances (50-70 km; 95-110 km; 160-190 km; 280-300 km; 420-440 km). Near approximately 450 km, they seem to follow a CSR process.

The result of the analysis of spatial patterns indicates that from 2002 to 2009, terrorist events tended to be highly clustered, mainly at relatively low distances with peaks at approximately 50 to 70 km for the GTD database. This suggests that the underlying processes which generate spatial clustering operate mainly at the local, and regional scale rather than at the global scale. Beyond approximately 450 km, the interaction among terrorist events becomes negligible for both RDWTI and GTD.

Spatial proximity might play an important role in explaining terrorist attacks in areas that have encountered terrorism (see Section 2.2.3). Furthermore, the influence of terrorist events becomes negligible beyond approximately 500 km. Note that since terrorist events exclusively occur in land — maritime terrorism (e.g. attacks on vessels or fixed platforms at sea) is not included in GTD — seas can be considered as "natural barriers", which ineluctably lead terrorists to act on land, and artificially generate a clustering process at different distances. Despite this potential artifice, the distribution of the land across the world cannot suffice to explain the observed patterns and further causal investigation is required.

4.4 Conclusion

The exploratory data analysis carried out throughout this Chapter revealed important discrepancies among the current terrorism databases. Based on the same study period (2002-2009) and study area (world), the number, localisation, and type (lethal/non-lethal) of terrorist events vary considerably from one database to another. As a result, inconsistencies among the databases with regard to the temporal and spatial patterns of terrorism have been revealed. The use of different sources of media reports, definition(s) of terrorist events, and data collection methods form the basis of the observed discrepancies. However, one cannot blame data providers alone. The phenomenon under study, terrorism, is inherently ambiguous and differences in its interpretation are inevitable (Section 2.1.1). In sum, one should remind that each database only reflects one particular interpretation of this complex phenomenon.

Despite the differences, similarities have been identified in the time and spatial dimensions. In time, the trend of the monthly number of fatalities due to terrorism appears consistent between GTD and RDWTI. In space, all databases indicate that terrorist events appear clustered at relatively low distances. This is evidence of the presence of temporal and spatial autocorrelation, which violates the assumption of the independence of events, and therefore, has to be taken into account in the modelling process. It also signifies that it is very likely that terrorism does not occur anywhere; terrorists tend to repeatedly target specific locations, which suggests the presence of local hotspots generated by factors that operate on a local scale. As a result, the underlying processes which might have caused the clustering patterns may be further investigated and modelled.

Terrorism is essentially a local phenomenon, as suggested by its tendency to form clustered patterns on a local scale. A better understanding of the local patterns of terrorism requires the identification of local drivers. Recall from Section 2.1.3 that the literature at country-level found that host countries with high per capita GDP may observe more terrorist attacks (Blomberg and Rosendorff, 2009; Tavares, 2004). Analogously, one might ask about the role of per capita GDP on terrorism on a local scale. Furthermore, the size of the population appears to play an important role in the occurrence of terrorism at country-level (Braithwaite and Li, 2007; Burgoon, 2006; Drakos and Gofas, 2006b; Dreher and Fischer, 2010; Freytag et al., 2010; Koch and Cranmer, 2007; Krueger and Laitin, 2008; Krueger and Maleckova, 2003). Whether these claims hold at local level is the subject of the next Chapter.

Undeniably, the identification of the causes of terrorism is challenging. Given the inherent ambiguity of the term and the tremendously complex interactions of its causes across multiple scales, the patterns revealed in the study are likely to be the results of socio-economic and/or geographic drivers. These promising results, along with the availability of powerful statistical tools, give hope of successful modelisation of major issues of terrorism, such as its propensity to generate deaths. Modelling the *local* spatio-temporal dynamics of lethal terrorism worldwide remains a challenging endeavour, but above all, a valuable task: "[e]ssentially, all models are wrong, but some are useful" Box and Draper (1987, p. 424).

Chapter 5

Modelling Lethal Terrorism in Space and Time

Terrorism today threatens the life of thousands of citizens around the world, as illustrated by the ongoing deadly attacks perpetrated by the Islamic State (IS) in Iraq, Syria, and Europe, Boko Haram (Arabic: جماعة أهل السنة للدعوة, “The Group of the People of Sunnah for Preaching and Jihad”) in Nigeria, and the al-Nusra Front or Jabhat al-Nusra (Arabic: جبهة النصرة لأهل الشام, “The Support Front for the People of Al-Sham”) in Syria for example. Yet the human losses caused by terrorism and the enormous costs required to combat it (NATO, 2008; The Washington Post, 2013) could be reduced through the use of statistical models able to bring insight into the mechanisms that drive lethal terrorism.

There is an urgent need to better understand why specific areas are targeted by terrorist events that generate losses of human lives, which I refer to as *lethal* terrorism. As previously discussed (Chapter 4), terrorism is essentially a local phenomenon. Modelling the fine-scale dynamics of lethal terrorism in the entire world might *a priori* appear too ambitious given the inherent complex nature of terrorism and its drivers. However, recent advances in Bayesian statistical modelling techniques have allowed scholars to successfully model equally complex social phenomena (e.g. crime or insurgencies) with high temporal and spatial accuracy (Mohler, 2014; Zammit-Mangion et al., 2012). In the same vein, this present Chapter will demonstrate that the spatial dynamics of lethal terrorism can be accurately modelled in both space and time. Moreover, the systematic approach suggested in this study provides a rigorous framework to assess several theories advanced to explain the spatial dynamics of terrorism, which include the *failed state theory* (Gros, 1996; Helman and Ratner, 1992) and *rough terrain theory* (Abadie, 2006; Ross, 1993).

Drawing from the SPDE approach (Section 3.3), distinctive features of the *lethality* of terrorism are modelled through two models: (i) the Bernoulli models the *propensity of terrorist attacks to be lethal*, which further equates to the *probability* of lethal attacks; (ii) the Poisson models the *number of expected lethal terrorist attacks*, which further refers to as the *number* of lethal attacks. Section 5.1 describes the selection of the database on terrorism and covariates. Section 5.2 provides details on the methods used in this present study, including the construction of the mesh and the specification of the Bernoulli and Poisson models. Section 5.3 presents the results and assesses theoretical work based on the findings of this study. Moreover, it describes a prior sensitivity analysis along with an assessment of the robustness of the results to changes in the mesh size. An assessment of the validity of the Bernoulli and Poisson models complements Section 5.3. Section 5.4 provides a summary of the findings, highlights the limitations of the work and suggests possible extension of this present study.

5.1 Data Selection

5.1.1 Selection of Terrorism Database

In order to build valuable, empirically based models, choosing an appropriate dataset is a crucial requirement (Zammit-Mangion et al., 2012). In the previous Chapter, I introduced the main characteristics of the currently available databases: GTD, GDELT, and RDWTI and described their properties through exploratory analysis tools. Following Sheehan (2012)'s approach, I define four *primary* and two *secondary* selection criteria to select the most suitable database, which will be used for modelling purpose. Primary criteria refer to attributes required in this present study: (i) *conceptual clarity*; (ii) clear *coding method*; (iii) information on *fatalities*; (iv) accurate *spatial resolution*. Secondary criteria include desirable but not necessary attributes: (i) high *number of variables*; (ii) wide *temporal scope*.

First, *conceptual clarity* refers to the degree of clarity with regard to the definition(s) used to gather and classify terrorist events[1]. While the definitions are clearly stated in GTD

[1] Recall that the concept of terrorism is intrinsically ambiguous and still being debated today (Beck and Miner, 2013). Inevitably, data-driven models depend on the methodological employed to collect data, whose methodological rigour might vary according to each database provider.

and RDWTI, GDELT, in contrast, does not provide a precise definition. Unlike RDWTI and GDELT, GTD allows the user to filter events according to one or multiple alternative definition(s) of terrorism.

Second, despite both GTD and RDWTI basing their *coding method* on specific, clearly stated criteria, potential biases may occur in the data due to reporting methods used by the media in particular. The collection of GTD events that occurred between 1998 and 2013 was based on a cross-referencing system, which used more than 4,000,000 news articles and 25,000 news sources (GTD, 2014). Nevertheless, since only events available from open sources or reported by the media are taken into account, under-reporting cannot be ruled out (Drakos, 2007; Drakos and Gofas, 2006a). More problematically in GDELT, its fully automated coding system may generate strong geographic bias, which lead Hammond and Weidmann (2014) to advise scholars not to use GDELT. Hammond and Weidmann observed that the correlation between GDELT and hand-coded data on local conflict is mediocre. GDELT over-reports violence in the proximity of capital and under-reports events in remote areas.

Third, both GTD and RDWTI report the number of *fatalities*, which is not the case in GDELT. GDELT is not suitable for our purpose since it does not provide the number of fatalities or information on lethality of terrorism. Fourth, high *spatial resolution* is crucial given the fine-scale desired level of analysis in this study. In contrast to GDELT and RDWTI, GTD is the only database providing the level of spatial accuracy of each event, through the variables called `specificity` and `vicinity` (Section 4.1.3).

Secondary criteria includes the availability of a high *number of variables*, which can be used to provide potential covariates for modelling purposes, or used as categorical variables in order to analyse and compare sub-samples. GTD provides a higher number of variables (132) compared to GDELT (57), and RDWTI (8). Since this study aims to incorporate a large number of events that occurred after 9/11, a wide *temporal scope* starting 2002 is desired. GDELT has a wide temporal scope since it is continuously updated. To a lesser extent, GTD provides data on a relative large time period, since it is updated every year. In contrast, RDWTI is not updated any more and includes events until 2009 only.

Table 5.1 Databases on terrorism (GTD, GDELT, and RDWTI): selection criteria.

Selection criteria	GTD	GDELT	RDWTI
Primary criteria			
1. Conceptual clarity	**high**	low	medium
2. Coding methodology	**high**	no	**high**
3. Fatalities	**yes**	no	**yes**
4. High spatial resolution	**yes**	no	no
Secondary criteria			
5. Number of variables	**high**	**high**	low
6. Temporal scope	medium	**high**	low

For each criterion, the elements in bold represent values at the highest rank. If the highest rank is shared by two databases, the corresponding elements are both indicated in bold.

Based on a ranking established according to the degree of fulfilment of the criteria (Table 5.1), GTD appears superior or equal to the other databases (GDELT and RDWTI) in 5 out of the 6 criteria. Therefore, I select GTD data for further modelling purposes. From now on the *terrorist attack* concept will be understood following the working definition provided by GTD.

Definition 5.1. (**Terrorist attack** (GTD, 2014)) *The threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation*. Moreover, at least two of the following three criteria must be present for an incident to be considered as terrorist attack: (i) *the act must be aimed at attaining a political, economic, religious, or social goal*; (ii) *there must be evidence of an intention to coerce, intimidate, or convey some other message to a larger audience (or audiences) than the immediate victims*; (iii) *the action must be outside the context of legitimate warfare activities*.

As a next step, data on terrorism is extracted from GTD (R code in Appendix A.2). From the 125,087 GTD terrorist events perpetrated from 1970 to 2013, 91% are accurately geolocalised in a city, town or village (`specificity`= $\{1,2\}$) or their spatial accuracy is not specified (`specificity`=0). Hence, events with both `specificity`=0 and `vicinity`=0 (not localised in the immediate vicinity a city) have been excluded. When the coordinates were missing, I used Google Earth™ to find the coordinates of the centroid of events based on the indication of the city name. Events that could not be found with Google Earth™ have been excluded (`R` code in Appendix A.5).

As a result, the sample extracted from GTD contains 57,100 observations. Although a considerable number of observations have been excluded, this culled selection prevents from gathering events whose spatial accuracy does not meet the level of spatial resolution required to conduct the present research. Moreover, since this research focused on post-9/11 terrorism, I have kept the 34,621 events that occurred between 2002 and 2013.

5.1.2 Selection of Covariates

Potentially relevant covariates are identified based on a thorough review of 43 studies carried out at country level by Gassebner and Luechinger (2011), who highlighted the main explanatory factors among 65 potential determinants of terrorism. Out of these, I consider covariates that satisfy two essential characteristics: (i) potential relationship with the *probability* or *number* of lethal attacks; (ii) availability at high spatial resolution, in order to model fine-scale spatial dynamics of terrorism worldwide. Seven spatial and space-time covariates meet these criteria and their potential association with terrorism (described in more detail below): satellite night light, population density, political regime, altitude, slope, travel time to the nearest large city, and distance to the nearest national border.

First, I assess the role of economic development on the lethality of terrorism, whose possible effects are still under debate. Most country-level empirical studies have not provided any evidence of a linear relationship between terrorism and gross domestic product (GDP) (Abadie, 2006; Drakos and Gofas, 2006b; Gassebner and Luechinger, 2011; Krueger and Laitin, 2008; Krueger and Maleckova, 2003; Piazza, 2006), without excluding possible non-linear relationship (Enders and Hoover, 2012). Case studies focused in the Middle East, including Israel and Palestine, showed that GDP is not significantly related to the number of suicide terrorist attacks (Berman and Laitin, 2008). Few studies, however, found that countries with high per capita GDP may encounter high levels of terrorist attacks (Blomberg and Rosendorff, 2009; Tavares, 2004). In line with the subnational nature of this study, I use *NOAA satellite lights at night* (*Version 4 DMSP-OLS*) as a covariate, which provides information about worldwide human activities on a yearly basis and at a high spatial resolution (30 arc-second grid) (Chen and Nordhaus, 2011; NOAA, 2014). This variable has been used as a proxy for socio-economic development measures such as per capita GDP estimation (Ebener et al., 2005; Elvidge et al., 2007; Henderson et al., 2009; Sutton and Costanza, 2002; Sutton et al., 2007). Note that three versions are available: *Average Visible*, *Stable Lights*, and *Cloud Free Coverages*. I use *Stable Lights*, which filter background noise and identify zero-cloud free observations (NOAA, 2014). In order to compare values of different years, I calibrate the data according to the procedure described in Elvidge et al. (2013, Chap.6).

Second, I assess the role of demography. Cities may provide more human mobility, anonymity, audiences and a larger recruitment pool in comparison to rural areas (Crenshaw, 1990, p. 115; Savitch and Ardashev, 2001). Large cities, in particular, offer a high degree of anonymity for terrorists to operate (Laqueur, 1999, p. 41). More specifically, densely populated areas appear vulnerable and are usually more prone to terrorism than sparsely populated areas (Coaffee, 2010; Crenshaw, 1981; Ross, 1993; Savitch and Ardashev, 2001; Swanstrom, 2002). In addition, locations that shelter high-value symbolic targets (buildings or installations), human targets (government officials, mayors, etc.), and public targets (public transports, shopping centres, cinemas, sport arenas, public venues, etc.) are particularly vulnerable to suicide terrorism (Hoffman, 2006, p. 167). Therefore, I use the *Gridded Population of the World* (*v3*), which provides population density on a yearly basis and at high-resolution (2.5 arc-minute grid) (CIESIN, 2005). Moreover, terrorists usually require free and rapid movement by rail or road in order to move from and to target points (Heyman and Mickolus, 1980; Wilkinson, 1979, p. 189). I compute the travel time from each terrorist event to the nearest large city (more than 50,000 inhabitants) based on the *Travel Time to Major Cities* (Nelson, 2008) at a high spatial resolution (30 arc-second grid).

Third, I assess the role of geographical variables: altitude, surface topography (slope), and distance to the nearest national border. Although the relationship between altitude, slope, and the lethality of terrorism is not straightforward, both variables provide an indication of the type of the geographical location, which could be a determining factor for terrorists regarding their choice of target (Ross, 1993). Moreover, Nemeth et al. (2014) suggested that distance to the nearest national border, altitude and slope might have an impact on terrorist activity. I extract both variables from *NOAA Global Relief Model* (ETOPO1), which provides altitude values at high spatial resolution (1 arc-minute grid) (Amante and Eakins, 2009).

Fourth, I assess the role of the level of democracy on the lethality of terrorism. I extract the level of democracy from *Polity IV Project, Political Regime Characteristics and Transitions, 1800-2014* (Polity IV) (Marshall et al., 2014). Polity IV informs about the level of freedom of press, and captures the level of democracy from -10 (hereditary monarchy) to $+10$ (consolidated democracy) for most independent countries from 1800 to 2014. Therefore, it has been commonly referred as proxy for measuring the type of regime or the extent of constraints in democratic institutions (Gleditsch, 2007; Li, 2005; Piazza, 2006). The level of democracy is expected to be positively associated with terrorism (Eubank and Weinberg, 2001; Ross, 1993; Schmid, 1992); the costs of terrorism in democracies tend to be reduced since terrorist groups might benefit from fewer constraints in speech, movement and association (Li, 2005). The effect of the level of democracy on terrorism might be however overestimated since terrorist events tend to be under-reported in non-democratic countries where the press is often not free (Drakos, 2007; Drakos and Gofas, 2006b).

Fifth, I assess the role of ethnicity. I compute the number of different ethnic groups from the *Georeferencing of ethnic groups* (GREG) database. GREG is the digitalised version of the Soviet *Atlas Narodov Mira* (ANM), and counts 1,276 ethnic groups around the world (Weidmann et al., 2010). Although ANM includes information dating back to the 1960s, it is still regarded as a reliable source for ethnicity across the world (Bhavnani and Choi, 2012; Morelli and Rohner, 2014). Although ethnic diversity does not necessarily lead to violence per se (Silberfein, 2003, p. 68), studies at country-level suggest that more terrorism may occur in ethnically fragmented societies (Gassebner and Luechinger, 2011; Kurrild-

Klitgaard et al., 2006), in countries with strong ethnic tensions, or may originate from ethnic conflicts in other regions (Basuchoudhary and Shughart, 2010).

Table 5.2 Terrorism's lethality (GTD 2002-2013): covariate correlation.

	pop	*lum*	*greg*	*pol*	*tt*	*alt*	*slo*
lum	0.25***						
greg	-0.06***	0.08***					
pol	-0.03***	0.03***	-0.03***				
tt	-0.14***	-0.45***	-0.06***	0.00			
alt	-0.20***	-0.29***	0.17***	-0.07***	0.13***		
slo	0.04***	-0.31***	0.08***	-0.08***	0.15***	0.42***	
distb	-0.14***	0.10***	0.10***	0.04***	-0.02***	0.08***	-0.15***

The values represent the Pearson correlation coefficient (ρ) computed for each pair of covariates: population density (*pop*), satellite night light (*lum*), ethnic diversity (*greg*), political regime (*pol*), travel time to the nearest major city (*tt*), altitude (*alt*), slope (*slo*), and distance to the nearest national border (*distb*). $^{***}p < 0.001$.

Table 5.2 shows that the correlation (Pearson correlation coefficient ρ) among political regime (*pol*) and travel time to the nearest major city (*tt*) is very weak and not significant ($\rho < 0.005$, $p > 0.1$). The correlation between the other pairs of covariates is always highly significant ($p < 0.001$) but weak ($0.02 \leq |\rho| < 0.30$) in most cases. However, the positive correlation between altitude (*alt*) and slope (*slo*) is relatively high and significant ($\rho = 0.42$, $p < 0.001$), since the computation of *slo* is directly derived from *alt*. In order to avoid multicollinearity, which may affect the accuracy of the estimation of each parameter (Zuur et al., 2010), I exclude *slo*. I believe that *alt* is a better proxy for estimating mountainous regions, where terrorist groups might have tactical advantages to carry out attacks (Abadie, 2006; Fearon and Laitin, 2003). While travel time to the nearest major city (*tt*) and satellite night light (*lum*) also exhibit (negative) significant and relatively high correlation value ($\rho = -0.45$, $p < 0.001$), I keep both *tt* and *lum*, since they highlight two distinct mechanisms, respectively: (i) link with socio-economic development (luminosity as proxy for per capita GDP estimation) (Ebener et al., 2005; Sutton et al., 2007) and (ii) tactical advantage of the use of communication network (rail system, highway, etc.) within or close to large cities (Heyman and Mickolus, 1980; Wilkinson, 1979, p. 189).

5.2 Method

5.2.1 Mesh Construction

The first operational step in the modelling process is the construction of the mesh. Recall that the SPDE and its discretised solutions represented by the GMRF are built on the top of a mesh, also-called Constrained Refined Delaunay Triangulation (CRDT) (Section 3.3) (Lindgren, 2012; Lindgren and Rue, 2013). The mesh is generated on the surface of the earth (approximated as a sphere), which avoids undesirable boundary effects (i.e. increase of variance near the boundary) that occur in projected maps (two-dimensional surface). [2]

Small triangles are used to cover the area of interest (land) while larger triangles cover areas where terrorism is assumed not to occur (seas, oceans). This would not be feasible with a traditional grid formed by regular squares, such as PRIO-GRID. As a result, the computational time is reduced without compromising the high spatial accuracy required in the study. The mesh (Figure 5.1) counts 9,697 vertices[3].

[2] The boundary effects in two-dimensional domains are due to the use of deterministic Neumann boundaries, which facilitate the construction of the mesh, but generate an increase of variance near the boundary. In `R-INLA`, they could be mitigated by avoiding sharp corners and extending the domain, however (Lindgren and Rue, 2015). [3] The mesh has been generated with `R-INLA` using the following parameters: `max.edge=c(3,1000)/180`, `cutoff=3/180`. The `cutoff` parameter is used to avoid building too many small triangles around clustered input locations, and `max.edge` specifies the maximum allowed triangle edge length in the inner domain and in the outer extension (Lindgren, 2012).

Fig. 5.1 Constrained Refined Delaunay Triangulation (CRDT) or *mesh* that covers the study area (earth's surface). The mesh counts 9,697 vertices. The following regions are highlighted: Americas (*left*), Europe, Africa and West Asia (*centre*), East Asia and Oceania (*right*).

5.2.2 Bernoulli Model

As a common practice in Bayesian spatial and spatio-temporal models, I suggest using a Bayesian hierarchical framework (HM), which is composed of three *stages*: (i) *observations*; (ii) *linear predictor*; (iii) *hyperparameters* (Section 1.2.3). Recall that the first stage considers the conditionally independent observations given the latent field and the hyperparameters of the model. The second stage describes the linear predictor, formed by the intercept, the coefficients of the covariates, and the latent Gaussian field (Section 3.3.2), whose spatio-temporal structure is modelled through the SPDE approach (Section 1.3.2). The third stage describes the priors assigned to the hyperparameters of the model.

While terrorist attacks are inherently discrete in space (terrorist attacks occur at specific locations on earth), I consider their lethality Y as a continuous phenomenon (in the sense of geostatistics (Cressie, 1991), see Section 3.2.2), which is observed at $\boldsymbol{s}$ locations (attacks) over the surface of the earth $\mathbb{S}^2$ at time $t \in \mathbb{R}$. The lethality is assumed to be the realisations of a continuously indexed space-time process $Y(\boldsymbol{s},t) \equiv \{y(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathscr{D} \subseteq \mathbb{S}^2 \times \mathbb{R}\}$, from which inference can be made about the process at any desired locations in $\mathscr{D}$ (Cameletti et al., 2013b). Hence, the hierarchical modelling framework is composed of three levels (`R` code in Appendix C.3):

$$y(\boldsymbol{s}_i,t)|\boldsymbol{\theta},\xi \sim \text{Bernoulli}(\pi(\boldsymbol{s}_i,t)) \tag{5.1a}$$

$$\text{logit}(\pi(\boldsymbol{s}_i,t))|\boldsymbol{\theta} = \eta(\boldsymbol{s}_i,t)|\boldsymbol{\theta} = \beta_0 + \boldsymbol{z}(\boldsymbol{s}_i,t)\boldsymbol{\beta} + \xi(\boldsymbol{s}_i,t) \tag{5.1b}$$

$$\boldsymbol{\theta} \sim p(\boldsymbol{\theta}), \tag{5.1c}$$

where the lethality $y(\boldsymbol{s}_i,t)$ is a dichotomous variable that takes the value 1 if the attack generated one or more deaths, and 0 if not (Equation 5.1a). The parameters to be estimated are $\boldsymbol{\theta} = \{\beta_0, \boldsymbol{\beta}, \sigma^2, \kappa, \rho\}$, which include the marginal variance σ^2 and scale parameter $\kappa > 0$ of its Matérn covariance function (Equation 3.32), and the temporal autocorrelation parameter $|\rho| < 1$ (described in more detail below) (Equations 5.1a, 5.1b, 5.1c).

In `R-INLA`, priors related to the variance of the Matérn covariance function are defined on τ instead of the marginal variance σ^2. As mentioned in Section 3.3.3, for a two-dimensional process, τ^2 and σ^2 are related so that: $\sigma^2 = \Gamma(\nu)/\Gamma(\alpha)4\pi\kappa^{2\nu}\tau^2$ (Lindgren et al., 2011). Hence, I use the default option for the stationary model in `R-INLA` as priors on τ and κ through prior distribution on $\log(\tau), \log(\kappa) \sim N(0,1)$, so that $\log(\kappa(\boldsymbol{s})) = \log(\kappa)$ and

$\log(\tau(\boldsymbol{s})) = \log(\tau)$. For prior sensitivity analysis, I compare the results with alternative priors on the parameters of the Matérn covariance function, which is further discussed in Section 5.3.3.

The conditional distribution of the linear predictor $\eta(\boldsymbol{s}_i,t) = \text{logit}(\pi(\boldsymbol{s}_i,t))$, given the parameters $\boldsymbol{\theta}$ (Equation (5.1b)), includes an intercept β_0, a vector of k covariates $\boldsymbol{z}(\boldsymbol{s}_i,t) = (z_1(\boldsymbol{s}_i,t),\ldots,z_k(\boldsymbol{s}_i,t))$ with coefficient vector $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_k)'$ and the GMRF $\xi(\boldsymbol{s}_i,t)$[4].

Based on GTD (2014), I extract the lethality of 35,917 accurately geolocalised terrorist attacks that occurred from year $t = 2002$ to $t = 2013$ in locations $\boldsymbol{s}_i$. In Equation 5.1a, I assume that $y(\boldsymbol{s}_i,t)$ follows a conditional Bernoulli distribution with probability $\pi(\boldsymbol{s}_i,t)$ of observing a lethal event and $1-\pi(\boldsymbol{s}_i,t)$ of observing a non-lethal event, given the GMRF ξ and parameters $\boldsymbol{\theta}$.

In order to minimise the complexity of the models, and consequently, reduce the computing time required to fit them, I assume a separable space-time covariance (Blangiardo and Cameletti, 2015, chap 7). Hence, the GMRF $\xi(\boldsymbol{s}_i,t)$ follows a first-order autoregressive process AR(1): $\xi(\boldsymbol{s}_i,t) = \rho\,\xi(\boldsymbol{s}_i,t-1) + \varepsilon(\boldsymbol{s}_i,t)$, with time independent zero-mean Gaussian field $\varepsilon(\boldsymbol{s}_i,t)$, $Cov(\xi(\boldsymbol{s}_i,t),\xi(\boldsymbol{s}_j,u)) = 0$ if $t \neq u$, and $Cov(\xi(\boldsymbol{s}_i),\xi(\boldsymbol{s}_j))$ if $t = u, \forall t,u \in \{2003,\ldots,2014\}$.

5.2.3 Poisson Model

In addition to modelling the probability of lethal terrorist attacks, one might identify locations that are more likely to encounter a higher number of lethal attacks over a year, which I further refer to the *number* of lethal terrorist attacks. The identification of such locations could be crucial for city planners, emergency managers, insurance companies, and property administrators for example, since this information could be used to better allocate resources used to prevent and counter terrorism (Nunn, 2007; Piegorsch et al., 2007). As in Equation 5.1, I use a three-stage Bayesian hierarchical modelling framework (`R` code in Appendix C.6):

$$y(\boldsymbol{s}_i,t)|\boldsymbol{\theta},\xi \sim \text{Poisson}(\mu(\boldsymbol{s}_i,t)) \tag{5.2a}$$

$$\log(\mu(\boldsymbol{s}_i,t))|\boldsymbol{\theta} = \eta(\boldsymbol{s}_i,t)|\boldsymbol{\theta} = \beta_0 + \boldsymbol{z}(\boldsymbol{s}_i,t)\boldsymbol{\beta} + \xi(\boldsymbol{s}_i,t) \tag{5.2b}$$

$$\boldsymbol{\theta} \sim p(\boldsymbol{\theta}). \tag{5.2c}$$

[4] Note that the linear predictor $\eta(\boldsymbol{s}_i,t)$ does not include any error term ($\varepsilon \sim N(0,\sigma_\varepsilon)$), since we one assumes that noise in the observations is negligible.

Based on GTD (2014), I consider the observed number of lethal attacks ($y(\boldsymbol{s}_i,t)$ in Equation 5.2a) that occurred in a period of 12 years ($t = 2002,\ldots,2013$) in 6,386 locations $\boldsymbol{s}_i$ within a 0.5° radius of cities' centroids. Since I model a "count" variable ($y(\boldsymbol{s}_i,t)$), it is convenient to aggregate events that occurred in very close locations within identical municipality areas. This has resulted in spatial aggregation reducing the number of observations from 35,917 (Equation 5.1) to 6,386 (Equation 5.2) (`R` code in Appendix C.5). Moreover, I assume that $y(\boldsymbol{s}_i,t)$ follows a Poisson distribution with parameter $\mu(\boldsymbol{s}_i,t)$, with $\log(\mu(\boldsymbol{s}_i,t)) = \eta(\boldsymbol{s}_i,t)$ and expected value $\mathbb{E}(y(\boldsymbol{s}_i,t)) = \mu(\boldsymbol{s}_i,t)$.

Equations 5.2b and 5.1b are structurally identical. As in the Bernoulli model (Section 5.2.2), I use the default option for the stationary model in `R-INLA` as priors on τ and κ through prior distribution on $\log(\tau), \log(\kappa) \sim N(0,1)$, so that $\log(\kappa(\boldsymbol{s})) = \log(\kappa)$ and $\log(\tau(\boldsymbol{s})) = \log(\tau)$. To assess prior sensitivity, I compare the results with alternative priors on the parameters of the Matérn covariance function, which is further discussed in Section 5.3.

5.2.4 Model Validity

Goodness-of-fit

Several approaches can be used to assess and compare Bayesian models (Gelman et al., 2014). A first step consists in measuring the degree to which the results of the models are attributable to the independent variables, which refers to internal validity. Internal validity is commonly assessed through measures of goodness-of-fit (GOF), which typically provide a summary of the discrepancy between the observed and predicted values. A thorough review of GOF methods would go beyond the scope of this thesis. For the sake of conciseness, I highlight few approaches that have been recently used in latent Gaussian spatio-temporal modelling applications.

Interestingly, frequentist model validation approaches are often applied in both Bayesian and non-Bayesian modelling studies (Gneiting, 2008). One relatively straightforward criterion is the root-mean-square error (RMSE), which could be expressed as: $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{p}_i - o_i)^2}$, with n the number of predictions, o_i the i$^{\text{th}}$ observation (e.g. an observed proportion of lethal attacks), and $\hat{p}_i$ the i$^{\text{th}}$ prediction (e.g. a predictive posterior mean of the probability of lethal attacks). The RMSE computes the square root of the mean of the squares of the deviations between the observations and corresponding predictions. This criterion has been used to assessing and comparing the performance of SPDE models in predicting the concentration of particulate matter with an aerodynamic diameter of less than 10 μm (PM_{10}) in Piemonte region, Italy (Cameletti et al., 2011, 2013b).[5]

Other related criteria have been used in the literature, including the mean absolute error (MAE), the relative bias (rBIAS), or the relative mean separation (rMSEP). The MAE can be expressed as: $\frac{1}{n}\sum_{i=1}^{n}|\hat{p}_i - o_i|$. It represents the average absolute difference between the predicted ($\hat{p}_i$) and observed (o_i) values. The rBIAS provides the average of the differences between $\hat{p}_i$ and o_i, relatively to the arithmetic mean of the observations $\bar{o}$. It could be defined as: $\frac{1}{n\bar{o}}\sum_{i=1}^{n}(\hat{p}_i - o_i)$. The rMSEP is formulated as: $\frac{\sum_{i=1}^{n}(\hat{p}_i - o_i)}{\sum_{i=1}^{n}(\bar{\hat{p}} - o_i)}$, with $\bar{\hat{p}}$, the arithmetic mean of the predicted values (Yip, 2010).

The differences between the model predictions and the observations are squared in the RMSE and taken as absolute values in the MAE, respectively. Therefore both RMSE and MAE cannot be negative by definition. In contrast, the rBIAS and rMSEP measure the relative difference between the model predictions and the observations and could therefore take positive, negative or zero values. The units in the RMSE and MAE are identical to the units of the observations and predictions while the rBIAS and rMSEP are unitless. For all these criteria (RMSE, MAR, rBIAS, and rMSEP), better predictions correspond to their corresponding lower values (Stephenson et al., 2004, p. 205).

The RMSE has been frequently used to globally assess the GOF of latent Gaussian spatio-temporal models (see e.g. Alegana et al. (2016); Hartikainen et al. (2011); Ingebrigtsen et al. (2014); Lenzi et al. (2016); Poggio et al. (2016)). The RMSE has the advantage of accurately capturing poor-performing models that exhibit large discrepancies between observations and predictions, since it inflates and penalises large discrepancies unlike MAE for example. Table 5.3 summarises the values of the RMSE and MAE used in assessing the internal validity of the models using different set of covariates (Section 5.3.1). The RMSE and MAE are computed for the Bernoulli and Poisson models, based on all observations and their corresponding predictive values (`R` code in Appendix C.4).

One might expect that the model with the highest number of covariates would exhibit the lowest RMSE values, since more variability in the data could be explained if a higher number of covariates are included. The RMSE values of the Bernoulli models are very similar. The Bernoulli model with 4 covariates — which corresponds to the selected model

[5] The approach suggested in Cameletti et al. (2013b) uses two different sets of data for the fitting and the validation process, respectively. Hence, their method allows assessing the predictive performance (external validity) of the models instead of the internal validity.

Table 5.3 Bernoulli and Poisson models: goodness-of-fit (RMSE and MAE)

Model	Nb. cov.	RMSE	MAE	Fit. time [days]	Nb. obs.
Bernoulli	0	0.561	0.530	9.0	35,917
Bernoulli	3	0.562	0.530	9.2	35,917
Bernoulli	4	0.565	0.482	9.2	35,917
Bernoulli	5	0.566	0.532	9.3	35,917
Poisson	0	25.79	6.422	6.3	6,386
Poisson	3	26.72	6.471	6.5	6,386
Poisson	4	26.68	6.466	6.5	6,386
Poisson	5	26.65	6.464	6.7	6,386

The goodness-of-fit (GOF) estimation of the models is provided by the root-mean-square error (RMSE) and the mean absolute error (MAE). The RMSE and MAE are computed for the Bernoulli and Poisson models using 0, 3, 4, or 5 covariates (column ‘Nb. cov.’). The computation of the MAE and RMSE includes all observations (column ‘Nb. obs.’) during the study period (2002-2013). The computational time (indicative) required to fit the models (column ‘Fit. time [days]’) is based on a 12-core Linux machine (99 GB of RAM) using `R-INLA`.

(Section 5.3.1) — shows the lowest MAE (0.482). The Bernoulli model with 0 covariate exhibits the lowest RMSE (0.561), while the model with 5 covariates exhibits the highest RMSE (0.566). A careful look at the results (Table 5.3) shows that the differences among the models with regard to the RMSE and MAE are very small. This reflect that the GMRF — which is included in all models used in this study — represents the dominant driver of the investigated process. In comparison with the GMRF, the covariates explain much less variation in the data.

The Poisson model with 0 covariate shows the lowest RMSE (25.79) and MAE (6.422), while the model with 3 covariates exhibits the highest RMSE (26.72) and MAE (6.471). Similar to the results of the Bernoulli models, both RMSE and MAE exhibit very similar values among the Poisson models. As mentioned above, little differences in the RMSE and MAE among the models indicate a relatively poor explanatory power of the covariates compared to that of the GMRF.

Model selection

While measuring how good a model fits the observed data is a central step in the modelling process, the model’s predictive ability is all the more important for policy makers who require accurate predictions in order to make informed decisions and implement effective counterterrorism strategies. Hence, the Bernoulli and Poisson models used further in the modelling process are selected according to their predictive accuracy. The estimation of predictive accuracy includes a bias intrinsic to the data used to fit the models (Gelman et al., 2014). Several approaches estimating out-of-sample predictive accuracy using within-sample fits while correcting for this bias include: the Akaike information criterion (AIC) (Akaike, 1974), deviance information criterion (DIC) (Spiegelhalter et al., 2002; van der Linde, 2005), and the widely applicable information criterion (also called Watanabe-Akaike information criterion) (WAIC) (Watanabe, 2010)[6].

Despite that none of the aforementioned criteria is able to compute within a reasonable time an unbiased (or approximately unbiased) and accurate estimation of out-of-sample predictions (Gelman et al., 2014), the WAIC appears superior than the DIC in assessing the predictive accuracy of Bayesian models (Vehtari et al., 2017). Unlike DIC, WAIC is invariant to parametrization and evaluates the predictions based on averaging posterior distribution rather than conditioning on a point estimation (Piironen and Vehtari, 2017) — this latter property is especially problematic for models where Fisher-information matrices are not invertible (so-called *singular* models) (Drton and Plummer, 2017). In addition, WAIC is asymptotically equivalent to Bayesian cross-validation.

Hence, the Bernoulli and Poisson models are selected through backward elimination, which initiates with all available covariates and, at each step, removes one by one the variable, whose loss minimises potential increase of the WAIC. The model with the lowest WAIC is selected. The selected Bernoulli model includes $k = 4$ standardised covariates with corresponding coefficients: satellite night light (β_{lum}), altitude (β_{alt}), ethnicity (β_{greg}), and travel time to the nearest large city (β_{tt}). The selected Poisson model includes $k = 5$ standardised covariates with corresponding coefficients: satellite night light (β_{lum}), altitude (β_{alt}), democracy level (β_{pol}), population density (β_{pop}), and travel time to the nearest large city (β_{tt}) (`R` code in Appendices C.4 and C.7). The results are discussed in further detail in Section 5.3.

[6] For an overview of the AIC, WAIC, and DIC, see e.g. Gelman et al. (2014). For a more comprehensive review, see e.g. Vehtari et al. (2012) or Vehtari et al. (2017).

5.3 Results

5.3.1 Bernoulli and Poisson Models: Posterior Estimates

The Bernoulli model (Table 5.4, columns "Bernoulli") suggests that terrorist attacks are more likely to be lethal far away from large cities, in higher altitude, in locations with higher ethnic diversity ($CI_{95\%}\ \beta_{tt}, \beta_{greg}, \beta_{alt} > 0$), and are less likely in areas with higher human activity ($CI_{95\%}\ \beta_{lum} < 0$) (`R` code in Appendices C.4 and C.7).

As an illustration, I compare the effect of an hypothetical 50% increase in the mean of luminosity (lum) on the probability of encountering lethal attacks (π). Since the covariates $\boldsymbol{\beta}$ are standardised, particular care should be taken to estimate such effect. Recall that NOAA (2014)'s satellite data on lights at night provide non-standardised values of luminosity from 0 (min) to 63 (max). Based on 35,917 observations used in the Bernoulli model, a 50% increase in the mean of luminosity corresponds to an increase from 34.3 to 51.5, or equivalently, from 0 to 0.73 on a standardised scale. In the benchmark scenario, all predictors, including lum and other covariates tt, $greg$, alt, and the GMRF (ξ) are held equal to 0. Hence, the linear predictor η (Equation 5.1b) equals 0 and $\pi = 1/(1+\exp(-\eta)) = 1/(1+1) = 0.5$. In the scenario including a 50% increase in the mean of luminosity, $lum = 0.73$, $\eta = \beta_{lum} \times lum = -0.11 \times 0.73 = -0.0803$, given that $\beta_{lum} = -0.11$ (Table 5.4, columns "Bernoulli"). Hence, $\pi = 1/(1+\exp-(-0.0803)) \cong 0.48$. As a result, an increase in the mean of luminosity by 50% decreases the probability of lethal attacks by approximately 2%.

In contrast, the Poisson model (Table 5.4, columns "Poisson") suggests that more economically developed areas, and locations with higher democratic levels ($CI_{95\%}\ \beta_{lum}, \beta_{pol} > 0$) and close to large cities ($CI_{95\%}\ \beta_{tt} < 0$) are more likely to encounter a higher number of lethal attacks. However, there is no significant relationship between altitude, population density, and the number of lethal attacks ($0 \in CI_{95\%}\ \beta_{alt}, \beta_{pop}$). The interpretation of these results is discussed in further detail in Section 5.4. Note that the results from the two models cannot be directly compared since they are based on a different number of observations, set of covariates, and spatial aggregation.

As with the Bernoulli model, I illustrate the effect of a 50% increase in the mean of luminosity on the expected number of lethal attacks (μ) estimated by the Poisson model, with all predictors, including lum and the other covariates tt, alt, pol, pop, and the GMRF (ξ) held equal to 0. In the benchmark scenario, η (Equation 5.2b) equals to 0, therefore

Table 5.4 Posterior estimates for the Bernoulli and Poisson models.

	Bernoulli (n=35,917)			Poisson (n=6,386)		
	mean	sd	95% CI	mean	sd	95% CI
Cov. (β)						
β_0	-0.58	0.13	(-0.83; -0.32)	-1.37	0.09	(-1.55; -1.19)
β_{lum}	-0.11	0.02	(-0.15; -0.06)	0.51	0.02	(0.47; 0.54)
β_{tt}	0.06	0.02	(0.02; 0.09)	-0.38	0.02	(-0.42; -0.34)
β_{greg}	0.04	0.02	(0.003; 0.08)			
β_{alt}	0.08	0.03	(0.03; 0.13)	0.03	0.02	(-0.001; 0.07)
β_{pol}				0.42	0.01	(0.39; 0.45)
β_{pop}				0.009	0.02	(-0.03; 0.04)
GMRF (ξ)						
ρ (AR1)	0.91	0.01	(0.88; 0.91)	0.91	0.01	(0.90; 0.93)
σ^2	2.27		(1.83; 2.73)	4.61		(4.11; 5.13)
r [km]	779		(643; 915)	238		(208; 267)

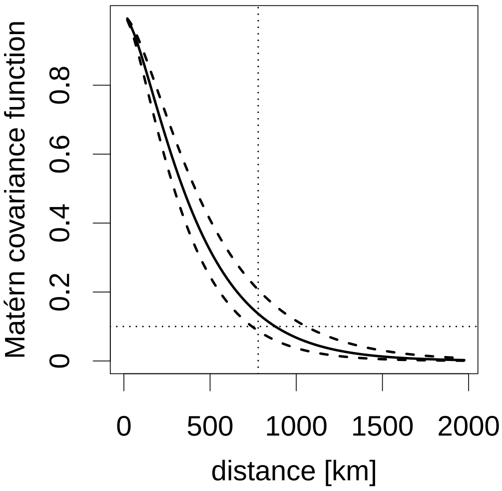
Mean, standard deviation, and 95% credible intervals (CI) of the intercept β_0, the coefficients of the standardised covariates ($\boldsymbol{\beta}$), the temporal (ρ) of the GMRF (ξ) and the spatial parameters of the Matérn correlation function (marginal variance σ^2 and range r) estimated in the Bernoulli ($n = 35,917$) and Poisson ($n = 6,386$) models. Note that the 95% credible intervals (CI) associated with σ^2 and range r correspond to 95% highest probability density intervals.

the expected number of lethal attacks $\mu = \exp(0) = 1$. Based on 6,386 observations used in the Poisson model, a 50% increase in the mean of luminosity corresponds to an increase from 7.45 to 11.2, or equivalently, from 0 to 0.25 in the corresponding standardised scale. Hence, $\eta = (\beta_{lum} \times lum)$, with $\beta_{lum} = 0.51$ (Table 5.4, columns "Poisson"). I obtain $\mu = \exp(0.51 \times 0.25) \cong 1.14$. Therefore, an increase in the mean of luminosity by 50% increases the expected number of lethal attacks by approximately 0.14.

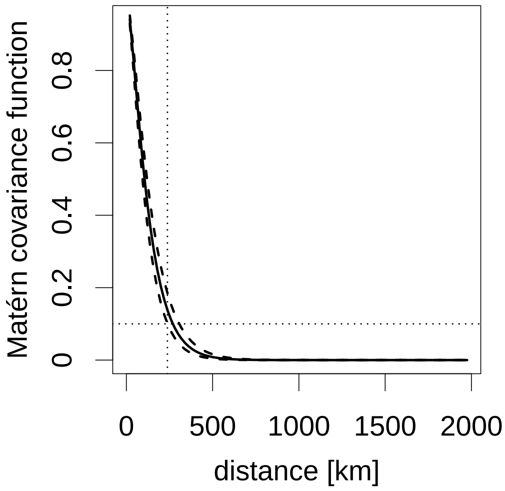
The analysis of the GMRF parameters (Table 5.4 and Figure 5.2) provides key insight into the spatial dependence structures of both the lethality of terrorism (Bernoulli model) and the frequency of lethal terrorism (Poisson model). The marginal variance σ^2 provides a measure of the variability of the Matérn covariance function. The range r is inversely proportional to κ, and also easier to interpret[7]. More specifically, r corresponds to the distance after which spatial correlation becomes progressively negligible.

Thus, the results suggest that the probability of lethal terrorist attacks is highly spatially correlated at short distances and the correlation decreases until reaching $\approx$ 779 km (posterior

[7] The range r is empirically estimated as: $\frac{\sqrt{8\nu}}{\kappa}$, assuming an order of the modified Bessel function $\nu > \frac{1}{2}$ (Lindgren et al., 2011).



(a) Bernoulli model: correlation function



(b) Poisson model: correlation function

Fig. 5.2 Matérn correlation function (*solid line*) with parameters $\nu = 1$ and $CI_{95\%}$ (*dashed lines*) for Bernoulli model (Figure 5.2a) with posterior mean range $r \approx 779$ km and Poisson model (Figure 5.2b) with posterior mean range $r \approx 238$ km. The posterior mean range (*vertical dotted line*) corresponds to correlation ≈ 0.1 (*horizontal dotted line*).

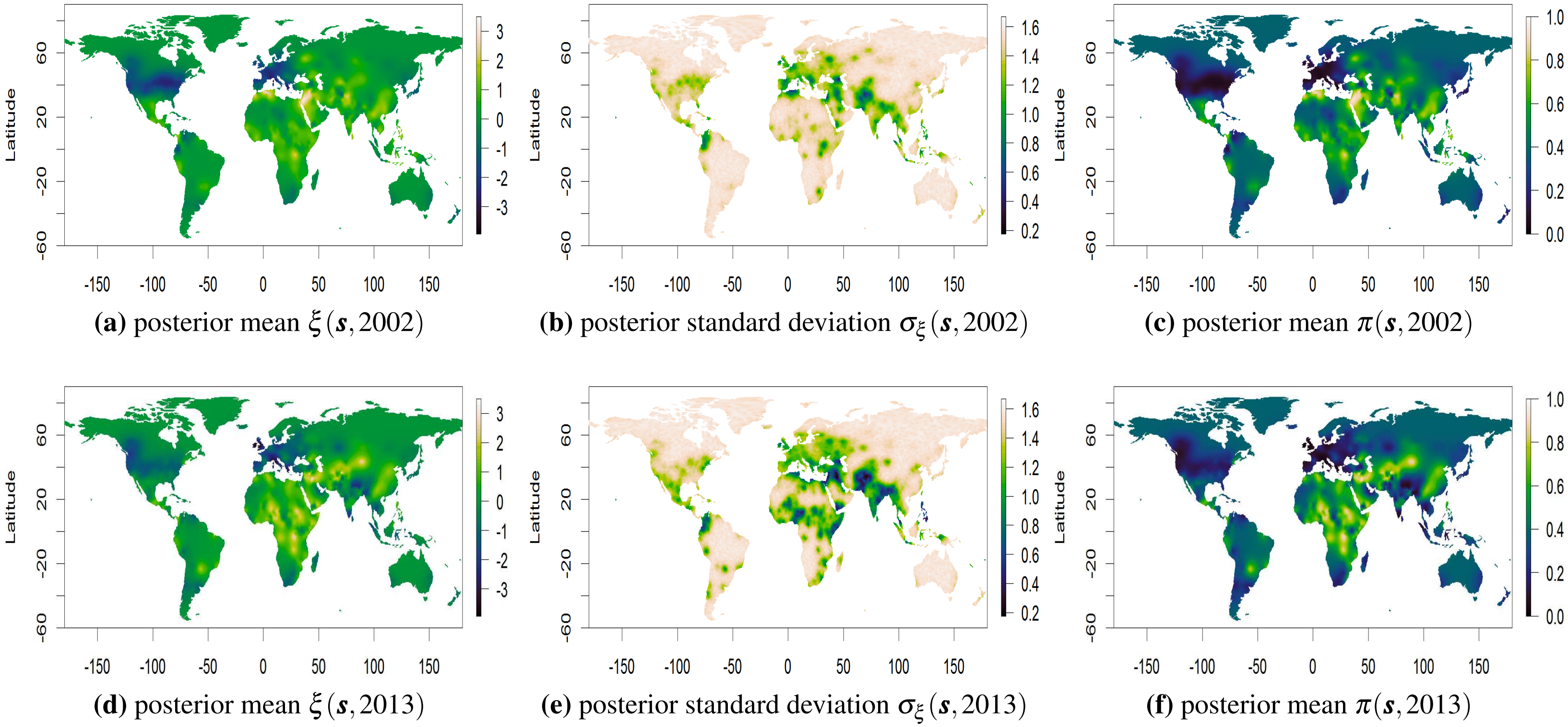
mean of r) (see *vertical dotted line* in Figure 5.2a). The spatial correlation with regard to the number of lethal attacks operates within more limited distances; the mean range is more then three times smaller than in the Bernoulli model (posterior mean of $r \approx 238$ km, see *vertical dotted line* in Figure 5.2b).

Similarities in the probability of lethal terrorist attacks are present within a radius three times higher than those observed with regard to the number of lethal attacks. For example, consider city A, which encountered a high proportion of lethal attack and also a high number of lethal attacks. The results suggest that cities very close (distance < 238 km) to city A are likely to encounter a large number of lethal attacks and a high proportion of lethal attacks. One expects that the proportion of lethal attack observed in cities far (238 km$>$ distance < 779 km) from city A is likely to be similar. However, the number of lethal attacks should not be associated with those observed in city A. Furthermore, both the proportion and the number of lethal attacks observed in cities situated at over 779 km from city A are not likely to be similar than those observed in city A.

5.3.2 Spatial Dynamics of Lethal Terrorism

As a further step, I examine changes in the estimated posterior marginals in both space and time. The uncertainty in the values of the GMRF is expressed through the posterior standard deviation of the random field ($\sigma_{\xi}(\boldsymbol{s},t)$)[8]. Providing an accurate measure of uncertainty is crucial for policy-makers to make informed decisions (Zammit-Mangion et al., 2013, p. 64) or to evaluate the impact of counterterrorism policies (Perl, 2007) for example. In effect, high values of $\sigma_{\xi}(\boldsymbol{s},t)$ suggest a poor estimation of $\xi(\boldsymbol{s},t)$, and hence, a high uncertainty in the estimation of the dependent variable (e.g. probability or number of lethal attacks).

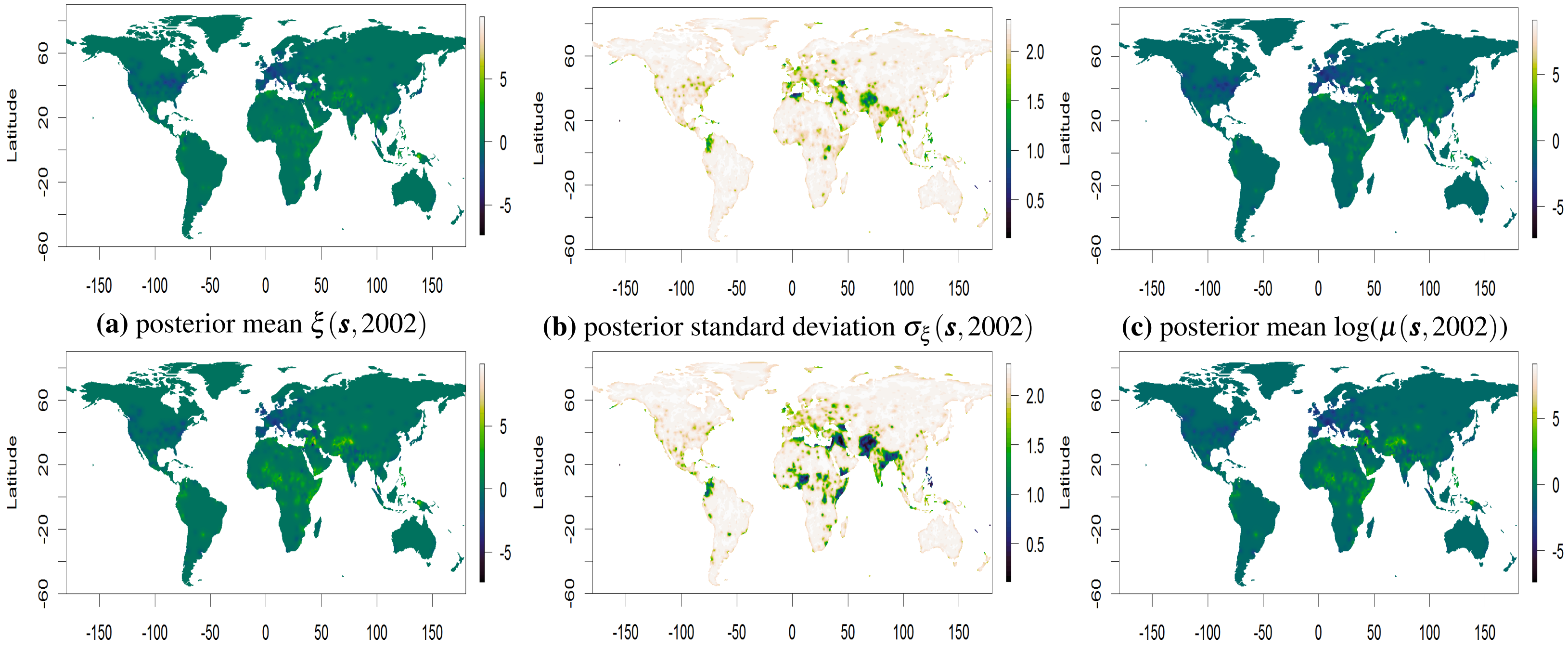
[8] Note that the posterior standard deviation of the GMRF σ_{ξ} should not be confused with standard deviation σ, the latter represents a parameter of the Matérn covariance function (Equation 3.32).



(a) posterior mean $\xi(\boldsymbol{s}, 2002)$
(b) posterior standard deviation $\sigma_{\xi}(\boldsymbol{s}, 2002)$
(c) posterior mean $\pi(\boldsymbol{s}, 2002)$

(d) posterior mean $\xi(\boldsymbol{s}, 2013)$
(e) posterior standard deviation $\sigma_{\xi}(\boldsymbol{s}, 2013)$
(f) posterior mean $\pi(\boldsymbol{s}, 2013)$

Fig. 5.3 Bernoulli model: posterior surface maps (2002-2013). GMRF posterior mean $\xi(\boldsymbol{s},t)$ (*left*), posterior standard deviation $\sigma_{\xi}(\boldsymbol{s},t)$ (*centre*), and posterior mean probability of lethal attack $\pi(\boldsymbol{s},t)$ (*right*) estimated in the 9,697 locations of mesh vertices and interpolated over a $1,444 \times 724$ grid. Illustrative projected maps provide values on land surface for years $t = 2002$ (*top*) and $t = 2013$ (*bottom*).



(a) posterior mean $\xi(\boldsymbol{s}, 2002)$ **(b)** posterior standard deviation $\sigma_{\xi}(\boldsymbol{s}, 2002)$ **(c)** posterior mean $\log(\mu(\boldsymbol{s}, 2002))$

(d) posterior mean $\xi(\boldsymbol{s}, 2013)$ **(e)** posterior standard deviation $\sigma_{\xi}(\boldsymbol{s}, 2013)$ **(f)** posterior mean $\log(\mu(\boldsymbol{s}, 2013))$

Fig. 5.4 Poisson model: posterior surface maps (2002-2013). GMRF posterior mean $\xi(\boldsymbol{s},t)$ (*left*), posterior standard deviation $\sigma_{\xi}(\boldsymbol{s},t)$ (*centre*), and (log) posterior mean expected number of lethal attacks $\mu(\boldsymbol{s},t)$ (*right*) estimated in the 9,697 locations of mesh vertices and interpolated over a $1,444 \times 724$ grid on land surface. Illustrative projected maps provide values for years $t = 2002$ (*top*) and $t = 2013$ (*bottom*).

Figures 5.3 and 5.4 show the posterior mean values of the GMRF (*left*) and highlight high uncertainty of its estimation in Siberia, the Amazonian region, Central Australia, or Greenland. High values of $\sigma_\xi(\boldsymbol{s},t)$ (*centre*) are explained by the absence and scarcity of data in these regions. In contrast, several regions in South America, Africa, Gulf Peninsula, India and Pakistan show an increase of the lethality of terrorism and the frequency of lethal terrorist attacks between 2002 (*top*) and 2013 (*bottom*), along with low values of $\sigma_\xi(\boldsymbol{s},t)$ (*centre*). As expected, uncertainty is much lower in regions which encountered a high number of terrorist attacks.

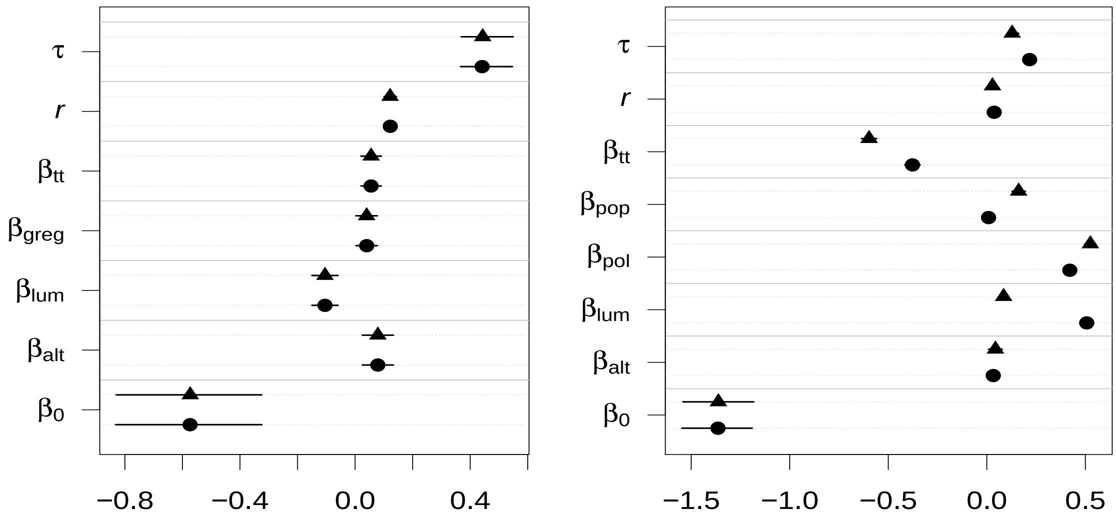
5.3.3 Prior and Mesh Sensitivity Analysis

The present study used the default Gaussian priors (multivariate Normal) provided by `R-INLA` for the parameters of the Matérn covariance function (κ,τ). As a robustness test, I ran a prior sensitivity analysis for both the Bernoulli and the Poisson models changing the prior distribution of σ^2 and range r, whose quantities can be more easily interpreted, and therefore attributed a prior distribution. With $\alpha = 2$, σ^2 and range r can be obtained through transformation, merely $r = \frac{\sqrt{8}}{\kappa}$ and $\sigma = \frac{1}{\sqrt{4\pi\tau}}$ (Lindgren and Rue, 2015; Lindgren et al., 2011)

For both the Bernoulli and the Poisson models, I set $\sigma^2 = 50$, which corresponds to a relatively large variance of the Matérn covariance function, since I assume that the spatial structure might exhibit considerable variation among areas that encountered a high number of lethal terrorist attacks (e.g. some locations in Iraq, Pakistan or Afghanistan) and those that did encounter only a few or none (e.g. some locations in Portugal, Brazil or Alaska).

Assuming that the lethality of terrorist attacks can spread over relatively large areas (e.g. through demonstration and imitation processes promoted by the media (Brosius and Weimann, 1991; Brynjar and Skjølberg, 2000; Enders et al., 1992), I set $r = 500$ km in the Bernoulli model. In contrast, I set $r = 100$ km in the Poisson model, since I believe that the number of lethal attacks is very specific to the characteristics related to the close neighbourhood in which terrorism occurs. The number of potential high-value, symbolic targets, human and public targets can vary widely across distant areas.

For example, one might reasonably assume that Baghdad shares important similarities with close Iraqi cities such as Abu-Grahib or Al-Fallujah (approximately 30-60 km from Baghdad). However, distant cities such as Al-Kasrah, Iraq (approximately 300 km from



(a) Bernoulli model: posterior estimates

(b) Poisson model: posterior estimates

Fig. 5.5 Bernoulli (Figure 5.5a) and Poisson models (Figure 5.5b): robustness to prior changes in spatial parameters. Estimation of the mean (default prior models: *point*; modified priors: *triangle*) and the 95% credible intervals (*line segment*) of the posterior distribution of the parameters: intercept β_0, covariates $\boldsymbol{\beta}$, and parameters r and τ of the Matérn covariance function. Modified prior values: Bernoulli ($\sigma^2 = 50$; $r = 500$ km); Poisson ($\sigma^2 = 50$; $r = 100$ km).

Baghdad) might exhibit important different characteristics, so the numbers of lethal attacks are likely to be independent of the levels observed in Baghdad.

The estimation of the posterior mean and the credible intervals of the coefficients $\boldsymbol{\beta}$ and parameters of the Matérn covariance function (τ, r) is illustrated for both Bernoulli (Figure 5.5a) and Poisson models (Figure 5.5b), where default priors (*point*) and modified priors (*triangle*) are specified. The parameters τ and r are not affected by a change in prior in both Bernoulli and Poisson models, albeit the estimated value of the mean of τ is less robust in the Poisson model. In the Bernoulli model, both posterior mean and credible intervals of all parameters are almost identical when estimated with the default or modified priors.

In the Poisson models, the direction of the effect of the estimated coefficients is not affected by changes in prior, albeit the estimated values of the mean of the parameters (especially β_{tt} and β_{lum}) are less robust. A higher sensitivity to change in prior is expected in the Poisson model since its number of observations ($n = 6{,}386$) is considerably reduced compared to the Bernoulli model ($n = 35{,}917$), thus the prior distribution has a higher influence on the posterior estimation, as illustrated in these results.

Changes in the spatial resolution of the mesh may have an impact on the results, a well-known problem referred to as the “modifiable areal unit problem” (MAUP) (Holt et al., 1996). As an additional robustness test, I consider two alternative meshes: *mesh*2 (Figure 5.6, *centre*) and *mesh*3 (Figure 5.6, *right*). Recall that the original mesh (Figure 5.6, *left*) used to fit all models is composed of 9,697 vertices. The alternative meshes used for robustness test are coarser, with 2,157 and 1,341 vertices for *mesh*2 and *mesh*3, respectively. I exclude testing robustness of models with meshes that count a number of vertices higher than the original mesh since computational costs will rise without providing any substantial gain given the relatively coarse spatial accuracy of the investigated phenomenon (for further information on the spatial resolution of GTD, see Section 4.1.3).

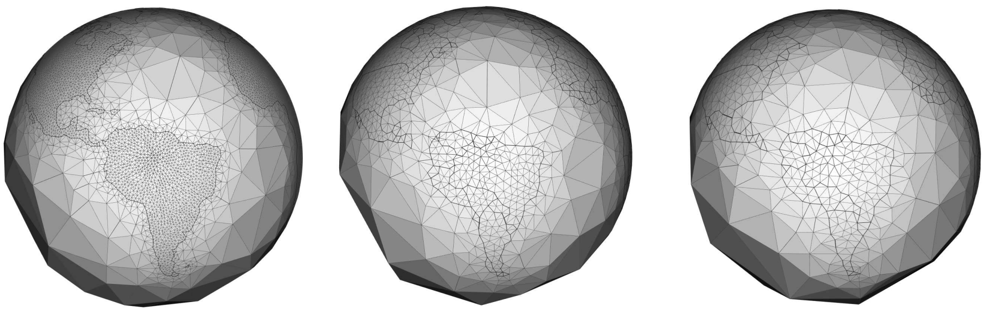


Fig. 5.6 Robustness test: change in the spatial resolution of the Constrained Refined Delaunay Triangulation (mesh). The original mesh (*left*) counts 9,697 vertices. For robustness test, meshes with rougher spatial resolution are used: *mesh*2 (*centre*) and *mesh*3 (*right*) counts 2,157 and 1,341 vertices, respectively.

All effects are robust to changes in mesh size, except altitude (β_{alt}) and population density (β_{pop}) (`R` code in Appendices C.4 and C.7). In the Bernoulli model using the original mesh, β_{alt} is positive and significant ($CI_{95\%}\ \beta_{alt} > 0$). In the corresponding models using *mesh*2 and *mesh*3, β_{alt} losses its significance ($0 \in CI_{95\%}\ \beta_{alt}$). In the Poisson model using the original mesh, β_{alt} is not significant ($0 \in CI_{95\%}\ \beta_{alt}$) but becomes negative ($CI_{95\%}\ \beta_{alt} < 0$) in the model using *mesh*2 (in the model using *mesh*3, β_{alt} remains not significant).

Population density is affected by changes in the mesh size only in the Poisson model. In the model using the original mesh, β_{pop} is not significant ($0 \in CI_{95\%}\ \beta_{pop}$) but becomes positive and significant ($CI_{95\%}\ \beta_{pop} > 0$) in the models using *mesh*2 and *mesh*3. Despite that most results are not affected by changes in mesh size, our sensitivity analysis indicates that some covariate coefficients appear less robust. As a general remark, I recommend choosing a mesh with the smallest number of vertices, which allows capturing spatial variation of the investigated phenomenon in the area of interest at the desired spatial scale. In this study, the original mesh (9,697 vertices) appears to best fit the spatial accuracy of the data, while *mesh*2 and *mesh*3 might be too coarse to capture fine-scale variability relative to the investigated process.

5.4 Conclusion

This study proposed a Bayesian hierarchical framework to model both the *probability* and *number* of lethal terrorist attacks that occurred across the world between 2002 and 2013. The statistical framework integrates spatial and temporal dependencies through a GMRF whose parameters have been estimated with `R-INLA`. Local-scale effects of some covariates that explain the *probability* and *number* of lethal terrorist attacks worldwide have been successfully captured.

Most country-level studies have not found significant linear relationship between the number of terrorist attacks and economic variables (Abadie, 2006; Drakos and Gofas, 2006b; Gassebner and Luechinger, 2011; Krueger and Laitin, 2008; Krueger and Maleckova, 2003). On a local scale, I showed that more economically developed areas tend to encounter more lethal attacks, which provides support to the theory advanced by Piazza (2006). The author suggests that more economically developed and literate societies with high standards of living may exhibit more lucrative targets, and therefore are expected to be more targeted by terrorist attacks, including lethal ones. Economically developed areas tend to provide a high level of protection of human targets, which may explain a lower proportion of lethal attacks despite being more prone to terrorism in general.

Similar to most country-level studies (Abadie, 2006; Eubank and Weinberg, 2001; Li, 2005; Schmid, 1992), I found that areas in democratic countries tend to encounter more lethal attacks compared to those in autocracies. The presence of freedom of speech, movement, and association in democratic countries might reduce the costs of conducting terrorist activities compared to those in autocratic countries (Li, 2005). In line with country-level (Gassebner and Luechinger, 2011; Kurrild-Klitgaard et al., 2006) and sub-national (Nemeth et al., 2014) findings, terrorist attacks are more likely to be lethal in ethnically diverse locations, perhaps due to stronger ethnic tensions (Basuchoudhary and Shughart, 2010)[9].

While most country-level studies found a significant positive linear relationship between population density and the number of terrorist attacks (Coaffee, 2010; Crenshaw, 1981; Ross, 1993; Savitch and Ardashev, 2001; Swanstrom, 2002), I did not find evidence that terrorist attacks are more likely to be lethal or that lethal attacks are more frequent in densely populated areas. Nevertheless, the effect of population on the number of lethal attacks becomes positive when coarser meshes are used (see Section 5.3.3).

The Euclidean distance from terrorist events to the nearest large city is positively associated with the lethality of terrorist attacks, while being negatively associated with the number of lethal attacks. Since targets are usually less secure in small cities and rural areas, this might facilitate deadly terrorist operations far from large urban centres, which is consistent with the findings. Reversely, large cities are expected to offer greater anonymity and larger recruitment pools, which might open the door for a higher number of lethal attacks. Also, lethal attacks within or close to large cities may have an impact on a larger audience, which is often a desired outcome (Laqueur, 1999, p. 41; Crenshaw, 1990, p. 115; Savitch and Ardashev, 2001). Furthermore, terrorists benefit from high density communication networks (road and rail) in large cities to move freely and rapidly from and to target points (Heyman and Mickolus, 1980 ; Wilkinson, 1979, p. 189). It is also not uncommon that terrorists target communication network infrastructure, as exemplified by the March 11, 2004 simultaneous attacks on several commuter trains in Madrid, Spain, which killed 191 people (Los Angeles Times, 2014).

Even though the number of lethal attacks itself does not appear to be associated with altitude, I found that terrorist attacks, when they occur, tend to be more lethal in higher altitude. Terrorists might be less constrained by governmental forces during terrorist operations launched in less accessible regions, such as mountains, which can provide safe havens to terrorists (Abadie, 2006; Ross, 1993). Terrorist groups can therefore benefit from knowledge of "rough" terrain (mountainous regions) to defeat the enemy, as illustrated by the successful attacks carried out by the Mujahedeen groups against the Soviet Union, and later, the Taliban against NATO (Buhaug et al., 2009). Therefore, one might reasonably assume that

[9] Note that, as pointed out in Esteban et al. (2012), one should acknowledge that ethnic diversity is only one possible measure of ethnic division among others, including *ethnic fractionalization* and *ethnic polarization*. Further analysis using different measures of ethnic division is required in order to assess the role of ethnic division sensu lato.

terrorists increase their killing efficiency in mountains, which is reflected in a higher proportion of "successful" lethal attacks. However, since human targets are usually more scarce in mountainous regions, the advantage of the terrain might not sufficiently compensate for the lack of targets, which in turn reduces the number of possible lethal attacks that can be planned and carried out. Complementary analysis would be required in order to confirm the plausibility of this interpretation of the results.

As with any statistical analyses of complex social phenomena, the outcome of this study should be taken with caution. Since this study aims to investigate terrorism across the entire world and at high spatial resolution, the availability of suitable covariates is limited. As a result, this work has ineluctably omitted numerous relevant drivers of both the lethality of terrorist attacks and their frequency, which include characteristics of each member of terrorist groups, ideology, beliefs, and cultural factors (Crenshaw, 1983, p. 29; Wilkinson, 1990, p. 151; Brynjar and Skjølberg, 2000; Richardson, 2006, pp. 92-93), or reciprocal interactions between counterterrorism and terrorism (English, 2010; Hoffman, 2002) for example. Despite the fact that the models do not allow for estimating the marginal effect of each potential unobserved factor, their aggregated effect has been however taken into account through the space-time dependence structure represented by the GMRF (ξ in equations (5.1b) and (5.2b)).

Second, one could reasonably expect some spatial variability in the probability and number of lethal terrorist attacks, especially within large cities that are regularly targeted by terrorists. However, since terrorist events from GTD are reported at the centroid of the nearest city in which they occurred, spatial variability within cities cannot be captured. Moreover, I assume that the spatial correlation in both the probability and number of lethal attacks depends only on the distance between the locations of terrorist attacks (stationarity) and is invariant to rotation (isotropy). This assumption might be too restrictive, since it would be equally reasonable to assume that the spatial correlation related to mass-casualty attacks extends into a larger spatial range, via a broader diffusion through media for example. Further studies might investigate the use of non-stationary models, which are currently being developed for the model class that may be fitted with INLA (Lindgren et al., 2011).

Third, the temporal unit exhibits limitations as well, since the study period is discretised into 12 years (2002-2013), even though GTD provides day, month, and year for most events. Access to more computational power may allow further analysis to investigate variation in a monthly or weekly basis. Moreover, for computational reasons, I assume no interaction between spatial and temporal dependencies in the models (i.e. *separable* space-time models), where the covariance structure can be written as the product of a purely spatial and a purely temporal covariance function for all space-time locations (Gneiting et al., 2006). In this research work, the random field (ξ) follows a simple autoregressive process (AR(1)) in time. In non-separable models, the dependencies structure in both space and time is usually highly complex (Harvill, 2010), and therefore more computationally demanding.

Fourth, subjective choices have been made throughout the entire modelling process, which might affect both the internal and external validity of our results. A major concern is the absence of consensus on the definition of terrorism (Beck and Miner, 2013; Jackson, 2016), and subjectivity is therefore inevitable (Hoffman, 2006, p. 23) (Section 2.1). In line with English (2010, pp. 24-25), I agree that is all the more important that studies on terrorism must clearly state how terrorism is understood. Accordingly, I use data from GTD, which uses a consistent definition used to classify acts as terrorist events perpetrated by non-state actors throughout the study period.

Moreover, as with any Bayesian analysis, this study involves a degree of subjectivity with regard to the choice of priors. Because of the relatively large number of observations, I am confident that the choice of priors does not influence the results. Furthermore, while the mesh may affect some covariates, most results are robust to change in mesh size, as confirmed by the sensitivity analysis (Section 5.3.3). The original mesh size has been chosen consistently with the spatial resolution of the data, which brings confidence that the results are robust.

Despite its aforementioned shortcomings, this study has proposed a rigorous framework to investigating the spatial dynamics of the probability and number of lethal terrorist attacks across the world and on a local scale. It has highlighted the role of major drivers of terrorism on a local scale and provided a measure of the uncertainty in the predictions. This research may provide complementary tools to enhance the efficacy of preventive counterterrorism policies. Ultimately, by assessing theories at a local scale, this study contributes to better understanding the spatial dynamics of lethal terrorism.

Chapter 6

Detecting Hotspot and Diffusion Processes of Lethal Terrorism

Intensive theoretical and empirical studies have been conducted since the 1980s to better understand the spatio-temporal processes observed in terrorism and related phenomena. Thus far, the literature has focused its investigation on *hotspot* and *diffusion* processes. However, scholars in terrorism have not made the most of recent space-time modelling techniques and failed to accurately quantify uncertainty. Moreover, studies at a national level did not capture local processes while local-level works have been focused on case studies, which restricted the generalisability of their results (Berrebi and Lakdawalla, 2007; Brown et al., 2004; Gao et al., 2013; LaFree et al., 2010, 2009; Nunn, 2007; Öcal and Yildirim, 2010; Piegorsch et al., 2007) (Section 2.2).

Drawing from the results of the modelling processes (Chapter 5), I suggest a Bayesian approach to detecting the local dynamics of terrorism over time and across the world divided into $1,440 \times 720$ regular grid-cells. I focus on identifying the processes of hotspot and diffusion of lethal terrorism that might be useful in the design and implementation of counterterrorism policies. Thus, the wide scope of this study along with its systematic approach allows generalisation, and thus, proposes a convenient framework to assessing theories on a local scale. Ultimately, this investigation aims to provide a better understanding of the local dynamics of lethal terrorism.

Section 6.1 suggests a definition of hotspot and escalation processes of the lethality and the number of lethal terrorist events. The suggested definition makes use of a measure of uncertainty through the estimation of the 95% credible intervals of the posterior probability of the lethality (Bernoulli model) and the number of lethal terrorist events (Poisson model). Areas that are likely to encounter a high proportion or a high number of lethal attacks are identified across the world. Furthermore, the spatial dynamics of hotspot and escalation processes are illustrated and analysed. Thus, the performance in detecting hotspot and escalation is assessed.

Similarly, Section 6.2 provides a definition of diffusion processes (including also negative diffusion, also-called *dissipation*) of the lethality and the number of lethal terrorist events, based on the measure of uncertainty computed through the estimation of the 95% credible intervals of the posterior probability of the lethality and the number of lethal terrorist events at successive time intervals. Hence, it highlights diffusion areas across the world and at high spatial resolution, and provides an assessment of the detection of diffusion processes. Section 6.3 assesses theories that have been advanced in the literature to explain hotspots and diffusion processes and identifies the main limits of the approach. Section 6.4 summarises the findings and suggests possible extensions of this present study.

6.1 Detection of Hotspot and Escalation

6.1.1 Concept and Definition

In this Section, I suggest a method which uses the uncertainty of the posterior probability distributions in order to identify hotspots in the probability of lethal attack $\pi(\cdot)$ and in the expected number of lethal events $\mu(\cdot)$, respectively. Below, I suggest a definition of hotspots of "high" *probability* of lethal attacks, which can be easily adjusted for hotspots of "high" *number* of lethal attacks (see below).

Definition 6.1. (**Hotspot of probability of lethal attack**) *Let* $\pi_{0.025}(a_{\mathbf{s}_i,t})$ *be the lower bound of the 95% credible intervals (CI) of the posterior mean probability of the attack(s) to be lethal in space-time location* $a_{\mathbf{s}_i,t} \subseteq A_{\mathbf{s},t} \subseteq \mathscr{D}$*, with space-time domain* $\mathscr{D}$*, the observed data* $\mathbf{y}$*, and threshold* υ*. A hotspot of lethality* $\mathscr{H}_{A_{\mathbf{s},t}}$ *covers the area* $A_{\mathbf{s},t}$ *formed by elements* $\{a_{\mathbf{s}_1,t},\ldots,a_{\mathbf{s}_i,t},\ldots,a_{\mathbf{s}_n,t}\}$ *contiguous*[1] *in space* $\mathbf{s}_i = \{\mathbf{s}_1,\ldots,\mathbf{s}_i,\ldots,\mathbf{s}_n\} \subseteq S$ *at a given time*

[1] Note that different approaches can be used to define contiguity (Section 3.3.1). In this study, the spatial domain (world) is discretised into a lattice formed by $1,440 \times 720$ regular grid-cells, where contiguous locations are understood as cells that share a common vertex or edge (first-order "queen" neighbourhood, see Figure 3.6).

$t \in T$, and that satisfy for each contiguous $a_{\mathbf{s}_i,t} \subseteq A_{\mathbf{s},t}$:

$$\mathscr{H}_{A_{\mathbf{s},t}} = \pi_{0.025}(a_{\mathbf{s}_i,t}) > \upsilon|\mathbf{y}, \quad a_{\mathbf{s}_i,t} \subseteq A_{\mathbf{s},t} \subseteq \mathscr{D}. \tag{6.1}$$

Recall from the Bernoulli modelling framework (Section 5.2.2, Equation 5.1) that the posterior mean probability is computed for each spatio-temporal locations $(\mathbf{s}_i, t)$, such that: $\pi(a_{\mathbf{s}_i,t}) = \text{logit}^{-1}(\beta_0 + \boldsymbol{z}(a_{\mathbf{s}_i,t})\boldsymbol{\beta} + \xi(a_{\mathbf{s}_i,t}))$, with β_0 the intercept, $\boldsymbol{\beta}$ the coefficients of the covariates $\boldsymbol{z}(a_{\mathbf{s}_i,t})$, and $\xi(a_{\mathbf{s}_i,t})$ the GMRF. In order to reduce the risk of identifying areas that are not likely to be "true" hotspots, one estimates the lower bound of the 95% CI of the posterior mean probability $\pi_{0.025}(a_{\mathbf{s}_i,t})$. As a result, this approach provides a point estimate of $\pi_{0.025}(a_{\mathbf{s}_i,t})$ that is unlikely to be lower than the true value. Hence, I suggest using $\upsilon = 0.5$, which corresponds to hotspot areas that are more likely to encounter lethal rather than non-lethal terrorist attacks. Hotspots of probability of lethal attacks are identified across the world, and illustrated from 2002 to 2013 in Africa and in the Middle East (Figure 6.1), which are the regions most affected by lethal terrorism.

Furthermore, I identify hotspots of "high" number of lethal attacks. Instead of using the Bernoulli modelling framework to estimate the lower bound of the 95% CI of the posterior mean probability $\pi_{0.025}(\cdot)$ of lethal attack, I compute the lower bound of the 95% CI of the posterior expected number of lethal events $\mu_{0.025}(\cdot)$ from the Poisson models (Section 5.2.3, Equation 5.2). As a result, the value of υ is adapted since it corresponds to a threshold of a count in the number of lethal attacks rather than a probability of lethal attack. Hence, I suggest using $\upsilon = 5$, which highlights cells with an expected number of lethal attacks greater than five (Figure 6.2). This threshold corresponds to the 90^{th} percentile of the number of lethal attacks observed in the sample ($n = 6,386$), but could be adjusted to any desired level.

6.1.2 Spatial Dynamics of Hotspot

The identification of hotspots of the probability and the number of lethal attacks is highly valuable since it allows highlighting areas more vulnerable to lethal terrorism, which calls for increased vigilance. The detection of hotspots uses the aforementioned definition (Section 6.1.1) applied on a regular lattice over the world ($1,440 \times 720$ grid-cells). Hence, one may observe intense activity of lethal terrorism from 2002 to 2013 in specific regions, including Iraq, materialised into several hotspots of high intensity (Figures 6.1 and 6.2).

The presence of hotspots in these areas reflects the intensification of terrorist activity that followed the invasion of Iraq in 2003 (*Operation Iraqi Freedom*) carried out by the US-led coalition. As early as 2002, the CIA director George Tenet and the US National Intelligence Council warned the US and UK government that radicalization and terrorism activity will increase in Iraq and outside its borders due to the invasion of Iraq. This was later confirmed by Britain's Royal Institute of International Affairs (Chatham House) shortly after the 2005 London bombings, and also by various reports from Israeli think tanks, as well as Saudi and French Intelligence (Chomsky, 2006, pp. 18-21).

Similarly, there are hotspots of probability and number of lethal attacks in some regions in Afghanistan (especially near latitude $\approx 35°$, longitude $\approx 66°$). Indeed, Afghanistan has been the theatre of numerous terrorist attacks since the US-led invasion in 2001 following 9/11. From 2003 to 2013, 3,539 terrorist attacks occurred mainly in the centre-east of Afghanistan, including the cities of Kabul, Jalalabad, Khogyani, and Sabari. From 2003 to 2013, 243 events (of which 157 were lethal) occurred in Kabul only. Most of these attacks were perpetrated by the Taliban. Even after the Taliban's withdrawal from Kabul in November 2001, lethal terrorist attacks did not cease in the city and within the country (Aljazeera, 2009). More particularly, highly lethal suicide bombings intensified from 2006 to 2013 (and further) (GTD, 2014).

Figure 6.1 highlights an interesting pattern, which emerges from 2010 to 2013 in Yemen, close to the city of Aden (latitude $\approx 13°$, longitude $\approx 45°$). In 2010, a hotspot of probability and number of lethal attacks is observed in the area of Aden. This corresponds to three lethal attacks among a total of three terrorist attacks recorded in GTD. Among them, the notorious bombing attack of Al Qaeda on the US Navy destroyer USS Cole on October 12, 2010, which killed 17 sailors and wounded more than 40, according to Navy Times (2010). The areas covered by hotspots increased in 2011. Indeed, GTD records 13 lethal attacks among 21 attacks in 2011 in Aden. All attacks have been perpetrated by Al Qaeda or by unknown perpetrators (GTD, 2014). The hotspots covering the city of Aden slightly reduces their size in 2012. Although 35 terrorist attacks occurred in 2012, 11 attacks were lethal, which exhibits a lower proportion of lethal attacks (31%) compared to 2011 (62%) (GTD, 2014). The hotspots keep reducing their size in 2013 (Figure 6.1).

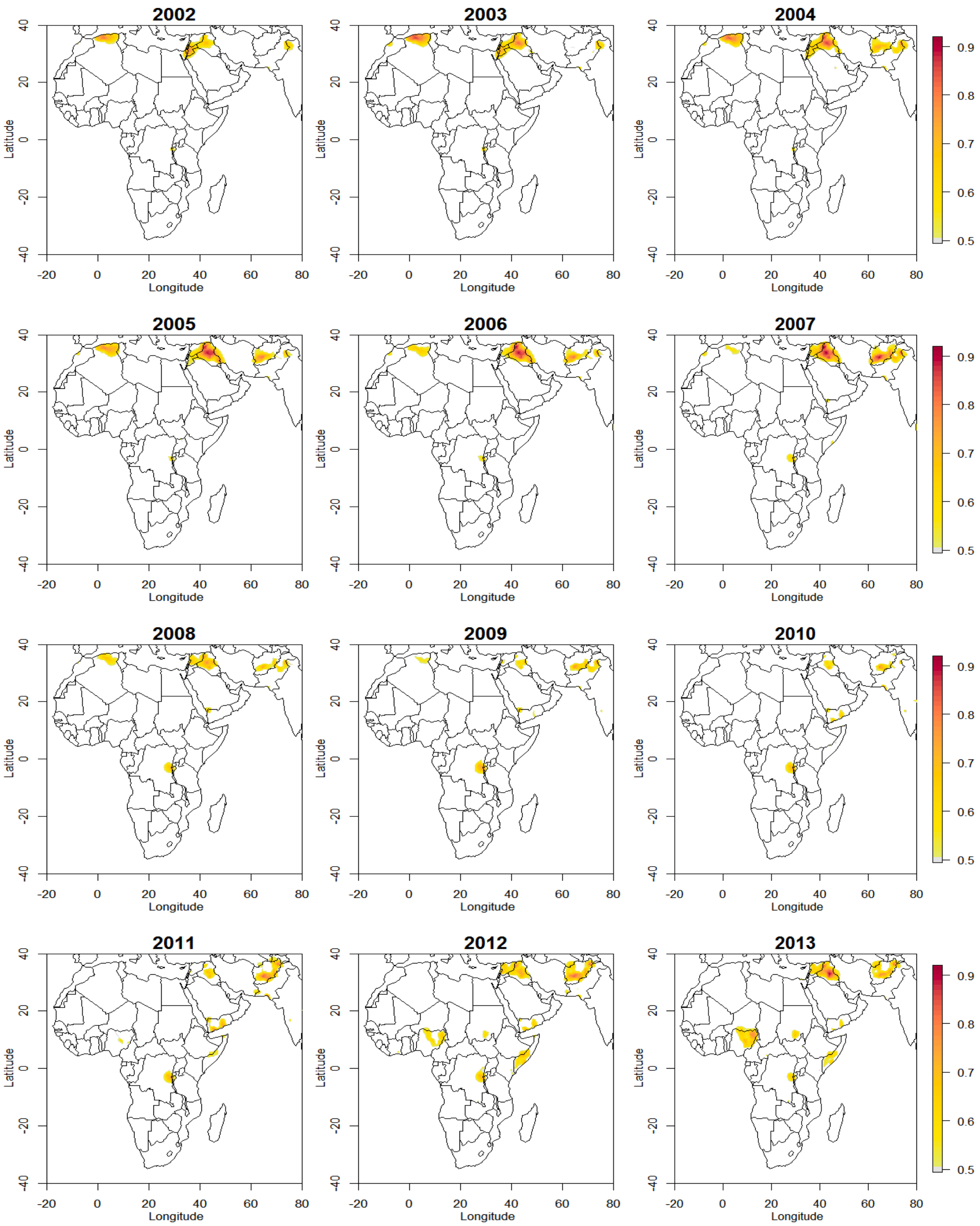


Fig. 6.1 The evolution of hotspots of the probability of lethal terrorist attacks is illustrated for areas that encountered intense terrorism activity from 2002 to 2013, such as the Democratic Republic of Congo, Nigeria, Iraq, Yemen, and Afghanistan. Within a $1,440 \times 720$ regular grid, hotspots are identified if the lower bound of the 95% CI of the posterior probability of lethal attack is lower or equal to 0.5. The highlighted values lie between 0.5 (*bright colour*) and 1 (*dark colour*).

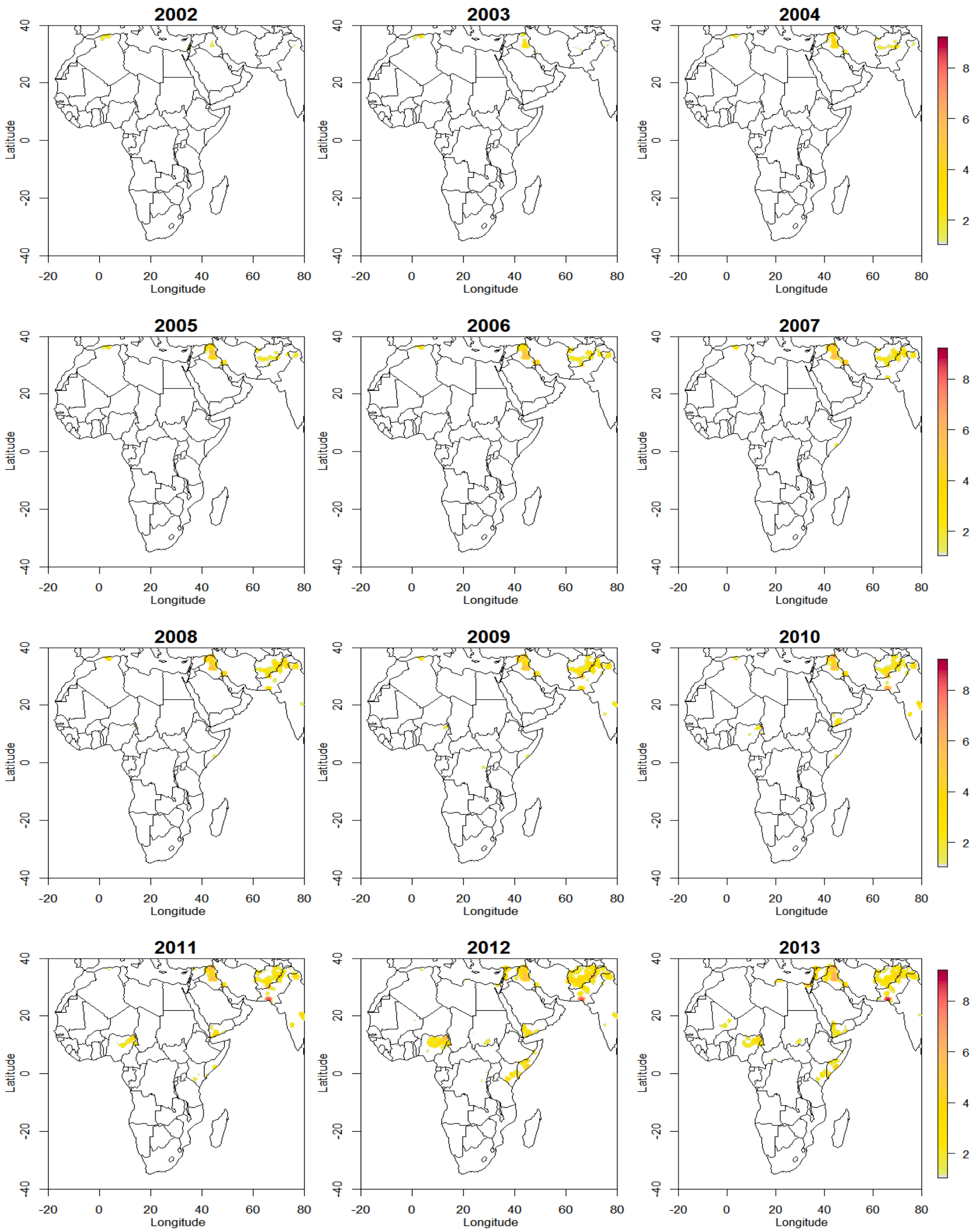


Fig. 6.2 The evolution of hotspots of the number of expected lethal terrorist attacks is illustrated for areas that encountered intense terrorism activity from 2002 to 2013, such as, Nigeria, Iraq, Yemen, and Afghanistan. Within a $1,440 \times 720$ regular grid, hotspots are identified if the lower bound of the 95% CI of the posterior expected number of lethal attack is lower or equal to 5. A log scale is used and the values lie between $\log(5) \approx 1.6$ (*bright colour*) and $\log(9,897) \approx 9.2$ (*dark colour*).

6.1.3 Escalation of Lethal Terrorism

Hotspots can be further distinguished by their potential for escalation, which would require further attention from a counterterrorism point of view. A hotspot, which occurs during a limited time period only and then disappears, is a *temporary* hotspot. In contrast, a *growing* hotspot is characterised by a significant and continuous increase in terrorist activity over time, which corresponds to an *escalation* process. Inversely, a significant and continuous decrease in terrorist activity over time refers to a *de-escalation* process (Anselin et al., 2000; Eck et al., 2005; Zammit-Mangion et al., 2012). The identification of de-escalation processes is also valuable, since it indicates that terrorist groups have considerably reduced their activity. Therefore, this information might be used to assess the efficiency of counterterrorism measures.

Drawing from the definition of hotspot (Section 6.1.1), escalation is measured as change over time in the point estimate of the posterior mean probability of lethal attacks based on its values measured at the bounds of its 95% CI in an area within a hotspot.

Definition 6.2. (**Escalation of lethality**) *Let $\pi_{0.025}(a_{\mathbf{s}_i,t+1})$ be the lower bound of the 95% credible intervals (CI) of the posterior mean probability of the attack(s) to be lethal estimated at time $t+1$ and $\pi_{0.975}(a_{\mathbf{s}_i,t})$ the higher bound of the 95% CI posterior mean probability of the attack(s) to be lethal at time t in a sub-area $a_{\mathbf{s}_i,t}$ of a hotspot $\mathscr{H}_{A_{\mathbf{s},t}}$, with $a_{\mathbf{s}_i,t} \subseteq A_{\mathbf{s},t} \subseteq \mathscr{D}$, with $\mathscr{D}$ the space-time domain, and the observed data* $\mathbf{y}$. *Escalation $\mathscr{E}_{a_{\mathbf{s}_i,t}}$ occurs from time t to $t+1$ in sub-area $a_{\mathbf{s}_i,t}$ of hotspot $\mathscr{H}_{A_{\mathbf{s},t}}$ if it satisfies*:

$$\mathscr{E}_{a_{\mathbf{s}_i,t}} = \pi_{0.025}(a_{\mathbf{s}_i,t+1}) - \gamma\pi_{0.975}(a_{\mathbf{s}_i,t}) > 0|\mathbf{y}, \quad a_{\mathbf{s}_i,t} \subseteq A_{\mathbf{s},t} \subseteq \mathscr{D}, \quad 0 < \gamma \leq 1. \tag{6.2}$$

Note that γ is added as *tolerance threshold* that defines the size of the overlap between the estimation of $\pi(\cdot)$ at time t and $t+1$[2]. Symmetrically, de-escalation occurred if $\mathscr{E}_{a_{\mathbf{s}_i,t}} < 0$. For example, assume an escalation of lethal terrorism in the city of Bonifacio, which is located within a hotspot that covers South Corsica, France from 2002 to 2003. Using a tolerance $\gamma = 0.75$, an escalation process occurs if:

$$\pi_{0.025}(\text{Bonifacio}, 2003) - 0.75 \times \pi_{0.975}(\text{Bonifacio}, 2002) > 0. \tag{6.3}$$

[2] If one uses $\gamma = 1$ (maximum), this signifies that no overlap is permitted, therefore the probability of identifying a "true" escalation process is high, however the probability of not identifying a "true" escalation is high as well. In order to slightly reduce the trade-off, I use $\gamma = 0.75$.

Furthermore, as for the probability of lethal attacks, I identify hotspots of high number of lethal attacks by replacing the posterior probability $\pi(\cdot)$ (Equation 6.1) with the posterior expected number of lethal events $\mu(\cdot)$. Similarly, I suggest using $\gamma = 0.75$. The value can be adapted according to the desired threshold that is required in the study. For the sake of clarity, the approach used to detect escalation and de-escalation processes is illustrated in Figure 6.3. It also highlights diffusion processes, which are described in more detail in Section 6.2.1.

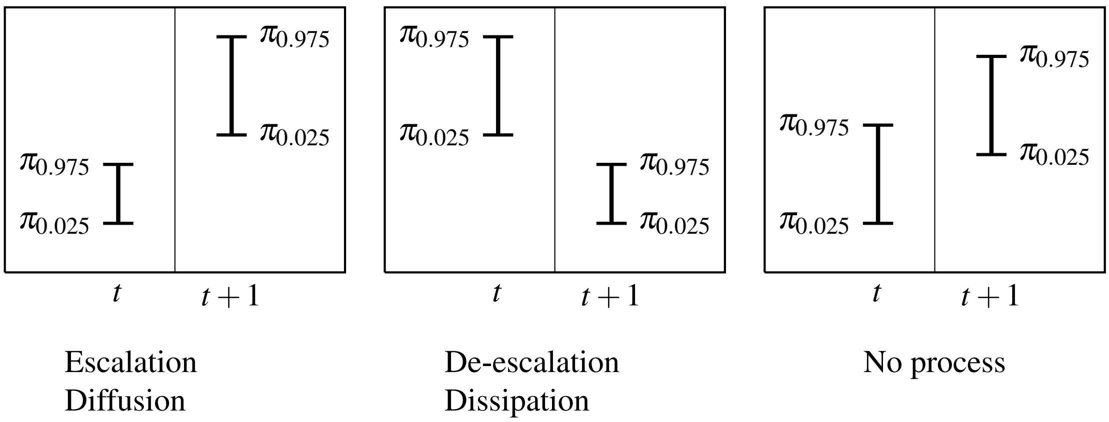


Fig. 6.3 Illustration of three possible dynamic processes. Within a $1,440 \times 720$ regular grid, the lower ($\pi_{0.025}$) and upper ($\pi_{0.975}$) bound of the 95% CI (*vertical segments*) of the posterior probability distribution is estimated in two successive time periods t and $t+1$. *Left*: $\pi_{0.025}(t+1) - \pi_{0.975}(t) > 0$ within a hotspot (*escalation*); within its neighbourhood (*diffusion*). *Centre*: $\pi_{0.025}(t) - \pi_{0.975}(t+1) > 0$ within a hotspot (*de-escalation*); within its neighbourhood (*dissipation*). *Right*: $\pi_{0.025}(t+1) < \pi_{0.975}(t) < \pi_{0.975}(t+1)$ (no process). Here, the tolerance threshold $\gamma = 1$ (Equations 6.2 and 6.4).

For the Bernoulli model, escalation processes (Figure 6.3, *left*) are identified when there is a positive difference between the lower bound of the 95% of the posterior probability distribution at time $t+1$ and the higher bound of the 95% CI of the posterior probability distribution at time t. In contrast, de-escalation processes (Figure 6.3, *centre*) correspond to a positive difference between the lower bound of the 95% of the posterior probability distribution at time t and the higher bound of the 95% CI of the posterior probability at time $t+1$. No specific dynamic process is observed if the upper bound of the posterior probability distribution at time t lies within the 95% CI of the posterior probability distributions at time $t+1$ (Figure 6.3, *right*).

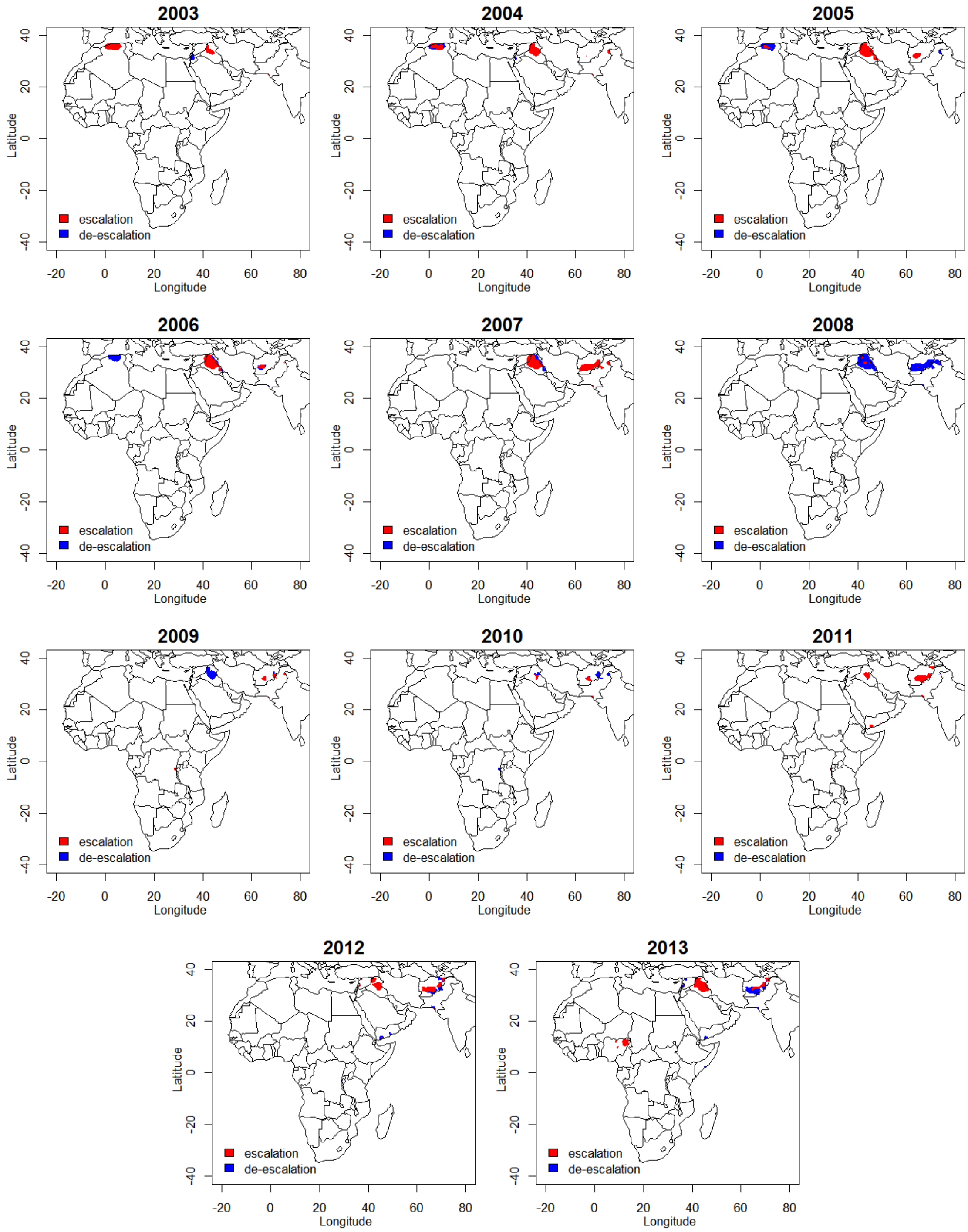


Fig. 6.4 The escalation (*red*) and de-escalation (*blue*) of the lethality of terrorism within hotspots are illustrated for areas that encountered considerable changes in terrorism activity from 2002 to 2013, such as: Nigeria, Iraq, Yemen, and Afghanistan. For example, escalation in year '2003' corresponds to escalation that occurred between year '2002' and '2003'. Escalation is defined according to Equation 6.2 with $\gamma = 0.75$.

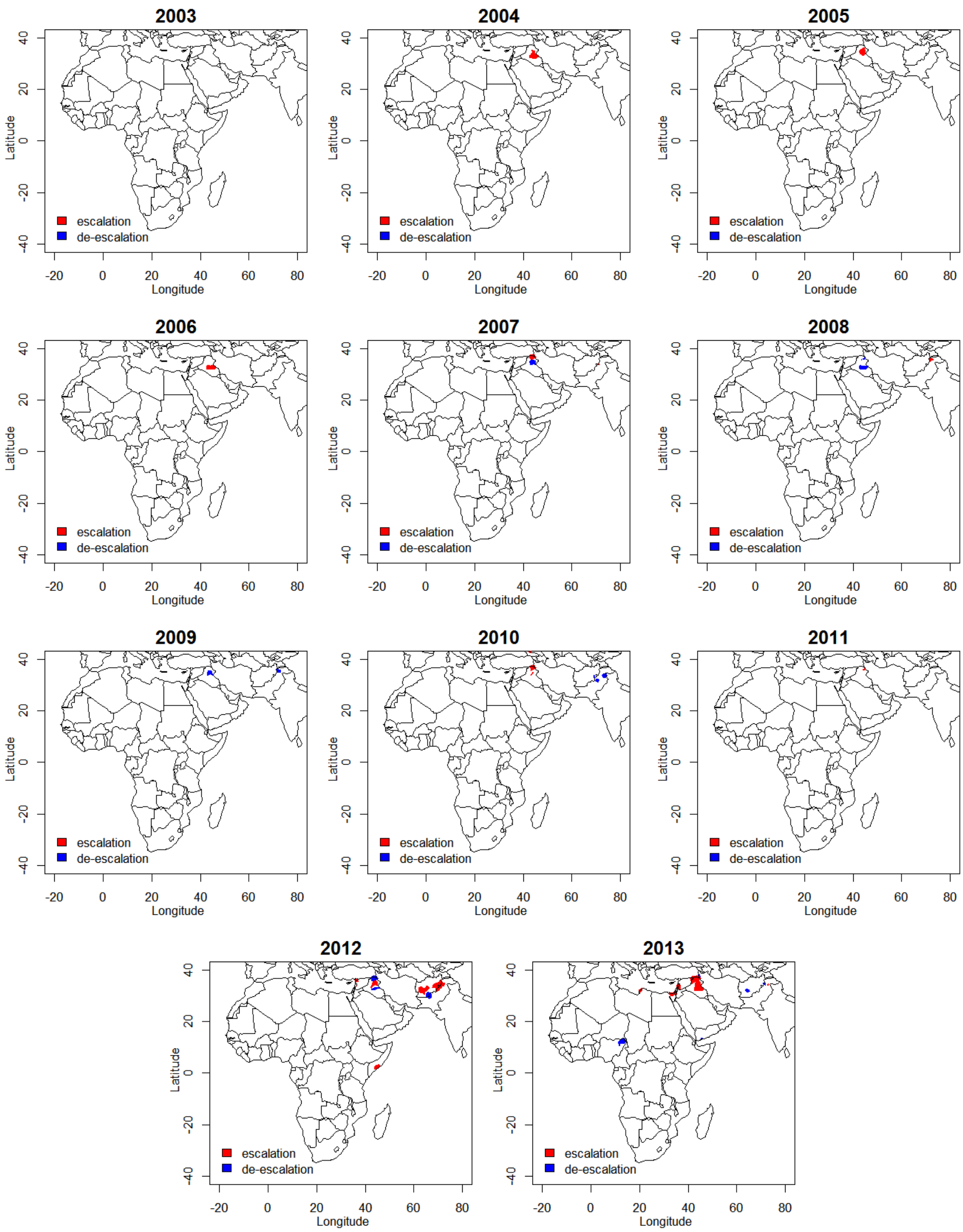


Fig. 6.5 Within a $1,440 \times 720$ regular grid, the escalation (*red*) and de-escalation (*blue*) of the number of lethal terrorist attacks within hotspots are illustrated for areas that encountered considerable changes in terrorism activity from 2002 to 2013, such as Iraq and Afghanistan. For example, escalation in year '2003' corresponds to escalation that occurred between year '2002' and '2003'. Escalation is defined according to Equation 6.2 with $\gamma = 0.75$.

Figures 6.4 and 6.5 illustrate the dynamics of escalation and de-escalation processes that occurred in Africa and in the Middle East from 2003 to 2013 in hotspots of probability of lethal terrorism, and respectively, in hotspots of the number of lethal terrorist attacks. For example, the de-escalation processes in the probability of lethal terrorism (Figure 6.4) that are labelled in year '2005' observed in Northern Algeria (latitude $\approx 36°$, longitude $\approx 3°$), correspond to escalation processes that started in 2004 and ended in 2005. Note that these de-escalation processes observed in Northern Algeria continued until 2006. The escalation (2002-2007) and de-escalation (2008-2009) processes observed in Iraq (Figure 6.4) mirror the changes in hotspots that occurred in the same period and in the same areas (Figure 6.1). From 2005 until 2007, an escalation process occurred in several locations in Afghanistan (latitude $\approx 35°$, longitude $\approx 66°$). From 2008, other areas encountered escalation processes within various hotspots, which reappeared between 2011 and 2013, consistently with the dynamics previously observed in the hotspot maps (Figure 6.1).

6.1.4 Assessing the Performance in Detecting Hotspot and Escalation

The performance of the approach suggested to identifying hotspot (Equation 6.1) and escalation (Equation 6.2) processes is assessed (R code in Appendix D.2). For each hotspot $\mathscr{H}_{A_{\mathbf{s},t}}$ identified in space-time area $A_{\mathbf{s},t}$, with time $t \in T$, $T = \{2002, \ldots, 2013\}$, the observed *proportion* $P_{\mathscr{H}_{A_{\mathbf{s},t}}}$ and *number* $N_{\mathscr{H}_{A_{\mathbf{s},t}}}$ of lethal attacks are computed. The performance in detecting hotspots of lethality is provided by the proportion of hotspots that are identified and exhibit $N_{\mathscr{H}_{A_{\mathbf{s},t}}} \geq 0.5$, consistently with the threshold $\upsilon = 0.5$. Analogously, the performance in detecting hotspots of the number of lethal attacks is provided by the proportion of hotspots that are identified and exhibit $P_{\mathscr{H}_{A_{\mathbf{s},t}}} \geq 5$, consistently with the threshold $\upsilon = 5$.

First, the performance in detecting hotspots of probability and number of lethal attacks is assessed (R code in Appendix D.3). Out of 166 identified hotspots of probability of lethal attacks, 127 (77%) show an observed proportion of lethal attacks that is greater or equal to 0.5. Moreover, out of 189 identified hotspots of number of lethal attacks, 123 (65%) exhibit an observed number of lethal attacks higher or equal to 5. On average, 70% ($\frac{127+123}{166+189}$) of the hotspots identified by the approach correspond to "true" hotspots. The approach used to identify hotspots of lethal terrorism can be considered as satisfactory[3].

[3] Note that the assessment shows that the detection of hotspot is highly *sensitive* (i.e. it has good ability to detect "true positive"). However, it does not assess the *specificity* of the approach (i.e. its aptitude in detecting "true negative").

Secondly, the performance in detecting escalation and de-escalation within hotspots of probability and number of lethal attacks, respectively is assessed. For each grid-cell $a_{\mathbf{s}_i,t}$ identified at time $t \in T$, $T = \{2002, \ldots, 2013\}$ within a hotspot $\mathscr{H}_{A_{\mathbf{s},t}}$, a ratio of the observed proportion P of lethal attacks ($R_{P_{a_{\mathbf{s}_i,t}}} = P_{a_{\mathbf{s}_i,t+1}} / P_{a_{\mathbf{s}_i,t}}$), and respectively, the observed number N of lethal attacks ($R_{N_{a_{\mathbf{s}_i,t}}} = N_{a_{\mathbf{s}_i,t+1}} / N_{a_{\mathbf{s}_i,t}}$), is computed using values gathered at two successive time periods.

Using a 10% threshold, escalation areas with $R_{P_{a_{\mathbf{s}_i,t}}} \geq 1.1$ and $R_{N_{a_{\mathbf{s}_i,t}}} \geq 1.1$ are considered as regions that encountered "true" escalation processes in the probability and number of lethal attacks, respectively. The performance of the identification of escalation processes is the proportion of areas that exhibit "true" escalation processes and that are identified by the approach among the total number of areas with escalation processes identified by the approach. Symmetrically, de-escalation areas with $R_{P_{a_{\mathbf{s}_i,t}}} \leq 0.9$ and $R_{N_{a_{\mathbf{s}_i,t}}} \leq 0.9$ are considered as regions that encountered "true" de-escalation processes in the probability and number of lethal attacks, respectively. Likewise, the performance of the identification of de-escalation processes is the proportion of areas that exhibit "true" de-escalation processes and that are identified by the approach among the total number of areas with de-escalation processes.

Out of 43 identified escalation processes of the probability of lethal attacks, 14 (33%) exhibited a significant observed escalation process. 19 (48%) among 40 de-escalation processes exhibited a significant de-escalation process. Moreover, out of 19 escalation processes of the number of lethal attacks identified by the model, 16 (84%) showed a significant escalation process. Furthermore, out of 13 de-escalation processes, 7 (54%) exhibited a significant observed de-escalation process. On average, the approach exhibits a moderate success in correctly detecting 49% ($\frac{14+19+16+7}{43+40+19+13}$) of the escalation/de-escalation processes of lethal terrorism.

The detection of escalation/de-escalation is challenging since it estimates changes in the activity of terrorism (lethality or number of lethal attacks) in rather small areas — individual grid-cells of approximately $0.25° \times 0.25°$ within hotspots —, while hotspots consist of an aggregation of grid-cells that may represent large areas, such as the persistent hotspots of lethality covering a large part of Iraq from 2003 to 2008 (Figure 6.1). Since one may reasonably expect fewer observations within each grid-cell than within hotspots, the identification of escalation processes is therefore more arduous than the identification of hotspots (see above). As a counterterrorism tool, the identification of hotspots appears more reliable than the detection of escalation processes at high-spatial resolution, as suggested in this present study.

6.2 Diffusion Processes

6.2.1 Concept and Definition

Changes in the activity of terrorism occur in time and also in space: terrorism may spread from hotspots towards neighbourhood regions. This particular type of spread induced by spatial proximity refers to the concept of *contagious diffusion* (Section 2.2.3). For the sake of concision, I will omit the adjective *contagious* from now on, instead using the word *diffusion* in reference of *contagious diffusion*.

Definition 6.3. **(Diffusion of lethality)** *Let $\pi_{0.975}(\mathscr{H}_{\mathrm{A}_{\mathrm{s},t}})$ be the upper bound of the 95% credible intervals (CI) of the posterior mean probability of the attack(s) to be lethal estimated at time t in a hotspot $\mathscr{H}_{\mathrm{A}_{\mathrm{s},t}}$, identified at $t \in T$, $T = \{2002, \ldots, 2013\}$ and $\mathscr{H}_{\mathrm{A}_{\mathrm{s},t}} \subseteq \mathscr{D}$, with $\mathscr{D}$ the space-time domain. Let $\pi_{0.025}(\mathscr{N}_{\mathrm{A}_{\mathrm{s},t+1}})$ be the lower bound of the 95% credible intervals (CI) of the posterior mean probability of the attack(s) to be lethal estimated at time $t+1$ in the neighbourhood $\mathscr{N}_{\mathrm{A}_{\mathrm{s},t}}$ of $\mathscr{H}_{\mathrm{A}_{\mathrm{s},t}}$. Diffusion $\mathscr{D}_{\mathscr{N}_{\mathrm{A}_{\mathrm{s},t}}}$ processes occur from time t to $t+1$ in the neighbourhood $\mathscr{N}_{\mathrm{A}_{\mathrm{s},t}}$ of hotspot $\mathscr{H}_{\mathrm{A}_{\mathrm{s},t}}$ if it satisfies*:

$$\mathscr{D}_{\mathscr{N}_{\mathrm{A}_{\mathrm{s},t}}} = \pi_{0.025}(\mathscr{N}_{\mathrm{A}_{\mathrm{s},t+1}}) - \gamma\pi_{0.975}(\mathscr{N}_{\mathrm{A}_{\mathrm{s},t}}) > 0, \quad \mathscr{N}_{\mathrm{A}_{\mathrm{s},t}} \subseteq \mathscr{D}, \quad 0 < \gamma \leq 1. \tag{6.4}$$

The concepts of diffusion/dissipation are analogous to escalation/de-escalation, respectively. As illustrated in Figure 6.3, in an escalation/de-escalation process, the variation is observed from time t to time $t+1$ within a hotspot, while in the diffusion/dissipation process, it is observed in the neighbourhood of a hotspot. One may apply the concept of diffusion with regard to the posterior probability or the posterior expected number of lethal terrorist attacks by using $\pi(\cdot)$ and $\mu(\cdot)$, respectively. Note that similar to escalation processes, a threshold γ is added to define the tolerance of the identification of diffusion processes. Furthermore, a decrease in the posterior probability or in the posterior expected number of lethal attacks that would occur in the neighbourhood of hotspots refers to a *dissipation* process. More formally, a dissipation process occurs if $\mathscr{D}_{\mathscr{N}_{\mathrm{A}_{\mathrm{s},t}}} < 0$ (Equation 6.4).

For example, consider a hotspot of probability of lethal attacks observed in 2002 in the city of Baghdad and suppose that its neighbourhood includes the city of Abu Grahib. Using the aforementioned definition and a tolerance $\gamma = 0.75$, a diffusion process occurs between 2002 and 2003 from Baghdad to Abu Grahib if:

$$\pi_{0.025}(\text{Abu Grahib}, 2003) - 0.75 \times \pi_{0.975}(\text{Abu Grahib}, 2002) > 0. \qquad (6.5)$$

In this work, a distance-based approach is used to define neighbourhood as 1°-width (approximately 111 km at the equator) buffer areas surrounding each hotspot identified across the world from 2002 to 2013 (R code in Appendix D.1)[4]. Furthermore, the posterior mean probability of the attacks to be lethal — using the upper bound, and respectively, the lower bound of the 95% CI of the posterior probability of lethal attacks — is averaged within the hotspots and within its corresponding neighbouring areas for each year. The same procedure is carried out with regard to the posterior expected number of lethal attacks.

6.2.2 Spatial Dynamics of Diffusion

Some areas e.g. in Afghanistan or Pakistan appear particularly susceptible to diffusion processes at specific time periods. Furthermore, processes of diffusion and dissipation of the probability of lethal attacks may occur alternately, as exemplified in Afghanistan (Figure 6.6): diffusion: 2004 and 2007; dissipation: 2005, 2006 and 2008. This alternating pattern suggests an intense struggle between roughly equal antagonistic forces: (i) highly lethal terrorist groups (mainly Taliban); (ii) counterterrorism forces (led by the US-coalition), where the latter appears to reduce the lethality of the former force only on a temporary basis.

Furthermore, a relatively high number of areas in the neighbourhood of hotspots of probability or number of lethal attacks do not exhibit a diffusion or a dissipation process (*grey* areas in Figures 6.6 and 6.7). This suggests that some hotspots do not spread further, which means that terrorism has been contained in a restricted area. This could be explained by several factors: efficient counterterrorism measures undertaken to limit potential spread and/or the desire of the terrorist group(s) involved to focus their attacks within particular areas and/or the incapacity of the terrorist group(s) involved to spread farther due to a lack of material or human resources. Note that in addition to diffusion processes, other mechanisms of spread may occur beyond the close neighbourhood. As discussed in Section 2.2.3,

[4] Recall that the definition of neighbourhood is subject to debate and could be constructed in various ways, according to its interpretation (see three examples of non distance-based neighbourhood in lattice data: Figures 3.4, 3.5 and 3.6).

relocation processes cannot be ruled out, but they are not considered in this study, since their identification remains challenging.

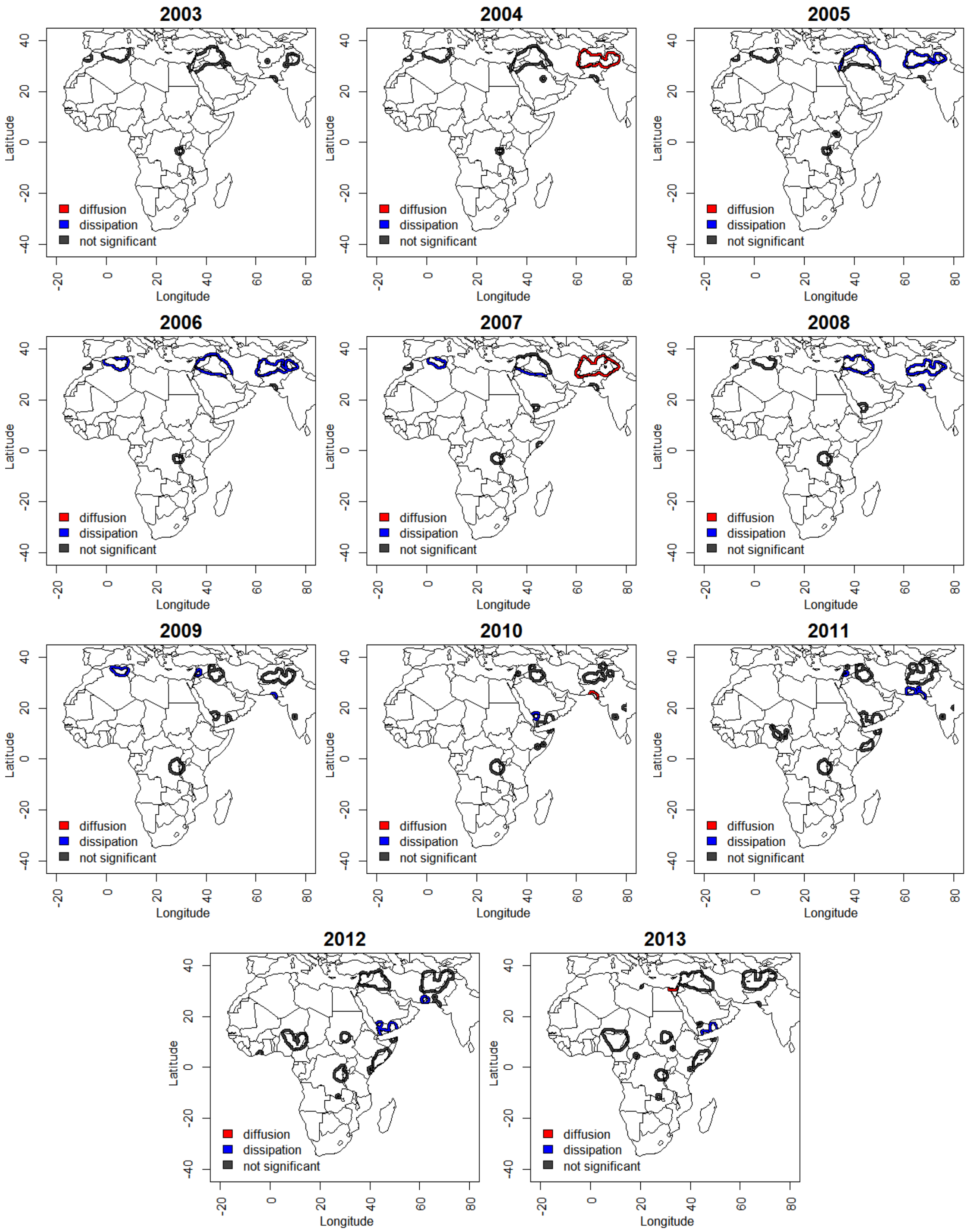


Fig. 6.6 Within a $1,440 \times 720$ regular grid, diffusion (*red*) and dissipation (*blue*) of the lethality of terrorism in the neighbourhood (1° buffer areas) of hotspots of the probability of lethal attacks is illustrated in countries intensely targeted by terrorism from 2002 to 2013. For example, diffusion in year '2003' corresponds to diffusion that occurred between year '2002' and '2003'. Diffusion is defined according to Equation 6.4 with $\gamma = 0.75$. The absence of a diffusion or dissipation process is labelled "not significant" (*grey*).

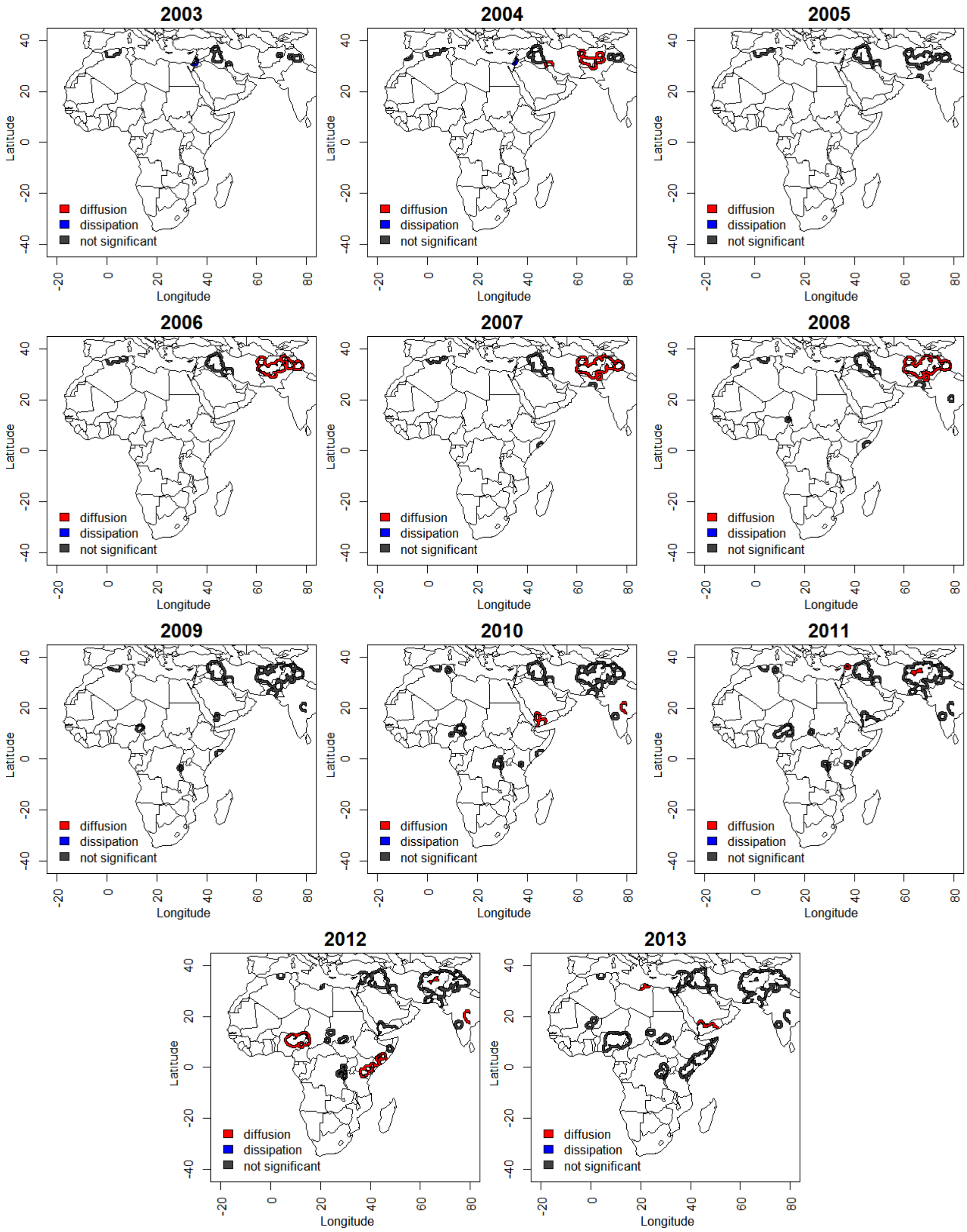


Fig. 6.7 Within a $1,440 \times 720$ regular grid, diffusion (*red*) and dissipation (*blue*) of the number of lethal terrorist events in the neighbourhood (1° buffer areas) of hotspots of the number of lethal attacks is illustrated in countries intensely targeted by terrorism from 2002 to 2013. For example, diffusion in year '2003' corresponds to diffusion that occurred between year '2002' and '2003'. Diffusion is defined according to Equation 6.4 with $\gamma = 0.75$. The absence of a diffusion or dissipation process is labelled "not significant" (*grey*).

6.2.3 Assessing the Performance in Detecting Diffusion

The performance of the approach used in identifying diffusion and dissipation in the probability and number of lethal attacks is assessed (R code in Appendix D.4). For each neighbourhood $\mathscr{N}_{A_{s,t}}$ of hotspots $\mathscr{H}_{A_{s,t}}$ and at time $t \in T$, with $T = \{2002, \dots, 2013\}$, the observed *proportion* $P_{\mathscr{N}_{A_{s,t}}}$ and *number* $N_{\mathscr{N}_{A_{s,t}}}$ of lethal attacks are computed. Hence, a ratio of the observed proportion of lethal attacks ($R_{P_{\mathscr{N}_{A_{s,t}}}} = P_{\mathscr{N}_{A_{s,t+1}}} / P_{\mathscr{N}_{A_{s,t}}}$), and respectively, the observed number of lethal attacks ($R_{N_{\mathscr{N}_{A_{s,t}}}} = N_{\mathscr{N}_{A_{s,t+1}}} / N_{\mathscr{N}_{A_{s,t}}}$), is computed for each successive time period within T.

Using a 10% threshold, neighbourhood areas with $R_{P_{\mathscr{N}_{A_{s,t}}}} \geq 1.1$ and $R_{N_{\mathscr{N}_{A_{s,t}}}} \geq 1.1$ are considered as regions that encountered "true" diffusion processes in the probability and number of lethal attacks, respectively. The performance of the identification of diffusion processes is the proportion of areas that exhibit "true" diffusion processes and that are identified by the approach among the total number of neighbourhood areas with diffusion processes. Symmetrically, neighbourhood areas with $R_{P_{\mathscr{N}_{A_{s,t}}}} \leq 0.9$ and $R_{N_{\mathscr{N}_{A_{s,t}}}} \leq 0.9$ are considered as regions that encountered "true" dissipation process in the probability and number of lethal attacks, respectively. Likewise, the performance of the identification of dissipation processes is the proportion of areas that exhibit "true" dissipation processes and that are identified by the approach among the total number of neighbourhood areas with dissipation processes.

Furthermore, the performance in detecting diffusion and dissipation processes in the lethality of terrorism and the number of lethal attacks is assessed. Out of 19 identified diffusion processes, 9 (47%) show an observed diffusion of the probability of lethal attacks. Out of 35 detected dissipation processes, 14 (40%) contain an observed dissipation of the probability of lethal attacks. Out of 84 detected diffusion processes, 39 (46%) exhibit an observed diffusion of the number of lethal attacks. Moreover, out of 48 identified dissipation processes, 7 (15%) show an observed dissipation of the number of lethal attacks.

On average, 37% ($\frac{9+14+39+7}{19+35+84+48}$) of the diffusion and dissipation processes have been correctly detected by the models. This relative low rate is partially due to the fact that this study considers the average value of diffusion and dissipation within the whole neighbourhood regions, which could represent large areas when hotspots cover vast territories (e.g. diffusion areas in Iraq from 2004 to 2007 in Figure 6.6). Consequently, some "true" diffusion processes might not be detected if an increase in the probability of lethal attacks in one or more regions in the neighbourhood were compensated by a decrease in the probability of lethal attacks in one or more other areas in the neighbourhood.

6.3 Theory Assessment

6.3.1 Hotspot Theories: An Empirical Assessment

Recall from Section 1.1 that one main objective of this thesis is to assess theoretical work that aims to explain the causes and the patterns of hotspots of lethal terrorism. More specifically, this section will assess theories described in Section 2.2.2, which include: the *repeat victimization theory*, *flag theory*, *social disorganisation theory*, *boost theory*, *terrorism geography theory* and *crime generator theory*. Since the predictions of theories may be observed from different spatial scales, I distinguish two approaches:

1. *global*: evidence that brings support or not to the assessed theory is based on a comparison between values obtained locally and values obtained from the entire domain (e.g. average luminosity in hotspots versus average luminosity in all locations of terrorist events in the world);

2. *local*: evidence that brings support or not to the assessed theory is based on a comparison between values obtained locally exclusively (e.g. average luminosity in hotspots versus average luminosity in the close neighbourhood of hotspots).

In other words, the *global* approach answers questions in *absolute* terms, while the *local* approach answers questions in *relative* terms. Whenever applicable I will assess theories from both points of view. Otherwise, I will use either the global or the local approach, depending on the suitability of each approach to assess a given theory. Drawing from Section 2.2.2, I briefly recall the main concept of each theory and describe the results of the assessment. Particular care is required in order to assess theories within the framework suggested by this research work.

First, the repeat victimization, crime generator, terrorism geography, and flag theories jointly posit that the characteristics of the location matter; the spatial location itself may drive the type and the risk of crime, including terrorism (Brantingham and Brantingham, 1995; Eck et al., 2005). Moreover, previous criminal activity increases the risk of crime in a given location, as predicted by the boost theory. The results of the analysis carried out in

Section 4.3 brought support to these theories. Terrorism does indeed not occur haphazardly in space; rather it tends to form clusters in specific locations. More specifically, I showed through a point pattern analysis (Section 4.3.4) that hotspots of terrorism are mainly generated *locally*, meaning that terrorist events tend to be spatially aggregated within very close distances, bringing further support to the aforementioned theories.

Secondly, lethal terrorism directly affects human lives and a better knowledge of its spatial dynamics is of major interest. It would be instructive to identify the main underlying driving factors that increase the risk of encountering hotspots of lethal terrorism in specific locations. Drawing from crime studies, the social disorganisation theory posits that more disorder and crime is more likely in locations with high residential instability, low socio-economic conditions and higher ethnic diversity (Shaw and McKay, 1942) (as cited in Steenbeek and Hipp, 2011). Analogous findings in terrorism suggest that social, economic, and demographic characteristics of the location directly influence its propensity to encounter terrorism (Nunn, 2007; Piegorsch et al., 2007).

The results from Chapter 5 bring only weak support to the predictions of social disorganisation theory with regard to ethnicity. From a global point of view, it appears that ethnically heterogeneous locations are more prone to encounter a higher proportion of lethal rather than non-lethal attacks. However, the number of lethal attacks does not appear to be influenced by ethnic diversity. The proportion of lethal attacks is more likely to be higher while their number is more likely to be smaller in less economically developed areas. Nevertheless, state instability appears a more robust factor to predict hotspots of lethal terrorism. Hotspots of lethal terrorism are more frequent in locations within failed states (more detail in Section 6.3.2).

Furthermore, I investigate the role of economic development with regard to the occurrence of hotspots of lethality of terrorism and hotspots of the number of lethal attacks. As discussed in Section 5.1.2, I used satellite night lights (luminosity) (NOAA, 2014) to estimating economic development at local level. Satellite night lights have been used as a proxy for per capita GDP estimation for example (Elvidge et al., 2007; Henderson et al., 2009; Sutton and Costanza, 2002). Hence, I compare: (i) the average values of luminosity taken at the location of GTD events that occurred in the world between 2002 and 2013 with those in hotspots and their neighbourhood (*global* assessment), and (ii) the average values of luminosity in hotspots with those in their neighbourhood (*local* assessment).

From a global point of view, there is strong evidence that both hotspots of probability of lethal attacks (Bernoulli model) and number of lethal attacks (Poisson model) occur mainly in more economically developed areas compared to other regions where terrorist events occurred (2002-2013) (Student's t-test: $p < 0.0005$) (R code in Appendix D.6). These results contradict the social disorganisation theory, which predicts more hotspots of terrorism in less economically developed areas. Rather, it brings support to Piazza (2006), who claims that more economically developed societies are expected to be more targeted by terrorist attacks since they provide more attractive targets.

At a local level, the results are more contrasted. There is no evidence that the average level of economic development differs between hotspots of probability of lethal attacks and their neighbourhood (Mann-Whitney's test: $p > 0.1$). However, the average level of economic development tends to be higher in hotspots of the number of lethal attacks and their neighbourhood (Mann-Whitney's test: $p < 0.007$). Figure 6.11 displays the distribution of the luminosity (proxy for per capita GDP) values averaged in hotspots (with and without diffusion areas), diffusion areas, and in locations of terrorist events that occurred in the world from 2002 to 2013 for the Bernoulli (*left*) and the Poisson (*right*) model.

6.3.2 Assessing Failed State Theory

The diffusion processes observed in terrorism have been under scrutiny from empirical and theoretical work since the 1980s. Mainstream theories include in particular *failed state* and *hierarchical* theories. Mainly derived from empirical studies, scholars have attempted to examine the role of economic development in the diffusion of terrorism as well (Section 2.2.3). However, the literature failed to assess these theories at sub-national level (Section 2.3). As a remedy, this Section suggests an approach to assess these theories on a *local-scale* and discusses its results within the context of the literature.

Failed state theory is assessed based on the analysis of the values of the Fragile States Index (FSI) (The Fund for Peace, 2015) obtained in the origin (hotspot) and in the area of diffusion. The FSI provides an annual estimation of the degree of which countries are considered as "failed" based on twelve social, economic and political indicators[5]. A *global* assessment compares the FSI values within hotspot and area of diffusion with worldwide average values. Furthermore, a *local* assessment of failed state theory compares the FSI

[5] For further details on FSI, see http://fsi.fundforpeace.org/indicators/.

values within hotspot and their neighbourhood (that may or may not encounter diffusion processes).

Recall that failed state theory (also-called *state collapse theory* (Zartman, 1995, pp. 1-13)) predicts that terrorism spreads from weak countries to neighbouring countries essentially because of the incapacity of weak states to control bordering areas. Moreover, the spread is expected to be mainly located in failed states as well, since they should be less capable of preventing terrorism entering their territories (Gros, 1996; Helman and Ratner, 1992). In order to assess the validity of these claims, two questions are addressed from both *local* and *global* points of view: (i) are hotspots more likely in failed states?; (ii) does the spread tend to direct mainly towards locations within failed states?

As a first step, I investigate the relationship between failed states and the diffusion in the probability of lethal attacks. I keep areas of diffusion and dissipation processes which have been identified (Equation 6.4) and confirmed by observed data from 2005 to 2013 (`R` code in Appendix D.5). Terrorism data starts from 2002, however data on failed state starts only from 2005 (The Fund for Peace, 2015), so the investigated time period is reduced accordingly. Out of 146 hotspots of lethality that have been identified from 2005 to 2013, 46 (32%) are surrounded by areas that exhibit an observed diffusion, as defined in Section 6.2.3. Next, the FSI is reported for each hotspot (based on the location of its centroid) and for its neighbourhood areas (based on an average within the area) (Table in Appendix D.1). Figure 6.8 shows the evolution of diffusion (*red*) and dissipation (*blue*) processes in the neighbourhood of hotspots (*yellow*), and the FSI, from 2005 to 2013.

As a second step, I identify a total of 175 hotspots of the number of lethal attacks that occurred during the study period. Among them, 114 (65%) are surrounded by areas that exhibit an observed diffusion or dissipation process, as defined in Section 6.2.3. Next, similar to the first step, the FSI is reported for each hotspot and for its neighbourhood areas (Table in Appendix D.2). Figure 6.9 shows the evolution of diffusion (*red*) and dissipation (*blue*) processes in the neighbourhood of hotspots (*yellow*), and the FSI, from 2005 to 2013.

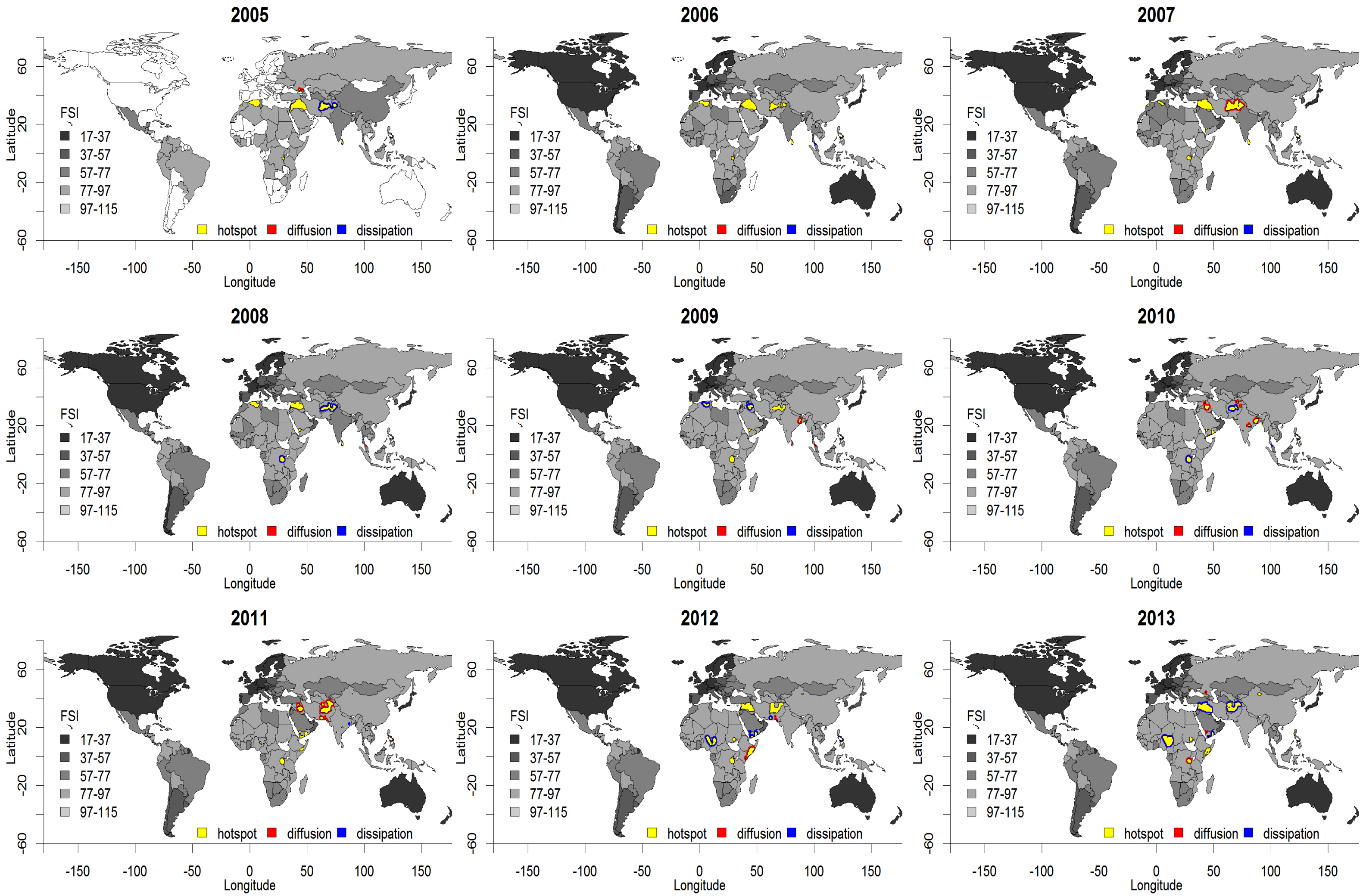


Fig. 6.8 Within a $1,440 \times 720$ regular grid, the observed diffusion (*red*) and dissipation (*blue*) processes that occurred across the world from 2005 to 2013 in the neighbourhood of hotspots (*yellow*) of probability of lethal attacks are highlighted. In addition, the Fragile States Index (FSI) is provided. High values of the FSI (*bright grey*) indicate states with high vulnerability to collapse or in the process of collapsing. Countries with missing FSI value are *blank*.

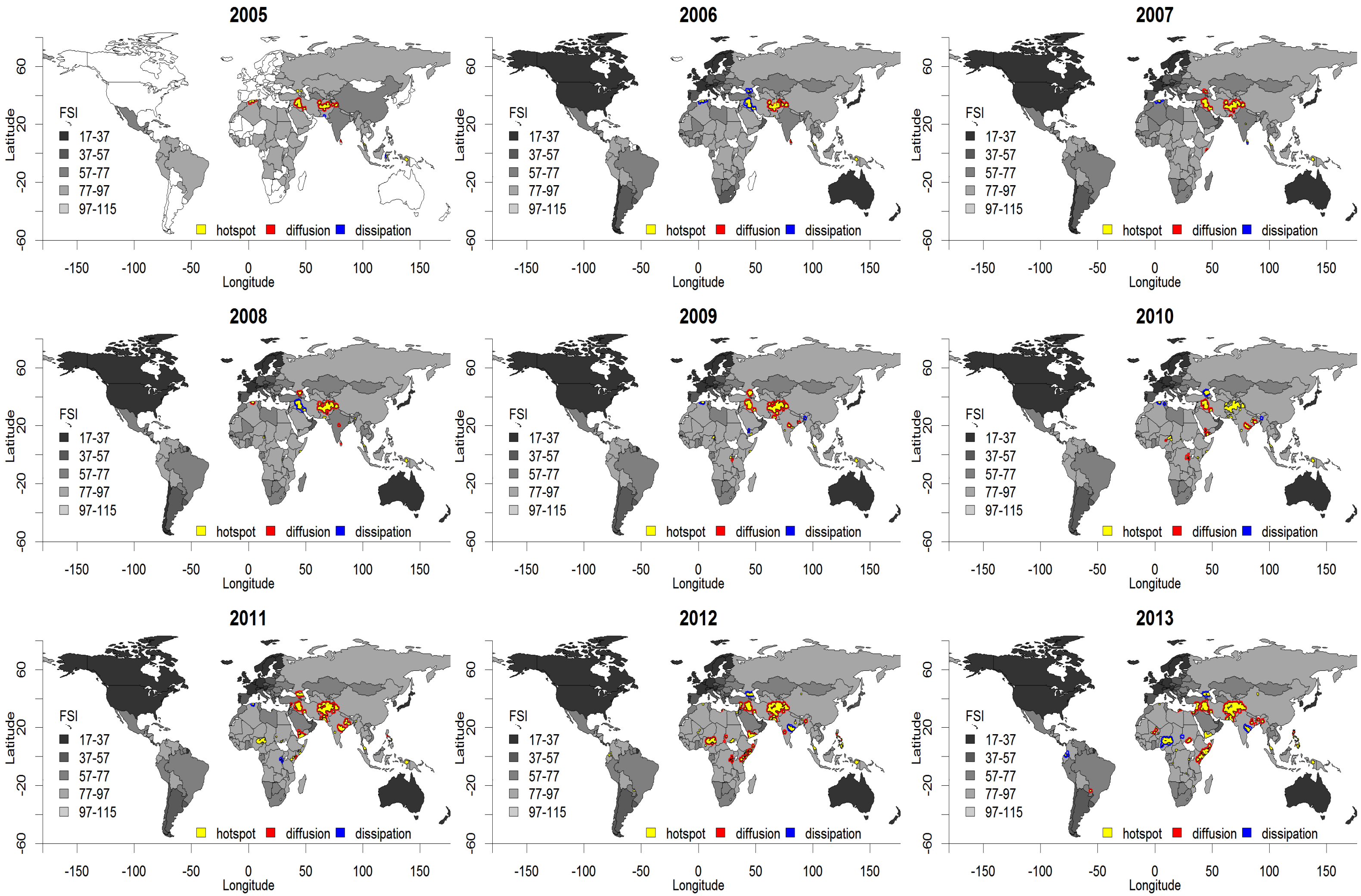


Fig. 6.9 Within a $1,440 \times 720$ regular grid, the observed diffusion (*red*) and dissipation (*blue*) processes that occurred across the world from 2005 to 2013 in the neighbourhood of hotspots (*yellow*) of the number of lethal attacks are highlighted. In addition, the Fragile States Index (FSI) is provided. High values of the FSI (*bright grey*) indicate states with high vulnerability to collapse or in the process of collapsing. Countries with missing FSI value are *blank*.

From a global point of view, the aforementioned questions (i) and (ii) are therefore positively answered; it appears that hotspots are more likely to appear in failed states and the spread tends to direct mainly towards locations within failed states. Note that I did not find evidence that failed state theory applies from a local point of view. Rather, diffusion tends to flow towards locations with lower FSI compared to the origins of the spread (hotspots) in the Poisson model (Mann-Whitney's test: $p < 0.001$). In the Bernoulli model, there is no evidence that the FSI is different between the origins and their areas of diffusion (Mann-Whitney's test: $p > 0.1$). Nevertheless, since hotspots and the diffusion areas of both the probability and number of lethal attacks occur within countries with very high FSI compared to the world's average FSI (Mann-Whitney's test: $p < 0.001$, Figure 6.10, and `R` code in Appendix D.5), the claims of failed state theory hold from a global perspective.

6.3.3 Assessing Hierarchical Theory

As famously known as Galton's problem (Section 2.2.3), the pertinent question arises as to whether the perceived spread is generated by spatial proximity only, from hotspot towards its neighbourhood, or rather, the spread is indirectly triggered by independent precipitating events, such as imitation processes for example. Recall that the hierarchical diffusion theory (Midlarsky et al., 1980) posits that "high status" states imitate terrorism that occurs in "low status" countries (Section 2.2.3).

A spread towards a country with a higher status than the country at the origin of the spread finds might be explained by terrorism geography theory (Nunn, 2007; Piegorsch et al., 2007). Similar to crime (Brantingham and Brantingham, 1995), terrorists aim at targeting symbolic and vulnerable targets (Savitch and Ardashev, 2001). Since attractive targets are usually well protected within more economically developed states, terrorist groups may use "safe havens" located within failed states to launch their attacks towards their prime targets located in economically developed neighbouring areas, which would be more difficult to carry out from within more economically developed areas. Hence, should we therefore expect to observe a tendency of terrorism to spread from locations within low status states towards locations within high status states?

In order to assess the plausibility of the theories, I used FSI as proxy for estimating the status of countries. The higher the status, the lower the FSI. The results from the previous Section suggest that hotspots and diffusion processes are mainly located in locations within low status countries (high FSI) compared to the world's average, which invalidate the hierar-

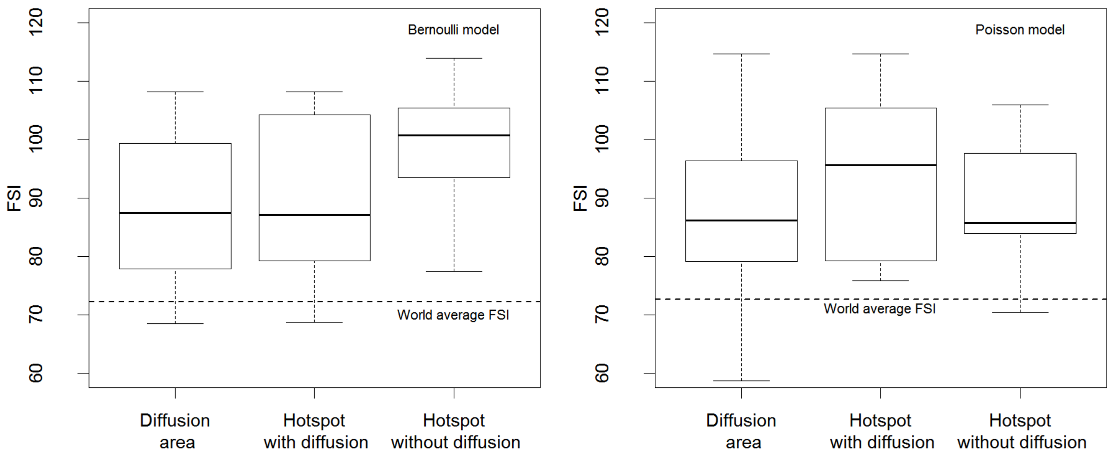


Fig. 6.10 The box plots illustrate the values of the Fragile States Index (FSI) in hotspot and diffusion areas. For each model (*left*: probability of lethal attacks (Bernoulli) and *right*: number of lethal attacks (Poisson)), three box plots show the 25^{th} ($Q1$) and 75^{th} percentiles ($Q3$) (*bottom* and *top* of the box, respectively), the median (*thick line segment* between $Q1$ and $Q3$), *upper whisker* ($\arg\min(max(FSI), Q3 + 1.5 \times (Q3 - Q1))$) and *lower whisker* ($\arg\max(min(FSI), Q1 - 1.5 \times (Q3 - Q1))$) of the FSI for: diffusion areas, hotspots with diffusion in their neighbourhood, and hotspots without diffusion in their neighbourhood. The world average FSI is represented by a *dashed line*.

chical diffusion and terrorism geography theories from a global point of view. From a local point of view however, it appears that diffusion processes related to the number of lethal attacks (Poisson model) tend to spread from locations with low status (high FSI) towards locations with higher status (lower FSI) (Mann-Whitney's test: $p < 0.001$). This latter result might be explained by non-contagious factors, as suggested by the hierarchical theory and terrorism geography, which appear valid with regard to the number of lethal attacks and from a local point of view exclusively. The results are illustrated in Figure 6.10.

The role of economic development in the occurrence and the spread of terrorism has eluded scholars. As discussed in Section 2.1.3, national-level studies showed that the relationship between per capita GDP and terrorism is not significant (Abadie, 2006; Drakos and Gofas, 2006a; Gassebner and Luechinger, 2011; Krueger and Laitin, 2008; Krueger and Maleckova, 2003; Piazza, 2006), or the relationship is at best not linear (Enders and Hoover, 2012).

In order to assess if the spread of terrorism is driven by economic development, two questions arise: (i) does terrorism tend to spread from hotspots towards less economically developed regions compared to the origin of the spread (local assessment)?; (ii) does terrorism tend to spread from hotspots towards less economically developed regions compared to other locations of terrorism in the world (global assessment)? In order to answer them, I compare the average values of luminosity within each hotspot with: (i) the values within diffusion areas of the probability and number of lethal attacks; (ii) with the values within other locations across the world that encountered terrorism during the study period.

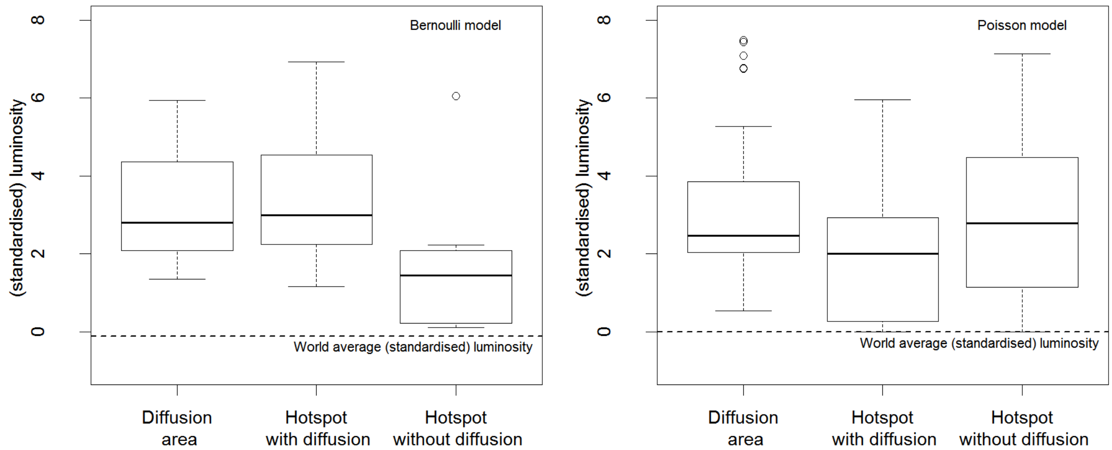


Fig. 6.11 The box plots illustrate the values of (standardised) luminosity in hotspot and diffusion areas. For each model (*left*: probability of lethal attacks (Bernoulli) and *right*: number of lethal attacks (Poisson)), three box plots show the 25$^{\text{th}}$ ($Q1$) and 75$^{\text{th}}$ percentiles ($Q3$) (*bottom* and *top* of the box, respectively), the median (*thick line segment* between $Q1$ and $Q3$), ($\arg\min(max(lum), Q3 + 1.5 \times (Q3 - Q1))$) and *lower whisker* ($\arg\max(min(lum), Q1 - 1.5 \times (Q3 - Q1))$) of the luminosity ($lum$) for: diffusion areas, hotspots with diffusion in their neighbourhood, and hotspots without diffusion in their neighbourhood. The world average (standardised) luminosity taken at the locations of the terrorist events (Bernoulli (*left*): $n = 35,917$; Poisson (*right*): $n = 6,386$) is represented by a *dashed line*. *Circles* indicate outliers.

From a global point of view, hotspots and diffusion processes of lethal terrorism are mainly present within more economically developed regions compared to other locations in the world where terrorism occurred (Student's t-test: $p < 0.0001$) (`R` code in Appendix D.6). From a local point of view, there is no evidence that lethal terrorism (both the lethality and the number of lethal attacks) spreads towards regions with higher or lower level of economic development compared to the origin of the spread (hotspot) (Mann-Whitney's test: $p > 0.1$). These contrasting results between the local and global points of view (illustrated in Figure 6.11) highlight the crucial role of the scale on which theories are assessed.

6.4 Conclusion

This Chapter has provided a rigorous framework used to analyse space-time processes of lethal terrorism. It has suggested definitions of hotspot, escalation, de-escalation, diffusion, and dissipation processes based on the posterior distribution resulting from a Bayesian analysis. Hotspots, escalation, and diffusion processes have been detected mainly in locations in Iraq, Yemen, Afghanistan during the study period, from 2002 to 2013. It appears that spatial proximity (*contagious diffusion*) facilitates diffusion at sub-national level, as observed at national-level by previous research works (Forsberg, 2014; Gao et al., 2013; Heyman and Mickolus, 1980; LaFree et al., 2009; Midlarsky et al., 1980; Neumayer and Plümper, 2010).

Furthermore, alternative theories that explain *non-contagious* diffusion processes have been examined. I showed that from a global point of view, hotspots, escalation, and diffusion processes occur mainly in failed states, which brings support to the failed state theory, complementary to the empirical findings from Piazza (2008) carried out at national level. Contrasting with country-level studies that did not find direct links between economic development (based on per-capita GDP data from the World Bank (2015)) and terrorism (Abadie, 2006; Blomberg et al., 2004; Drakos and Gofas, 2006b; Gassebner and Luechinger, 2011; Krueger and Maleckova, 2003; Piazza, 2006), I showed that globally, areas with a higher level of economic development tend to be more affected by hotspots and diffusion processes, which does not fit with the social disorganisation theory with regard to the role of socio-economic conditions but provides support to Piazza (2006), who predicts a higher intensity of terrorism in more economically developed societies.

Most importantly, this work pointed out the importance of the scale on which theories are assessed. In general, most claims derived from the failed state and social disorganisation theories that are valid from a global point of view do not hold locally. The hierarchical theory and the terrorism geography theory may explain the processes related to the number of lethal attacks from a local point of view. However, the diffusion processes of the probability of lethal attacks failed to be explained by the investigated theories at a local level.

Despite the efforts to built a rigorous framework to assess theories on a local scale, this suggested approach is subject to important shortcomings, which need to be addressed since they might restrict the generalisability of the results. First, the framework is based on the results of the modelling processes described in Chapter 5. As mentioned in Section 5.4, the models are limited by several factors, including: resolution in space (city-level) and time (yearly aggregation), narrow view on the concept of terrorism (based on GTD), omission of relevant drivers of terrorism, assumption on spatial stationarity, non-separability of the space-time model, and subjectivity in the choice of prior and in detection's threshold to name but some.

It is worth remembering that the detection of hotspot and diffusion processes is ineluctably dependent on the scale on which the phenomenon is observed. Furthermore, subjectivity cannot be excluded in the threshold values used to measure the performance of the detection of hotspot and diffusion processes. They need to be carefully chosen according to the tolerance level required that might differ from one study to another. Also, the results form a comparison of values (e.g. luminosity or FSI) between two neighbouring areas are not necessarily consistent with those from a comparison between e.g. hotspots and world's average. For example, I showed that diffusion of lethal terrorism occurs mainly in less economically developed and low status areas compared to other regions in the world. However, I also showed that from a local point of view, diffusion flows from hotspots towards more economically developed neighbouring areas and locations within countries with higher status.

Chapter 7

Conclusion

7.1 Introduction

This thesis has presented applied research work which combines a computationally efficient spatial modelling approach that expresses a Gaussian Random Markov Field (GMRF) as the solution to a Stochastic Partial Differentiation Equation (SPDE) (Lindgren et al., 2011; Rue et al., 2009), analysing the spatial dynamics of non-state lethal terrorism that occurred in the world from 2002 to 2013. This analysis has exclusively focused on terrorism perpetrated by non-state actors (non-state terrorism) — for the sake of conciseness the term "non-state" has been omitted throughout this work — since it has used the *Global Terrorism Database* (GTD) (GTD, 2014), which excludes terrorist events perpetrated by state actors (state terrorism). Furthermore, this study has used various complementary data sources that provided socio-economic and geographic covariates at high spatial resolution (Amante and Eakins, 2009; CIESIN, 2005; Marshall et al., 2014; Nelson, 2008; NOAA, 2014; Weidmann et al., 2010).

Both the underlying processes behind the *probability* and *number* of lethal terrorist attacks have been considered in continuous space and represented by a Gaussian Field (GF) discretised into a GMRF. The combination of the SPDE approach with the Integrated Nested Laplace Approximation (INLA) has successfully modelled the fine-scale processes of lethality of terrorism and the number of lethal attacks in both space and time. Hence, *local* hotspot, escalation, and diffusion processes have been identified and analysed. Furthermore, theories that explain these processes have been assessed on a local-scale within a wide spatial (world) and temporal (2002-2013) framework.

The suggested methodology includes seven main elements, which can be summarised as follows in a chronological order or appearance in the thesis:

1. Data selection: selection of terrorism database and covariates that cover the space-time domain (space: world; time: years 2002-2013) at high spatial resolution;
2. Terrorist events selection: selection of temporally and spatially accurate terrorist events perpetrated by non-state actors;
3. Discretisation: representation of the infinite dimensional SPDE into a discretised field (GMRF) with piecewise linear basis functions over a triangular mesh that covers the domain;
4. Model selection: selection of the most parsimonious Bernoulli (probability of lethal attack) and Poisson (number of lethal events) space-time models;
5. Bayesian inference: estimation of the posterior distribution of the parameters of the Bernoulli and Poisson space-time models;
6. Detection of spatial dynamics: identification and analysis of hotspots and diffusion processes of lethal terrorism;
7. Theory assessment: assessment of theories of hotspot and diffusion on a local scale from both global and local perspectives.

The objectives set out in Section 1.1 have been achieved. This present work has suggested that both the lethality of terrorism and the number of lethal terrorist attacks are spatially and temporally *autocorrelated*. In other terms, lethal terrorism does not appear locally at random (*complete spatial randomness*); rather it tends to form clusters, also called hotspots, which can further expand through diffusion processes. The occurrence of lethal terrorism has been (partially) explained by economic, geographic, and demographic factors.

Furthermore, the Bayesian approach used in this study has accurately quantified the *uncertainty* of the spatial dynamics of lethal terrorism, which represents key information that could be used by policy-makers to prevent and combat terrorism. The rigorous framework implemented in this work allows the assessment of theories on a *local* scale, which represents a considerable improvement with regard to *country*-level assessment, which cannot capture local-scale processes of terrorism.

7.2 Summary

Chapter 2 has introduced key concerns related to the study of terrorism, including the epistemology, causality and rationality of terrorism, which are still debated within the field of terrorism studies. This chapter has stressed the important impact of these debates on the entire modelling process, as well as the interpretation and generalisability of the results of this present study. Furthermore, it has clarified important concepts specific to space-time processes, such as spatial autocorrelation, hotspot and diffusion through a selective review of the literature in terrorism and related areas of research, including crime and conflict. It has identified the literature gap and described the aim of this present study.

Chapter 3 gradually introduced the reader to the concepts of statistical modelling through a brief exploration of techniques starting from the most simple temporal models to more sophisticated spatio-temporal model recently used in the literature. The review has highlighted the limits of each approach used so far and has argued in favour of the application of recent advances in spatial statistics and inferential techniques. Furthermore, it has described the main approaches (point process, geostatistics, and lattice) used to represent spatial data. Hence, it has pointed out the suitability of combining the SPDE approach with accurate and fast inference techniques (INLA) in modelling lethal terrorism through the probability and the number of lethal attacks.

Chapter 4 described the results of an exploratory analysis of data on terrorism based on the main current databases that provide geolocalised data on terrorism worldwide. The characteristics of the data from each provider have been compared and their issues have been highlighted. The analysis of the temporal trends has indicated discrepancies among the database providers, despite the reinforcement effects commonly observed in all databases. Furthermore, the presence of spatial correlation has been identified and analysed globally through Moran's I, and locally, through Getis-Ord $G^*_{\mathbf{s}_i}$. Furthermore, second-order properties of spatial point patterns of terrorism have been explored through the pair correlation function (*pcf*), which has suggested a strong clustering activity on a local scale.

Chapter 5 provided justification for the choice of using GTD as database on terrorist events perpetrated by non-state actors and based on specific criteria, stating the theoretical ground of including specific covariates in the statistical models. The construction of the mesh has been described along with the specification of the Bernoulli and Poisson models, which have been used to model the probability of the lethality of terrorist attacks and the number of lethal attacks, respectively. The results, which apply to non-state terrorism exclusively, have confirmed the contagious nature of lethal terrorism in both space and time. Moreover, they have indicated that the number of lethal attacks tends to be higher in more economically developed areas, close to large cities, and within democratic countries. In contrast, attacks appear more likely to be lethal far away from large cities, at higher altitudes, in less economically developed areas, and in locations with higher ethnic diversity. A sensitivity analysis has indicated that most results are robust to changes in prior and in mesh size.

Chapter 6 focused on the detection of hotspot and diffusion processes. Drawing from the results obtained in Chapter 5, hotspot, escalation and diffusion processes have been defined and their dynamics analysed. The performance in the detection of these processes has been better with regard to hotspots compared to escalation and diffusion processes. Escalation processes have been relatively difficult to measure since they occur in small areas within hotspots and might therefore encounter too few observations to be properly estimated. Another issue has been raised in the identification of "true" diffusion processes, whose values are averaged within wide areas when surrounding large hotspots, which could therefore obscure diffusion and dissipation processes compensating their antagonistic effects when averaged within the neighbourhood of hotspots.

Furthermore, the results from Chapter 6 have indicated that diffusion processes tend to occur mainly in failed states and in more economically developed areas in comparison to world's averages. From a global point of view, these findings have brought support to the failed state theory but contradicted the social disorganisation theory (in its socio-economic part) in particular. From a local point of view, the claims of the failed state theory with regard to the diffusion of the number of lethal attacks do not hold. The number of lethal attacks appears to spread towards neighbouring areas that are located within stronger states in comparison with the characteristics observed at the origin of the spread. Thus, the hierarchical theory provides suitable explanations of the diffusion of number of lethal terrorist attacks at a local level.

7.3 Shortcomings and Potential Improvements

As acknowledged in Sections 5.4 and 6.4, this work is subject to important shortcomings. The main fundamental issue that concerns all empirical studies of terrorism remains the absence of any consensual definition(s) of the concept of terrorism (Beck and Miner, 2013;

Jackson, 2016). This is certainly the main barrier to a wide generalisability of the results of this present work. Therefore, it is worth remembering that this work is based on a relatively narrow view of terrorism, where subjectivity in its interpretation cannot be excluded, and is therefore rightly open to criticism. The choice of GTD as the terrorism database is a major step to mitigate this issue.

The results of the model used to capture the spatial dynamics of lethal terrorism have relied on the assumption of using exhaustive and accurate data on terrorism and covariates. Keep in mind that this assumption does not reflect the relatively poor quality of currently available geolocalised data on worldwide terrorism. It is worth reiterating that the findings of this study cannot be generalised to state terrorism, which has been excluded from the analysis. Along with the potential inaccuracies present in the data, one cannot exclude the omission of relevant drivers that occur e.g. at the individual and group level, such as psychological processes (Crenshaw, 1983, p. 29; Wilkinson, 1990, p. 151; Brynjar and Skjølberg, 2000; Richardson, 2006, pp. 92-93) and the reciprocal interactions between counterterrorism and terrorism, for example (English, 2010; Hoffman, 2002).

As mentioned in Sections 5.4 and 6.4, this study might benefit from further improvements. One might reasonably assume that spatial correlation related to mass-casualty attacks extends into a larger spatial range, via a broader diffusion through media for example. Consequently, the (second-order) stationarity assumptions might be relaxed by using non-stationary covariance functions, such as the Sampson-Guttorp method, parametric non-stationary covariance function, convolution method, mixture of different metric models, or non-stationary adaptive spectrum (Yip, 2010). However, the covariance function might lose its crucial property of positive definiteness in its non-stationary version, which additionally, may result in higher computational costs required to mitigate the issues (Bolin, 2012, p. 7). The applicability of non-stationary models in the analysis of terrorism opens questions left for future investigation.

Moreover, finer temporal scales (quarter, month, week, or day) might be used along with more complex non-separable models. Furthermore, the incorporation of additional covariates that might be available in the future would allow the assessment of complementary theories that explain diffusion processes through *non-contagious* factors, e.g. modern mass media (Wilkinson, 1979, p. 103; Mazur, 1982; Crenshaw, 1991; Martin, 1990, pp. 161-162; Weimann, 2008; Nacos, 2010), financial support (Heyman and Mickolus, 1980), transnational collaboration (Heyman and Mickolus, 1980; Crenshaw, 1983, p. 17; Brynjar and

Skjølberg, 2000), freedom of movement and transportation (Heyman and Mickolus, 1980; Wilkinson, 1979, p. 189), and the dynamics of the terrorist group (e.g. tactical change) (LaFree et al., 2012; Raghavan et al., 2013).

Subjectivity remains in the choice of prior used in the Bayesian hierarchical framework, which might have affected the results of the Bernoulli (Equation 5.1c) and Poisson (Equation 5.2c) models. Moreover, the thresholds used in the detection of hotspots, escalation, and diffusion processes are not exempt from subjectivity. In order to mitigate any potential negative effects on the results, efforts have been undertaken to choose reasonable values in line with the purpose of the study and consistent with the scale in which the phenomenon is studied. Further studies are required to investigate the effects of variation in the threshold values on the performance of the detection of the processes under investigation.

Furthermore, it should be mentioned that the suggested approach drawing from the Bernoulli and Poisson models to detecting hotspot (Equation 6.1), escalation (Equation 6.2), and diffusion (Equation 6.4) processes represents only one among many other possible approaches. More particularly, the approach could be improved by e.g. taking into account complex space-time dependencies present in the data, while relaxing theoretical assumptions through the combination of non-parametric and parametric approaches, as suggested in Cameletti et al. (2013a). Future studies might be carried out in order to compare the performance of several methods used in the identification of spatio-dynamic processes of terrorism.

7.4 Policy Implications and Conclusion

Despite its aforementioned shortcomings, this study has proposed a rigorous framework to investigating the spatial dynamics of lethal terrorism across the world. It has provided a measure of uncertainty in the predictions, which is crucial for policy-makers to make informed decisions (Zammit-Mangion et al., 2013, p. 64) or to evaluate the impact of counterterrorism policies (Perl, 2007). From a counterterrorism point of view, the risk that a given location will encounter a lethal attack based on the posterior mean probability of lethal attacks of, say 80%, within a [10% - 90%] credible interval (CI) cannot be equally considered if the posterior mean probability of lethal attacks is also 80% but within a [75% - 90%] CI. Despite having the same point estimate in the mean of the risk, more vigilance is required in the case of the latter prediction, which shows less uncertainty in its prediction.

This valuable information on uncertainty is all the more important when resources are limited and need to be allocated in areas that are the most likely to be at high risk (Porter and White, 2012).

Despite the potential consequences of terrorism far beyond the location of the attacks, terrorism still occurs at very specific locations. As shown in this work, the probability and the number of lethal attacks are not uniformly distributed within countries; rather terrorism aggregates into sub-national clusters. Moreover, some locations are more prone to escalation (i.e. intensification of terrorism within hotspots) or diffusion processes, which might increase the level of terrorist activity in the neighbourhood of hotspots. Ignoring these properties of terrorism might lead to ill-advised policy recommendations.

While important counterterrorism efforts have been undertaken, most strategies employed to prevent and combat terrorism have not been accurately assessed (Lum et al., 2009). Ineluctably, country-level assessments neither allow capturing the risk of lethal terrorism — which is inherently localised and tends to considerably vary within each country —, nor the efficiency of counterterrorism measures. As a result, country-level analyses are irrelevant for most counterterrorism implementations that require valuable information on a local-scale in order to implement efficient preventive measures, e.g. surveillance using closed circuit television (CCTV) — whose efficiency and real impact on crime reduction has been under controversy in Europe and in the US (Galdon Clavell et al., 2012; Welsh and Farrington, 2004).

This study could therefore benefit policy-makers needing flexible and efficient decision-support tools through its fine-scale level of analysis. Drawing from the results of this analysis, a careful monitoring of the spatial dynamics of the attacks perpetrated by lethal terrorist groups currently in activity, such as Al-Qaeda, ISIS, or Boko Haram, might prevent further attacks and eventually reduce the number of victims worldwide. Ultimately, the findings of this PhD thesis contribute to improving our understanding of the spatial dynamics of lethal terrorism that occurred across the world from 2002 to 2013.

References

Abadie, A. (2006). Poverty, political freedom and the roots of terrorism. *The American Economic Review*, 96:50–56.

Abrahamsen, P. (1997). *A review of Gaussian random fields and correlation functions*. Norsk Regnesentral/Norwegian Computing Center.

Ackerman, W. V. and Murray, A. T. (2004). Assessing spatial patterns of crime in Lima, Ohio. *Cities*, 21(5):423–437.

Adler, R. J. and Taylor, J. E. (2009). *Random Fields and Geometry. Springer Monographs in Mathematics*. Springer.

Akaike, H. (1974). A new look at the statistical model identification. *IEE Transactions on Automatic Control*, 19:716–723.

Aldrich, J. (1997). RA Fisher and the making of maximum likelihood 1912-1922. *Statistical Science*, 12(3):162–176.

Alegana, V. A., Kigozi, S. P., Nankabirwa, J., Arinaitwe, E., Kigozi, R., Mawejje, H., Kilama, M., Ruktanonchai, N. W., Ruktanonchai, C. W., Drakeley, C., Lindsay, S. W., Greenhouse, B., Kamya, M. R., Smith, D. L., Atkinson, P. M., Dorsey, G., and Tatem, A. J. (2016). Spatio-temporal analysis of malaria vector density from baseline through intervention in a high transmission setting. *Parasites & Vectors*, 9(1):637.

Alexander, Y. and Pluchinsky, D. A. (1992). *Europe's red terrorists: The fighting communist organizations*. Frank Cass and Company Ltd.

Aljazeera (2009). Timeline: Taliban in Afghanistan. Key events and developments related to Taliban. http://www.aljazeera.com/news/asia/2009/03/2009389217640837.html. [Accessed 12 June 2015].

Amante, C. and Eakins, B. (2009). ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA. http://www.ngdc.noaa.gov/mgg/global/global.html. [Accessed 7 December 2014].

Amnesty International (2013). Annual report: The state of the world's human rights. http://www.amnesty.org/en/annual-report/2013/introduction. [Accessed 12 June 2014].

Andresen, M. A. (2006). Crime measures and the spatial analysis of criminal activity. *British Journal of Criminology*, 46(2):258–285.

Anselin, L. (1990). Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science*, 30(2):185–207.

Anselin, L., Cohen, J., Cook, D., Gorr, W., and Tita, G. (2000). Spatial analyses of crime. *Criminal justice*, 4(2):213–262.

Anselin, L. and Griffith, D. A. (1988). Do Spatial Effecfs Really Matter in Regression Analysis? *Papers in Regional Science*, 65(1):11–34.

Argomaniz, J. (2011). *The EU and Counter-Terrorism: Politics, Polity and Policies After 9/11*. Routledge, London, UK, 1st edition.

Asensio, M. and Ferragut, L. (2002). On a wildland fire model with radiation. *International Journal for Numerical Methods in Engineering*, 54(1):137–157.

Baddeley, A. (2008). Analysing spatial point patterns in R. Technical report, Technical report, CSIRO, 2010. Version 4. Available at www. csiro. au/resources/pf16h. html.

Baddeley, A. (2010). *Modeling Strategies*, pages 339–370. CRC Press, Boca Raton, FL, USA.

Baddeley, A. and Turner, R. (2014). *spatstat: Spatial Point Pattern analysis, model-fitting, simulation, tests*. R package version 1.34.1.

Banerjee, S., Gelfand, A. E., Finley, A. O., and Sang, H. (2008). Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(4):825–848.

Barros, C. P. (2003). An intervention analysis of terrorism: The Spanish ETA case. *Defence and Peace Economics*, 14:401–412.

Bartels, R. (1982). The rank version of von Neumann's ratio test for randomness. *Journal of the American Statistical Association*, 77(377):40–46.

Bashir, U., Ul-Haq, I., Gillani, A. H., and Muhammad, S. (2013). Influence of terrorist activities on financial markets: Evidence from KSE. *Financial Assets and Investing*, 2:5–13.

Basuchoudhary, A. and Shughart, W. F. (2010). On ethnic conflict and the origins of transnational terrorism. *Defence and Peace Economics*, 21(1):65–87.

Beck, C. J. and Miner, E. (2013). Who gets designated a terrorist and why? *Social Forces*, 91:837–872.

Beck, N., Gleditsch, K. S., and Beardsley, K. (2006). Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly*, 50(1):27–44.

Behlendorf, B., LaFree, G., and Legault, R. (2012). Microcycles of violence: Evidence from terrorist attacks by ETA and the FMLN. *Journal of Quantitative Criminology*, 28:49–75.

Benmelech, E., Berrebi, C., and Klor, E. F. (2012). Economic conditions and the quality of suicide terrorism. *Journal of Politics*, 74:113–128.

Bennett, A. F. (2005). *Inverse modeling of the ocean and atmosphere*. Cambridge University Press.

Berman, E. and Laitin, D. D. (2008). Religion, terrorism and public goods: Testing the club model. *Journal of Public Economics*, 92(10):1942–1967.

Bernardo, J. M. and Smith, A. F. (2010). *Bayesian theory*. John Wiley & Sons.

Berrebi, C. and Lakdawalla, D. (2007). How does terrorism risk vary across space and time? an analysis based on the Isreali experience. *Defence and Peace Economics*, 18:113–131.

Berrebi, C. and Ostwald, J. (2011). Earthquakes, hurricanes, and terrorism: Do natural disasters incite terror? *Public Choice*, 149:383–403.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36:192–236.

Besag, J. (1975). Statistical analysis of non-lattice data. *The Statistician*, pages 179–195.

Bhavnani, R. and Choi, H. J. (2012). Modeling civil violence in Afghanistan: Ethnic geography, control, and collaboration. *Complexity*, 17(6):42–51.

Bilal, A. R., Talib, N. B. A., Haq, I. U., Khan, M. N. A. A., and Islam, T. (2012). How terrorism and macroeconomic factors impact on returns: A case study of Karachi stock exchange. *World Applied Sciences Journal*, 19:1575–1584.

Bivand, R. (2015). Creating neighbours. http://cran.r-project.org/web/packages/spdep/vignettes/nb.pdf. [Accessed 12 March 2015].

Black, N. (2013). When have violent civil conflicts spread? Introducing a dataset of substate conflict contagion. *Journal of Peace Research*, 50(6):751–759.

Blangiardo, M. and Cameletti, M. (2015). *Spatial and Spatio-temporal Bayesian Models with R-INLA*. John Wiley & Sons.

Blomberg, S. B., Hess, G. D., and Weerapana, A. (2004). Economic conditions and terrorism. *European Journal of Political Economy*, 20(2):463–478.

Blomberg, S. B. and Rosendorff, P. B. (2009). A gravity model of globalization, democracy and transnational terrorism. In Hess, G. D., editor, *Guns and Butter*, pages 25–156. Cambridge MIT Press, Cambridge, MA, USA.

Blum, A., Asal, V., Wilkenfeld, J., Steinbruner, J., Ackerman, G., Gurr, T. R., Stohl, M., Post, J. M., Sinai, J., LaFree, G., Dugan, L., Franke, D., Stanislawski, B. H., Sheffer, G., Lichbach, M. I., and Sandler, T. (2005). Nonstate actors, terrorism, and weapons of mass destruction. *International Studies Review*, 7:133–170.

Bochner, S. (1955). *Harmonic analysis and the theory of probability*. University of California press.

Bogen, K. T. and Jones, E. D. (2006). Risks of mortality and morbidity from worldwide terrorism: 1968-2004. *Risk Analysis*, 26:45–59.

Bolin, D. (2012). *Models and methods for random fields in spatial statistics with computational efficiency from Markov properties*. PhD thesis, Lund University.

Bolin, D. and Lindgren, F. (2011). How do markov approximations compare with other methods for large spatial data sets? *arXiv preprint arXiv:1106.1980*.

Bollerslev, T. (1986). Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Borum, R. (2011). Radicalization into violent extremism i: A review of social science theories. *Journal of Strategic Security*, 4(4):2.

Box, G. E. P. and Draper, N. R. (1987). *Empirical Model Building and Response Surfaces*. John Wiley & Sons, New York, NY, USA.

Boyle, M. J. (2009). Bargaining, fear, and denial: Explaining violence against civilians in Iraq 2004–2007. *Terrorism and Political Violence*, 21(2):261–287.

Braithwaite, A. and Li, Q. (2007). Transnational terrorism hot spots: Identification and impact evaluation. *Conflict Management and Peace Science*, 24:281–296.

Brandt, P. T. and Sandler, T. (2012). A Bayesian Poisson vector autroregression model. *Political Analysis*, 20:292–315.

Brandt, P. T. and Williams, J. T. (2001). A linear Poisson autoregressive model: The Poisson ar(p) model. *Political Analysis*, 9:164–184.

Brantingham, P. and Brantingham, P. (1995). Criminality of place. *European Journal on Criminal Policy and Research*, 3(3):5–26.

Brooks, S., Gelman, A., Jones, G., and Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*. CRC press.

Brosius, H.-B. and Weimann, G. (1991). The contagiousness of mass-mediated terrorism. *European Journal of Communication*, 6(1):63–75.

Brown, D., Dalton, J., and Hoyle, H. (2004). Spatial forecast methods for terrorist events in urban environments. *Intelligence and Security Informatics*, 3073:426–435.

Brunsdon, C., Fotheringham, S., and Charlton, M. (1998). Geographically weighted regression-modelling spatial non-stationarity. *The Statistician*, 47:431–443.

Brynjar, L. and Skjølberg, K. (2000). Why terrorism occurs: a survey of theories and hypotheses on the causes of terrorism.

Buhaug, H., Gates, S., and Lujala, P. (2009). Geography, rebel capability, and the duration of civil conflict. *Journal of Conflict Resolution*, 53(4):544–569.

Buhaug, H. and Gleditsch, K. S. (2008). Contagion or confusion? Why conflicts cluster in space. *International Studies Quarterly*, 52:215–233.

Buhaug, H. and Rød, J. K. (2006). Local determinants of African civil wars, 1970–2001. *Political Geography*, 25(3):315–335.

Burgoon, B. (2006). On welfare and terror. *Journal of Conflict Resolution*, 50:176–203.

Cameletti, M., Ignaccolo, R., and Bande, S. (2011). Comparing spatio-temporal models for particulate matter in piemonte. *Environmetrics*, 22(8):985–996.

Cameletti, M., Ignaccolo, R., and Sylvan, D. (2013a). Assessment and visualization of threshold exceedance probabilities in complex space–time settings: A case study of air quality in Northern Italy. *Spatial Statistics*, 5:57–68.

Cameletti, M., Lindgren, F., Simpson, D., and Håvard (2013b). Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *Advances in Statistical Analysis*, 97:109–131.

Carlin, B. P. and Louis, T. A. (2008). *Bayesian methods for data analysis*. CRC Press, third edition edition.

Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, 46(3):167–174.

Casetti, E. (1972). Generating models by the expansion method: Application to geographical research. *Geographical Analysis*, 4:81–91.

CEACS (2013). Explaining terrorist and insurgent behavior. http://www.march.es/ceacs/proyectos/dtv/datasets.asp. [Accessed 12 June 2014].

Chainey, S., Tompson, L., and Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, 21(1):4–28.

Chalk, P. (1995). The liberal democratic response to terrorism. *Terrorism and Political Violence*, 7(4):10–44.

Chavanasporn, W. (2010). *Application of stochastic differential equations and real option theory in investment decision problems*. PhD thesis, University of St Andrews, UK.

Chen, X. and Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21):8589–8594.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327–335.

Chiozza, G. (2002). Is there a clash of civilizations? evidence from patterns of international conflict involvement, 1946-1997. *Journal of Peace Research*, 39:711–734.

Chomsky, N. (2002). *Pirates and Emperors, Old and New: International Terrorism in the Real World*. Pluto Press, new edition edition.

Chomsky, N. (2003). *Power and Terror: Post-9/11 Talks and Interviews*. Seven Stories Press, New York, junkerman,john and masakazu,takei edition.

Chomsky, N. (2006). *Failed States: The Abuse of Power and The Assault on Democracy*. Penguin Books.

CIESIN (2005). Gridded Population of the World, Version 3 (GPWv3): Population Density Grid. Data from Center for International Earth Science Information Network (CIESIN). Columbia University and Centro Internacional de Agricultura Tropical. http://dx.doi.org/10.7927/H4XK8CG2. [Accessed 12 December 2014].

Clauset, A., Young, M., and Gleditsch, K. S. (2007). On the frequency of severe terrorist events. *Journal of Conflict Resolution*, 51:58–87.

Cliff, C. and First, A. (2013). Testing for contagion/diffusion of terrorism in state dyads. *Studies in Conflict & Terrorism*, 36(4):292–314.

Coaffee, J. (2010). Protecting vulnerable cities: the UK's resilience response to defending everyday urban infrastructure. *International Affairs*, 86(4):939–954.

Cohen, J. and Tita, G. (1999). Diffusion in homicide: Exploring a general method for detecting spatial diffusion processes. *Journal of Quantitative Criminology*, 15(4):451–493.

Cohen, L. E. and Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American sociological review*, pages 588–608.

Collier, P. and Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4):563–595.

Congdon, P. (2007). *Bayesian statistical modelling*. John Wiley & Sons, Chichester, England, 2nd edition.

Conteh-Morgan, E. (2004). *Collective political violence: An introduction to the theories and cases of violent conflicts*. Routledge, New York, NY, USA.

Costa, A. C. C., Codeço, C. T., Honório, N. A., Pereira, G. R., Pinheiro, C. F. N., and Nobre, A. A. (2015). Surveillance of dengue vectors using spatio-temporal Bayesian modeling. *BMC medical informatics and decision making*, 15(1):93.

Crelinsten, R. (2009). *Counterterrorism*. Polity Press, Cambridge, UK.

Crenshaw, M. (1981). The causes of terrorism. *Comparative Politics*, 13:379–399.

Crenshaw, M. (1983). *Introduction: Reflections on the Effects of Terrorism*, pages 1–37. Wesleyan University Press, Middletown, CT, USA.

Crenshaw, M. (1990). *The Causes of Terrorism*, pages 113–126. St. Martin's Press, Inc, New York, NY, USA.

Crenshaw, M. (1991). How terrorism declines. *Terrorism and Political Violence*, 3:69–87.

Crenshaw, M. (2000). The psychology of terrorism: An agenda for the 21st century. *Political psychology*, 21(2):405–420.

Crenshaw, M. (2014). Terrorism research: The record. *International Interactions*, 40(4):556–567.

Cressie, N. (1991). *Statistics for Spatial Data*. John Wiley & Sons, London, UK, 1st edition.

Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*. John Wiley & Sons, Hoboken, NJ, USA, 1st edition.

Cronin, A. K. (2003). Behind the curve: Globalization and international terrorism. *International Security*, 27:30–58.

Cryer, J. D. and Chan, K.-S. (2008). *Time Series Analysis With Applications in R*. Springer, New York, NY, USA, 2nd edition.

Cseke, B., Zammit-Mangion, A., Heskes, T., and Sanguinetti, G. (2015). Sparse approximate inference for spatio-temporal point process models. *Journal of the American Statistical Association*, (just-accepted):1–52.

Da Prato, G., Debussche, A., and Temam, R. (1994). Stochastic burgers'equation. *Nonlinear Differential Equations and Applications NoDEA*, 1(4):389–402.

Darmofal, D. (2015). *Spatial Analysis for the Social Sciences*. Analytical Methods for Social Research. Cambridge University Press.

Davis, P. J. (1959). Leonhard Euler's Integral: A Historical Profile of the Gamma Function: In Memoriam: Milton Abramowitz. *The American Mathematical Monthly*, 66(10):849–869.

Davis, W. W., Duncan, G. T., and Siverson, R. M. (1978). The dynamics of warfare: 1816-1965. *American Journal of Political Science*, 22(4):772–792.

De la Calle, L. and Sánchez-Cuenca, I. (2011). The quantity and quality of terrorism the DTV dataset. *Journal of Peace Research*, 48(1):49–58.

Dellaportas, P. and Roberts, G. O. (2003). An introduction to MCMC. In *Spatial statistics and computational methods*, pages 1–41. Springer.

Deng, H. and Wickham, H. (2011). Density estimation in R. Technical report, had.co.nz.

Diggle, P. (1985). A kernel method for smoothing point process data. *Applied Statistics*, 34(2):138–147.

Diggle, P. and Gabriel, E. (2010). *Handbook of spatial statistics*, chapter Spatio-Temporal Point Processes, pages 449–462. CRC press.

Diggle, P. J. (2007). *Spatio-Temporal Point Processes: Methods and Applications*, pages 2–45. Chapman & Hall/CRC, Boca Raton, FL, USA.

Diggle, P. J. (2014). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. CRC Press, Boca Raton, FL, USA, 3rd edition.

D'Orsogna, M. R. and Perc, M. (2014). Statistical physics of crime: A review. *Physics of life reviews*.

Drake, C. J. M. (1998). *Terrorists' Target Selection*. St. Martin's Press, New York, NY, USA, 1st edition.

Drakos, K. (2007). The size of under-reporting bias in recorded transnational terrorist activity. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(4):909–921.

Drakos, K. and Gofas, A. (2006a). The devil you know but are afraid to face: Underreporting bias and its distorting effects on the study on terrorism. *Journal of Conflict Resolution*, 50:714–735.

Drakos, K. and Gofas, A. (2006b). In search of the average transnational terrorist attack venue. *Defence and Peace Economics*, 17:73–93.

Dreher, A. and Fischer, J. A. V. (2010). Government decentralization as a disincentive for transnational terror? An empirical analysis. *International Economic Review*, 51:981–1002.

Drton, M. and Plummer, M. (2017). A Bayesian information criterion for singular models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(2):323–380.

Dugan, L. and Chenoweth, E. (2013). Government actions in terror environments (GATE): A methodology that reveals how governments behave toward terrorists and their constituencies. In *Handbook of Computational Approaches to Counterterrorism*, pages 465–486. Springer.

Ebener, S., Murray, C., Tandon, A., and Elvidge, C. C. (2005). From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery. *International Journal of Health Geographics*, 4(1):5.

Eck, J., Chainey, S., Cameron, J., and Wilson, R. (2005). Mapping crime: Understanding hotspots. chapter 1: Crime hot spots: What they are, why we have them, and how to map them. Technical report, U.S. National Institute of Justice.

Eck, J. E., Gersh, J. S., and Taylor, C. (2000). *Finding crime hot spots through repeat address mapping*, pages 49–64. Thousand Oaks: Sage Publications, London, UK.

Eidsvik, J., Finley, A. O., Banerjee, S., and Rue, H. (2012). Approximate Bayesian inference for large spatial datasets using predictive process models. *Computational Statistics & Data Analysis*, 56:1362–1380.

Eidsvik, J., Shaby, B. A., Reich, B. J., Wheeler, M., and Niemi, J. (2014). Estimation and prediction in spatial models with block composite likelihoods. *Journal of Computational and Graphical Statistics*, 23(2):295–315.

Elbakidze, L. and Jin, Y. (2012). Victim countries of transnational terrorism: An empirical characteristics analysis. *Risk Analysis*, 32:2152–2164.

Elster, J. (2009). *Reason and Rationality*. Princeton University Press, Princeton, NJ, USA, 1st edition.

Elvidge, C. D., Hsu, F.-C., Baugh, K. E., and Ghosh, T. (2013). *National Trends in Satellite Observed Lighting: 1992-2012*, pages 97–120. CRC Press, Boca Raton, FL, USA.

Elvidge, C. D., Safran, J., Tuttle, B., Sutton, P., Cinzano, P., Pettit, D., Arvesen, J., and Small, C. (2007). Potential for global mapping of development via a nightsat mission. *GeoJournal*, 69(1-2):45–53.

Encyclopaedia Britannica Online (2014). “diffusion”. Encyclopaedia Britannica Inc. [Accessed 17 December 2014].

Enders, W. and Hoover, G. A. (2012). The nonlinear relationship between terrorism and poverty. *American Economic Review*, 102:267–272.

Enders, W., Parise, G. F., and Sandler, T. (1992). A time-series analysis of transnational terrorism: Trends and cycles. *Defence and Peace Economics*, 3(4):305–320.

Enders, W. and Sandler, T. (1991). Causality between transnational terrorism and tourism: The case of Spain. *Studies in Conflict & Terrorism*, 14(1):49–58.

Enders, W. and Sandler, T. (1993). The effectiveness of anti-terrorism policies: A vector-autogression-intervention analysis. *American Political Science Review*, 87:829–844.

Enders, W. and Sandler, T. (1996). Terrorism and foreign direct investment in Spain and Greece. *Kyklos*, 49:331–352.

Enders, W. and Sandler, T. (1999). Transnational terrorism in the post-cold war era. *International Studies Quarterly*, 43:145–167.

Enders, W. and Sandler, T. (2000). Is international terrorism becoming more threatening? A time series investigation. *Journal of Conflict Resolution*, 44:307–322.

Enders, W. and Sandler, T. (2002). Patterns of transnational terrorism, 1970-1999: Alternative time-series estimates. *Intenational Studies Quarterly*, 46:145–165.

Enders, W. and Sandler, T. (2005). Transnational terrorism 1968-2000: Thresholds, persistence and forecasts. *Southern Economic Journal*, 71:467–482.

Enders, W. and Sandler, T. (2006). Distribution of transnational terrorism among countries by income class and geography after 9/11. *International Studies Quarterly*, 50(2):367–393.

Enders, W., Sandler, T., and Gaibulloev, K. (2011). Domestic versus transnational terrorism: Data, decomposition, and dynamics. *Journal of Peace Research*, 48:319–337.

Enders, W. and Su, X. (2007). Rational terrorists and optimal network structure. *The Journal of Conflict Resolution*, 51:33–57.

Engene, J. O. (2007). Five decades of terrorism in Europe: The TWEED dataset. *Journal of Peace Research*, 44:109–121.

English, R. (2010). *Terrorism: How to respond*. Oxford University Press.

Esteban, J., Mayoral, L., and Ray, D. (2012). Ethnicity and conflict: Theory and facts. *Science*, 336(6083):858–865.

Eubank, W. and Weinberg, L. (2001). Terrorism and democracy: Perpetrators and victims. *Terrorism and Political Violence*, 13(1):155–164.

Evans, D. (2001). Spatial analyses of crime. *Geography*, 86(3):211–223.

Evans, L. C. (2010). *Partial Differential Equations. Graduate Studies in Mathematics*, volume 19. American Mathematical Society, Providence, RI, USA, 2nd edition.

Faber, J., Houweling, H. W., and Siccama, J. G. (1984). Diffusion of war: Some theoretical considerations and empirical evidence. *Journal of Peace Research*, 21:277–288.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *The American Political Science Review*, 97(1):75–90.

Fokianos, K. (2012). *Count Time Series Models*, pages 315–347. Elsevier, Oxford, UK.

Forsberg, E. (2014). Diffusion in the study of civil wars: A cautionary tale. *International Studies Review*, 16:188–198.

Fotheringham, A. S., Brunsdon, C., and Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons Ltd, West Sussex, UK, 1st edition.

Fouedjio, F. (2016). Second-order non-stationary modeling approaches for univariate geostatistical data. *Stochastic Environmental Research and Risk Assessment*, pages 1–20.

Freytag, A., Krueger, J. J., and Meierrieks, D. (2010). The origins of terrorism: Cross-country estimates on socio-economic determinants of terrorism. Economics of Security Working Paper Series 27, DIW Berlin, German Institute for Economic Research.

Fuentes, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association*, 102(477):321–331.

Furrer, R., Genton, M. G., and Nychka, D. (2012). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*.

Gaibulloev, K. and Sandler, T. (2008). Growth consequences of terrorism in Western Europe. *Kyklos*, 61:411–424.

Galdon Clavell, G., Lojo, L. Z., and Romero, A. (2012). CCTV in Spain: An empirical account of the deployment of video-surveillance in a Southern-European country. *Information Polity*, 17(1):57–68.

Galton, F. (1889). Galton's comment on: Tylor's method of investigating the development of institutions applied to the laws of marriage and descent. *Journal of the Royal Anthropological Institute*, 18(3):245–272.

Gao, P., Guo, D., Liao, K., Webb, J. J., and Cutter, S. L. (2013). Early Detection of Terrorism Outbreaks Using Prospective Space-Time Scan Statistics. *The Professional Geographer*, 65:676–691.

Gassebner, M. and Luechinger, S. (2011). Lock, stock, and barrel: A comprehensive assessment of the determinants of terror. *Public Choice*, 149:235–261.

GDELT (2013). *GDELT Data Format Codebook V 1.03*. V 1.03.

Geirsson, O. P., Hrafnkelsson, B., and Simpson, D. (2015). Computationally efficient spatial modeling of annual maximum 24-h precipitation on a fine grid. *Environmetrics*, 26(5):339–353.

Gelman, A., Hwang, J., and Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6):997–1016.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741.

Gibbs, J. P. (1989). Conceptualization of terrorism. *American Sociological Review*, pages 329–340.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). *Markov chain Monte Carlo in practice*. Chapman and Hall, London, UK, 1st edition.

Gleason, J. M. (1980). A Poisson model of incidents of international terrorism in the United States. *Terrorism*, 4:259–265.

Gleditsch, K. S. (2007). Transnational dimensions of civil war. *Journal of Peace Research*, 44(3):293–309.

Gneiting, T. (2008). Editorial: probabilistic forecasting. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(2):319–321.

Gneiting, T., Genton, M. G., and Guttorp, P. (2006). Geostatistical space-time models, stationarity, separability, and full symmetry. *Monographs On Statistics and Applied Probability*, 107:151.

Gneiting, T. and Guttorp, P. (2010). *Handbook of spatial statistics*, chapter Continuous Parameter Spatio-Temporal Processes, pages 427–436. CRC press.

Gordon, R. A. (2015). *Applied statistics for the social and health sciences. Study Guide*. Cram101 Textbook Reviews. 1st edition.

Gorman, D. M., Gruenewald, P. J., and Waller, L. A. (2013). Linking places to problems: Geospatial theories of neighborhoods, alcohol and crime. *GeoJournal*, 78(3):417–428.

Griffith, D. A. (2003). *Spatial Autocorrelation and Spatial Filtering*. Springer-Verlag, Heidelberg, Germany.

Griffith, D. A. and Layne, L. J. (1999). *A Casebook for Spatial Statistical Data Analysis: A compilation of Analyses of Different Thematic Data Sets*. Oxford University Press, New York, NY, USA, 1st edition.

Gros, J.-G. (1996). Towards a taxonomy of failed states in the New World Order: decaying Somalia, Liberia, Rwanda and Haiti. *Third World Quarterly*, 17(3):455–472.

Grubesic, T. H. and Mack, E. A. (2008). Spatio-temporal interaction of urban crime. *Journal of Quantitative Criminology*, 24(3):285–306.

GTD (2014). Global Terrorism Database (GTD) Codebook: Inclusion Criteria and Variables. http://www.start.umd.edu/gtd/downloads/Codebook.pdf. [Accessed 9 December 2014].

Hägerstraand, T. (1970). What about people in regional science? *Papers in regional science*, 24(1):7–24.

Hakim, S. and Shachmurove, Y. (1996). Spatial and temporal patterns of commercial burglaries: The evidence examined. *American Journal of Economics and Sociology*, 55(4):pp. 443–456.

Hall, D. B. (2000). Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics*, 56:1030–1039.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, Princeton, NJ, USA, 1st edition.

Hamilton, L. C. and Hamilton, J. D. (1983). Dynamics of terrorism. *International Studies Quarterly*, 27:39–54.

Hammersley, J. M. and Handscomb, D. C. (1964). *Monte carlo methods*, volume 1. Methuen London.

Hammond, J. and Weidmann, N. B. (2014). Using machine-coded event data for the micro-level study of political violence. *Research & Politics*, 1(2):2053168014539924.

Hartikainen, J., Riihimäki, J., and Särkkä, S. (2011). Sparse spatio-temporal Gaussian processes with general likelihoods. *Artificial Neural Networks and Machine Learning–ICANN 2011*, pages 193–200.

Harvey, A. C. (1993). *Time Series Models*. MIT Press, Cambridge, MA, USA, 2nd edition.

Harvill, J. L. (2010). Spatio-temporal processes. *Wiley interdisciplinary reviews: computational statistics*, 2(3):375–382.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

Helman, G. B. and Ratner, S. R. (1992). Saving failed states. *Foreign policy*, 89(3):3–20.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2009). Measuring economic growth from outer space. no. w15199. In *National Bureau of Economic Research*.

Heyman, E. and Mickolus, E. (1980). Observations on “why violence spreads”. *International Studies Quarterly*, 24:299–305.

Hill, S. and Rothchild, D. (1986). The contagion of political conflict in africa and the world. *The Journal of Conflict Resolution*, 30:716–735.

Hoffman, B. (2001). Change and continuity in terrorism. *Studies in Conflict and Terrorism*, 24(5):417–428.

Hoffman, B. (2002). Rethinking terrorism and counterterrorism since 9/11. *Studies in Conflict and Terrorism*, 25(5):303–316.

Hoffman, B. (2006). *Inside Terrorism*. Columbia University Press, New York, NY, USA, rev. and expanded edition.

Holden, R. T. (1986). The contagiousness of aircraft hijacking. *American Journal of Sociology*, 91:874–904.

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

Holt, D., Steel, D., Tranmer, M., and Wrigley, N. (1996). Aggregation and ecological effects in geographically based data. *Geographical analysis*, 28(3):244–261.

Hsiang, S. M., Burke, M., and Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, 341(6151):1235367.

Huntington, S. (1993). The clash of civilizations? *Foreign Affairs*, 72:22–49.

Hyndman, R. J. (2011). Moving averages. In *International Encyclopedia of Statistical Science*, pages 866–869. Springer.

Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point patterns*. John Wiley & Sons Ltd, Chichester, West Sussex, UK, 1st edition.

Ingebrigtsen, R., Lindgren, F., and Steinsland, I. (2014). Spatial models with explanatory variables in the dependent structure. *Spatial Statistics*, 8:20–38.

International Court of Justice (1986). Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America). Judgment of 27 June 1986. http://www.icj-cij.org/docket/?sum=367&p1=3&p2=3&case=70&p3=5. [Accessed 26 August 2015].

Isham, V. (2010). *Spatial Point Process Models*, pages 283–298. CRC Press, Boca Raton, FL, USA.

Jackson, R. (2016). *The Ashgate Research Companion to Political Violence*, chapter Critical Accounts of Terrorism, pages 47–62. Routledge.

Jacquez, G. M. (1996). A k nearest neighbour test for space-time interaction. *Statistics in medicine*, 15(18):1935–1949.

Jenkins, B. M. (1975). International terrorism: A new mode of conflict. In Carlton, D. and Schaerf, C., editors, *California Seminar on Arms Control and Foreign Policy. Number 48*, London, UK. Croom Helm.

Jenkins, B. M. (1983). New mode of conflict. Technical report, RAND, Santa Monica, CA, USA.

Jenkins, B. M. and Johnson, J. (1975). International terrorism: A chronology, 1968-1974. Technical report, DTIC Document, Santa Monica, CA, USA.

Jensen, M. (2013). Discussion Point: The Benefits and Drawbacks of Methodological Advancements in Data Collection and Coding: Insights from the Global Terrorism Database (GTD). goo.gl/yMQosr. [Accessed 8 March 2015].

Johnson, S. (2008). Repeat burglary victimisation: a tale of two theories. *Journal of Experimental Criminology*, 4(3):215–240.

Johnson, S. D., Bernasco, W., Bowers, K. J., Elffers, H., Ratcliffe, J., Rengert, G., and Townsley, M. (2007). Space–time patterns of risk: a cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology*, 23(3):201–219.

Johnson, S. D., Lab, S. P., and Bowers, K. J. (2008). Stable and fluid hotspots of crime: differentiation and identification. *Built Environment (1978-)*, pages 32–45.

Jongman, A. J. (1992). Trends in international and domestic terrorism in Western Europe, 1968–1988. *Terrorism and Political Violence*, 4(4):26–76.

Jordan, F. (2008). Predicting target selection by terrorists: A network analysis of the 2005 London underground attacks. *International journal of critical infrastructures*, 4(1):206–214.

Junkerman, J. (2002). Power and terror: Noam chomsky in our times. Film directed by John Junkerman. Documentary.

Kaspi, H., Ramanan, K., et al. (2013). Spde limits of many-server queues. *The Annals of Applied Probability*, 23(1):145–229.

Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103(484):1545–1555.

Kegley, C. W. (1990). *The Characteristics, Causes, and Controls of International Terrrorism: An Introduction*, pages 1–9. St. Martin's Press, Inc, New York, NY, USA.

Kellert, S. H. (1994). *In the wake of chaos: Unpredictable order in dynamical systems*. University of Chicago press.

King, R., Papathomas, M., and Thomas, L. (2014). Bayesian inference.

Kingman, J. F. C. (1992). *Poisson processes*, volume 3. Oxford university press.

Kirchgässner, G., Wolters, J., and Hassler, U. (2013). *Introduction to Modern Time Series Analysis*. Springer, Heidelberg, Germany, 2nd edition.

Klein, J. L. (1997). *Statistical Visions in Time: A History of Time Series Analysis, 1662-1938*. Cambridge University Press, Cambridge, UK, 1st edition.

Koch, M. T. and Cranmer, S. (2007). Testing the "Dick Cheney" hypothesis: do governments of the left attract more terrorism than governments of the right? *Conflict Management and Peace Science*, 24(4):311–326.

Krieger, T. and Meierrieks, D. (2011). What causes terrorism? *Public Choice*, 147:3–27.

Krige, D. G. (1951). A statistical approach to some mine valuation and allied problems on the Witwatersrand. *J. Chem., Metal. and Mining Soc. South Africa*, 52(6):119–139.

Krueger, A. B. and Laitin, D. D. (2008). *Kto Kogo?: A Cross-Country Study of the Origins and Targets of Terrorism*, pages 148–173. Cambridge University Press, Cambridge.

Krueger, A. B. and Maleckova, J. (2003). Education, poverty and terrorism: Is there a causal connection? *The Journal of Economic Perpectives*, 17:119–144.

Kurrild-Klitgaard, P., Justesen, M., and Klemmensen, R. (2006). The political economy of freedom, democracy and transnational terrorism. *Public Choice*, 128:289–315.

Kyung, M., Gill, J., and Casella, G. (2011). New findings from terrorism data: Dirichlet process random-effects models for latent groups. *Journal of the Royal Statistical Society. Series C. Applied Statistics*, 60:701–721.

LaFree, G. (2005). Are national crime trends converging? Evidence for homicide victimization rates, 1956 to 1998. *Sociological Quarterly*, 1(46):191–211.

LaFree, G. (2010). The Global Terrorism Database: Accomplishments and challenges. *Perspectives on Terrorism*, 4(1).

LaFree, G. and Dugan, L. (2004). How does studying terrorism compare to studying crime? In Deflem, M., editor, *Terrorism and Counter-Terrorism: Criminological perspectives*, volume 5, pages 53–74. Elsevier Ltd, Oxford, UK.

LaFree, G. and Dugan, L. (2007). Introducing the Global Terrorism Database. *Terrorism and Political Violence*, 19(2):181–204.

LaFree, G., Dugan, L., Xie, M., and Singh, P. (2012). Spatial and temporal patterns of terrorist attacks by ETA 1970 to 2007. *Journal of Quantitative Criminology*, 28:7–29.

LaFree, G., Morris, N. A., and Dugan, L. (2010). Cross-national patterns of terrorism comparing trajectories for total, attributed and fatal attacks, 1970–2006. *British Journal of Criminology*, 50(4):622–649.

LaFree, G., Yang, S.-M., and Crenshaw, M. (2009). Trajectories of terrorism. *Criminology & Public Policy*, 8(3):445–473.

Lai, B. (2007). "draining the swamp": an empirical examination of the production of international terrorism, 1968–1998. *Conflict Management and Peace Science*, 24(4):297–310.

Lai, Y.-S., Zhou, X.-N., Utzinger, J., and Vounatsou, P. (2013). Bayesian geostatistical modelling of soil-transmitted helminth survey data in the People's Republic of China. *Parasit Vectors*, 6:359.

Lakdawalla, D. and Zanjani, G. (2005). Insurance, self-protection, and the economics of terrorism. *Journal of Public Economics*, 89:1891–1905.

Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics*, 34(1):1–14.

Laqueur, W. (1999). *The New Terrorism: Fanaticism and the Arms of Mass Destruction*. Oxford University Press, New York, NY, USA, 1st edition.

Law, J., Quick, M., and Chan, P. (2014). Bayesian spatio-temporal modeling for analysing local patterns of crime over time at the small-area level. *Journal of Quantitative Criminology*, 30(1):57–78.

Leetaru, K. and Schrodt, P. A. (2013). GDELT: Global data on events, location, and tone, 1979–2012. In *Paper presented at the ISA Annual Convention*, volume 2, page 4.

Lenzi, A., Pinson, P., Clemmensen, L. H., and Guillot, G. (2016). Spatial models for probabilistic prediction of wind power with application to annual-average and high temporal resolution data. *Stochastic Environmental Research and Risk Assessment*, pages 1–17.

Lewis, E., Mohler, G., Brantingham, P. J., and Bertozzi, A. L. (2012). Self-exciting point process models of civilian deaths in Iraq. *Security Journal*, 25:244–264.

Li, Q. (2005). Does democracy promote or reduce transnational terrorist incidents? *Journal of Conflict resolution*, 49(2):278–297.

Lichbach, M. I. (1992). Nobody cites nobody else: Mathematical models of domestic political conflict. *Defence and Peace Economics*, 3(4):341–357.

Lindgren, F. (2012). Continuous domain spatial models in R-INLA. *The ISBA Bulletin*, 19(4):14–20.

Lindgren, F. and Rue, H. (2013). Bayesian spatial and spatiotemporal modelling with R-INLA. *Journal of Statistical Software*.

Lindgren, F. and Rue, H. (2015). Bayesian spatial modelling with R-INLA. *Journal of Statistical Software*, 63(19):1–25.

Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Linke, A. M., Witmer, F. D., and O'Loughlin, J. (2012). Space-time granger analysis of the war in Iraq: A study of coalition and insurgent action-reaction. *International Interactions*, 38(4):402–425.

Loftin, C. (1986). Assaultive violence as a contagious social process. *Bulletin of the New York Academy of Medicine*, 62(5):550.

Loh, J. M. (2008). A valid and fast spatial bootstrap for correlation functions. *The Astrophysical Journal*, 681(1):726.

Longley, P. A., Goodchild, M. F., and Maguire, D. J. (2001). *Geographic Information Systems and Science*. John Wiley & Sons, Ltd, Chichester, UK.

Los Angeles Times (2014). Spain still healing 10 years after Madrid '11-M' train bombings. http://articles.latimes.com/2014/mar/09/world/la-fg-spain-bombings-20140309. [Accessed 12 July 2014].

Lum, C., Haberfeld, M. M., Fachner, G., and Lieberman, C. (2009). Police activities to counter terrorism: What we know and what we need to know. In *To protect and to serve*, pages 101–141. Springer.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer research*, 27(2 Part 1):209–220.

Marshall, M. G., Jaggers, K., and Gurr, T. R. (2014). Polity IV project: Political regime characteristics and transitions, 1800-2013. http://www.systemicpeace.org/inscrdata.html. [Accessed 12 January 2015].

Martin, J. L. (1990). The media's role in international terrorism. In Kegley, C. W., editor, *International Terrorism: Characteristics, Causes, Controls*, chapter The Media's Role in International Terrorism, pages 158–162. St. Martin's Press, Inc.

Matérn, B. (1960). Spatial Variation: Meddelanden fran Statens Skogsforskningsinstitut. *Lecture Notes in Statistics*, 49(5).

Matheron, G. (1963). Principles of geostatistics. *Economic geology*, 58(8):1246–1266.

Matheron, G. (1971). *The theory of regionalized variables and its applications*, volume 5. École national supérieure des mines.

Mazur, A. (1982). Bomb threats and the mass media: Evidence for a theory of suggestion. *American Sociological Review*, 47:407–411.

Medina, R. M., Siebeneck, L. K., and Hepner, G. F. (2011). A geographic information systems (GIS) analysis of spatiotemporal patterns of terrorist incidents in Iraq 2004-2009. *Studies in Conflict & Terrorism*, 34:862–882.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.

Mickolus, E. (2003). *International terrorism attributes of terrorist events (ITERATE)*. Vinyard Software, Dunn Loring, VA, USA. Data codebook.

Mickolus, E. F. (1980). *Transnational Terrorism: A Chronology of Events, 1968-1979*. Greenwood Press, Westport, CT, USA.

Mickolus, E. F. (1993). *Terrorism, 1988-1991: A Chronology of Events and a Selectively Annotated Bibliography*, volume Bibliographies and Indexes in Military Studies. Greenwood Press, Westport, CT, USA.

Mickolus, E. F., Sandler, T., and Murdock, J. (1988). *International Terrorism in the 1980s: A Chronology (volume 1: 1980-1983)*. Iowa State University Press, Ames, IA, USA.

Mickolus, E. F., Sandler, T., and Murdock, J. (1989). *International Terrorism in the 1980s: A Chronology (volume 2: 1984-1987)*. Iowa State University Press, Ames, IA, USA.

Mickolus, E. F. and Simmons, S. L. (1997). *Terrorism, 1992-1995: A Chronology of Events a Selectively Annotated Bibliography*. Greenwood Press, Westport, CT, USA.

Mickolus, E. F. and Simmons, S. L. (2002). *Terrorism, 1996-2001: A Chronology of Events and a Selectively Annotated Bibliography (2 volumes)*. Greenwood Press, Westport, CT, USA.

Mickolus, E. F. and Simmons, S. L. (2006). *Terrorism, 2002-2004: A Chronology (3 volumes)*. Greenwood Press, Westport, CT, USA.

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

Miller, H. J. (2005). A measurement theory for time geography. *Geographical Analysis*, 37(1):17–45.

Mitchell, S. M. and Moore, W. H. (2002). Presidential uses of force during the Cold War: Aggregation, truncation, and temporal dynamics. *American Journal of Political Science*, pages 438–452.

Mohler, G. (2013). Modeling and estimation of multi-source clustering in crime and security data. *The Annals of Applied Statistics*, 7:1525–1539.

Mohler, G. (2014). Marked point process hotspot maps for homicide and gun crime prediction in Chicago. *International Journal of Forecasting*, 30(3):491–497.

Mohler, G., Short, M. B., Brantingham, P. J., Schoenberg, F. P., and Tita, G. E. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108.

Möller, A., Thorarinsdottir, T. L., Lenkoski, A., and Gneiting, T. (2015). Spatially adaptive, bayesian estimation for probabilistic temperature forecasts. *arXiv preprint arXiv:1507.05066*.

Morelli, M. and Rohner, D. (2014). Resource concentration and civil wars. In *National Bureau of Economic Research*.

Morgenstern, A. P., Velásquez, N., Manrique, P., Hong, Q., Johnson, N., and Johnson, N. (2013). Resource letter MPCVW-1: Modeling political conflict, violence, and wars: A survey. *American Journal of Physics*, 81(11):805–814.

Most, B. A. and Starr, H. (1980). Diffusion, reinforcement, geopolitics, and the spread of war. *The American Political Science Review*, 74:932–946.

Mueller, J. (2005). Six rather unusual propositions about terrorism. *Terrorism and Political Violence*, 17:487–505.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of econometrics*, 33(3):341–365.

Murphy, A. B. (2003). The space of terror. In Cutter, S. l., Richardson, D. B., and Wilbanks, T. J., editors, *The Geographical Dimension of Terrorism*, chapter The Space of Terror, pages 47–52. Routledge.

Musenge, E., Chirwa, T. F., Kahn, K., and Vounatsou, P. (2013). Bayesian analysis of zero inflated spatiotemporal HIV/TB child mortality data through the INLA and SPDE approaches: applied to data observed between 1992 and 2010 in rural North East South Africa. *International Journal of Applied Earth Observation and Geoinformation*, 22:86–98.

Mwalili, S. M., Lesaffre, E., and Declerck, D. (2008). The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. *Statistical Methods in Medical Research*, 17(2):123–139.

Nacos, B. L. (2010). Revisiting the contagion hypothesis: Terrorism, news coverage, and copycat attacks. *Perspectives on Terrorism*, 3(3):3–13.

NATO (2008). The mechanics of terrorism. http://www.nato.int/docu/review/2008/04/AP_COST/EN/index.htm. [Accessed 10 February 2015].

Nauenberg, M. (2001). *Isaac Newton's Natural Philosophy*. MIT Press.

Navy Times (2010). 10 years after Cole bombing, a different navy. http://archive.navytimes.com/article/20101011/NEWS/10110316/10-years-after-Cole-bombing-different-Navy. [Accessed 18 June 2015].

Nelson, A. (2008). Estimated travel time to the nearest city of 50,000 or more people in year 2000. http://bioval.jrc.ec.europa.eu/products/gam/. [Accessed 10 June 2014].

Nelson, A. L., Bromley, R. D., and Thomas, C. J. (2001). Identifying micro-spatial and temporal patterns of violent crime and disorder in the British city centre. *Applied Geography*, 21(3):249–274.

Nemeth, S. C. (2010). *A Rationalist Explanation of Terrorist Targeting*. PhD thesis, University of Iowa, USA.

Nemeth, S. C., Mauslein, J. A., and Stapley, C. (2014). The primacy of the local: Identifying terrorist hot spots using Geographic Information Systems. *The Journal of Politics*, 76:304–317.

Neumayer, E. and Plümper, T. (2010). Galton's problem and contagion in international terrorism along civilizational lines. *Conflict management and peace science*, 27(4):308–325.

NOAA (2014). *Version 4 DMSP-OLS Nighttime Lights Time Series*. National Oceanic and Atmospheric Administration, National Geophysical Data Center.

Nunn, S. (2007). Incidents of terrorism in the United States, 1997-2005. *Geographical Review*, 97:89–111.

O'Brien, C. C. (1983). *Terrrorism under Democratic Conditions: The Case of the IRA*, pages 91–104. Wesleyan University Press, Middletown, CT, USA.

Öcal, N. and Yildirim, J. (2010). Regional effects of terrorism on economic growth in Turkey: A geographically weighted regression approach. *Journal of Peace Research*, 47(4):477–489.

O'Loughlin, J. and Witmer, F. D. (2011). The localized geographies of violence in the North Caucasus of Russia, 1999–2007. *Annals of the Association of American Geographers*, 101(1):178–201.

O'Loughlin, J., Witmer, F. D., Linke, A. M., and Thorwardson, N. (2010). Peering into the fog of war: The geography of the Wikileaks Afghanistan war logs, 2004-2009. *Eurasian Geography and Economics*, 51(4):472–495.

Ord, J. K. and Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical analysis*, 27(4):286–306.

Oxford English Dictionary Online (2014). "diffusion, n.". Oxford University Press. [Accessed 17 December 2014].

Pape, R. A. (2006). Suicide Terrorism and Democracy: What We've Learned Since 9/11. Technical report, CATO Institute.

Pease, K., Britain, G., et al. (1998). *Repeat victimisation: Taking stock*. Number 90 in Crime Detection and Prevention Paper. Home Office Police Research Group, London, UK.

Perl, A. (2007). Combating terrorism: The challenge of measuring effectiveness. Technical report, Technical Re-port RL33160, Congressional Research Services. Available at http://fpc.state.gov/documents/organization/57513.pdf.

Philibert, J. (2005). One and a half century of diffusion: Fick, Einstein, before and beyond. *Diffusion Fundamentals*, 2(1):1–10.

Piazza, J. (2006). Rooted in poverty? terrorism, poor economic development, and social cleavages. *Political Violence*, 18:159–177.

Piazza, J. A. (2008). Incubators of terror: Do failed and failing states promote transnational terrorism? *International Studies Quarterly*, 52(3):469–488.

Picard, R. G. (1986). News coverage as the contagion of terrorism: Dangerous charges backed by dubious science. *Political Communication*, 3(4):385–400.

Piegorsch, W. W., Cutter, S. L., and Hardisty, F. (2007). Benchmark analysis for quantifying urban vulnerability to terrorrst incidents. *Risk Analysis*, 27:1411–1425.

Piironen, J. and Vehtari, A. (2017). Comparison of Bayesian predictive methods for model selection. *Statistics and Computing*, 27(3):711–735.

Pion-Berlin, D. and Lopez, G. A. (1991). Of victims and executioners: Argentine state terror, 1975-1979. *International Studies Quarterly*, 35(1):63–86.

Poggio, L., Gimona, A., Spezia, L., and Brewer, M. J. (2016). *Digital Soil Mapping Across Paradigms, Scales and Boundaries*, chapter Example of Bayesian Uncertainty for Digital Soil Mapping, pages 181–193. Springer Environmental Science and Engineering. Springer, Singapour.

Porter, M. D. and White, G. (2012). Self-exciting hurdle models for terrorist activity. *The Annals of Applied Statistics*, 6(1):106–124.

PSU (2014). Lesson 5.1: Decomposition models.STAT 510: Applied time series analysis. https://onlinecourses.science.psu.edu/stat510/node/69.

Raghavan, V., Galstyan, A., and Tartakovsky, A. G. (2013). Hidden Markov models for the activity profile of terrorist groups. *The Annals of Applied Statistics*, 7(4):2402–2430.

Raleigh, C. and Dowd, C. (2015). Armed Conflict Location and Event Data Project (ACLED) Codebook. http://www.acleddata.com/wp-content/uploads/2015/01/ACLED_Codebook_2015.pdf. [Accessed 21 November 2015].

Raleigh, C. and Hegre, H. (2009). Population size, concentration, and civil war. A geographically disaggregated analysis. *Political Geography*, 28(4):224–238.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010a). Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47(5):651–660.

Raleigh, C. and Urdal, H. (2007). Climate change, environmental degradation and armed conflict. *Political Geography*, 26(6):674–694.

Raleigh, C., Witmer, F., O'Loughlin, J., and Denemark, R. A. (2010b). A review and assessment of spatial analysis and conflict: The geography of war. *The international studies encyclopedia*, 10:6534–6553.

RAND (2011). Database scope. http://www.rand.org/nsrd/projects/terrorism-incidents/about/scope.html. [Accessed 12 June 2014].

Rapoport, D. C. (1996). Editorial: The media and terrorism: Implications of the Unabomber case. *Terrorism and Political Violence*, 8(1):7–9.

Ratcliffe, J. H. (2002). Aoristic signatures and the spatio-temporal analysis of high volume crime patterns. *Journal of Quantitative Criminology*, 18(1):23–43.

Ratcliffe, J. H. (2004). The hotspot matrix: A framework for the spatio-temporal targeting of crime reduction. *Police practice and research*, 5(1):5–23.

Ratcliffe, J. H. (2006). A temporal constraint theory to explain opportunity-based spatial offending patterns. *Journal of Research in Crime and Delinquency*, 43(3):261–291.

Ratcliffe, J. H. and Rengert, G. F. (2008). Near-repeat patterns in Philadelphia shootings. *Security Journal*, 21(1):58–76.

Rey, S. J., Mack, E. A., and Koschinsky, J. (2012). Exploratory space-time analysis of burglary patterns. *Journal of Quantitative Criminology*, 28(3):509–531.

Rezendes, S. and O'Sullivan, P. (1986). Terrain and the effectiveness of guerrillas. In *Papers & Proceedings of Applied Geography Conferences*, volume 9, pages 234–239.

Richardson, L. (2006). *What Terrorists Want*. John Murray, London, UK, 1st edition.

Richardson, R., Kottas, A., and Sansó, B. (2015). Flexible Integro-Difference Equation Modeling for Spatio-Temporal Data. Technical report, Tech. Rep. UCSC-SOE-14-10, Department of Applied Mathematics and Statistics, University of California, Santa Cruz.

Ridout, M., Demétrio, C. G., and Hinde, J. (1998). Models for count data with many zeros. In *Proceedings of the XIXth International Biometric Conference*, volume 19, pages 179–192.

Rodrigues, A., Diggle, P., and Assuncao, R. (2010). Semiparametric approach to point source modelling in epidemiology and criminology. *Journal of the Royal Statistical Society - Series C Applied Statistics*, 59:533–542.

Rogerson, P. and Sun, Y. (2001). Spatial monitoring of geographic patterns: an application to crime analysis. *Computers, Environment and Urban Systems*, 25(6):539–556.

Rogerson, P. A. (2015). Maximum Getis-Ord statistic adjusted for spatially autocorrelated data. *Geographical Analysis*, 47(1):20–33.

Roncek, D. (2000). *Schools and Crime*, pages 153–165. Thousand Oaks: Sage Publications, London, UK.

Ross, J. I. (1993). Structural causes of oppositional political terrorism: Towards a causal model. *Journal of Peace Research*, 30:317–329.

Ross, M. H. and Homer, E. (1976). Galton's problem in cross-national research. *World Politics*, 29(1):1–28.

Rue, H. and Held, L. (2005). *Gaussian Markov random fields: theory and applications*. CRC Press.

Rue, H. and Held, L. (2010). *Handbook of spatial statistics*, chapter Discrete Spatial Variation, pages 171–200. CRC press.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the royal statistical society: Series b (statistical methodology)*, 71(2):319–392.

Rue, H. and Tjelmeland, H. (2002). Fitting Gaussian Markov random fields to Gaussian fields. *Scandinavian Journal of Statistics*, 29(1):31–49.

Salehyan, I. (2006). *Rebels Without Borders: States Boundaries, Transnational Opposition, and Civil Conflict*. PhD thesis, University of San Diego, USA.

Salehyan, I. (2015). Best practices in the collection of conflict data. *Journal of Peace Research*, 52(1):105–109.

Sambanis, N. (2008). Terrorism and civil war. In Keefer, P. and Loayza, N., editors, *Terrorism, Economic Development, and Political Openness*, pages 174–206. Cambridge University Press, New York, NY, USA.

Sánchez-Cuenca, I. and De la Calle, L. (2009). Domestic terrorism: The hidden side of political violence. *Annual Review of Political Science*, 12:31–49.

Sandler, T. (2013). The analytical study of terrorism: Taking stock. *Journal of Peace Research*, 51:1–15.

Sandler, T. and Arce, D. G. (2003). Terrorism & game theory. *Simulation Gaming*, 34:319–337.

Sandler, T. and Enders, W. (2007). Applying analytical methods to study terrorism. *International Studies Perspectives*, 8(3):287–302.

Sandler, T. and Lapan, H. E. (1988). The calculus of dissent: An analysis of terrrorists' choice of targets. *Synthese*, 76:245–261.

Santifort, C., Sandler, T., and Brandt, P. T. (2013). Terrorist attack and target diversity: Changepoints and their drivers. *Journal of Peace Research*, 50:75–90.

Savitch, H. V. and Ardashev, G. (2001). Does terror have an urban future? *Urban Studies*, 38:2515–2533.

Schmid, A. P. (1992). Terrorism and democracy. *Terrorism and Political Violence*, 4(4):14–25.

Schmid, A. P. and Jongman, A. J. (1988). *Political Terrorism: A New Guide to Actors, Authors, Concepts, Data Bases, Theories, and Literature*. Transaction Books, New Brunswick, NJ, USA, 2nd edition.

Schrödinger, E. (1926). Quantisierung als eigenwertproblem. *Annalen der physik*, 385(13):437–490.

Schrodt, P. A., Yonamine, J., and Bagozzi, B. E. (2013). *Data-based Computational Approaches to Forecasting Political Violence*, pages 129–162. Springer Science+Business Media, New York, NY, USA.

Schutte, S. and Weidmann, N. B. (2011). Diffusion patterns of violence in civil wars. *Political Geography*, 30(3):143–152.

Shaw, C. R. and McKay, H. D. (1942). *Juvenile delinquency and urban areas: A study of rates of delinquents in relation to differential characteristics of local communities in American cities*. Sage Publications Inc, Chicago, Illinois, USA.

Sheehan, I. S. (2012). *Assessing and Comparing Data Sources for Terrorism Research*, pages 13–40. Springer, New York, NY, USA.

Sherman, L. W., Gartin, P. R., and Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27–56.

Siebeneck, L. K., Medina, R. M., Yamada, I., and Hepner, G. F. (2009). Spatial and temporal analyses of terrorist incidents in Iraq, 2004-2006. *Studies in Conflict & Terrorism*, 32:591–610.

Silberfein, M. (2003). Insurrections. In Cutter, S. l., Richardson, D. B., and Wilbanks, T. J., editors, *The Geographical Dimension of Terrorism*, chapter Insurrections, pages 67–73. Routledge.

Silke, A. (2004). *Research on Terrorism*, chapter An Introduction to Terrorism Research, pages 1–29. Frank Cass.

Simpson, D., Illian, J., Lindgren, F., Sørbye, S., and Rue, H. (2016). Going off grid: computationally efficient inference for log-Gaussian Cox processes. *Biometrika*, 103(1):49–70.

Sørbye, S. H. (2014). Introduction to INLA: Latent Gaussian models and simple examples in R-INLA. Spatial Modelling with INLA.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.

Stanton, A., Thart, A., Jain, A., Vyas, P., Chatterjee, A., and Shakarian, P. (2015). Mining for Causal Relationships: A Data-Driven Study of the Islamic State. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2137–2146. ACM.

START (2014). Data collection methodology. http://www.start.umd.edu/gtd/. [Accessed 12 June 2014].

Steen, J., Liesch, P. W., Knight, G., and Czinkota, M. R. (2006). The contagion of international terrorism and its effects on the firm in an interconnected world. *Public Money & Management*, 26:305–312.

Steenbeek, W. and Hipp, J. R. (2011). A longitudinal test of social disorganization theory: Feedback effects among cohesion, social control, and disorder. *Criminology*, 49(3):833–871.

Stein, M. L., Chen, J., Anitescu, M., et al. (2013). Stochastic approximation of score functions for Gaussian processes. *The Annals of Applied Statistics*, 7(2):1162–1191.

Stein, M. L., Chi, Z., and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(2):275–296.

Stephenson, J., Gallagher, K., and C, H. C. (2004). *Geological Prior Information: Informing Science and Engineering*, chapter Beyond kriging: dealing with discontinuous spatial data fields using adaptive prior information and Bayesian partition modelling, pages 183–206. Geological Society, London, UK.

Stevens, J. W. (2009). What is bayesian statistics. *What is*.

Subrahmanian, V. S., Mannes, A., Sliva, A., Shakarian, J., and Dickerson, J. P. (2013). *Computational Analysis of Terrorist Groups: Lashkar-e-Taiba*. Springer Science+Business Media, New York, NY, USA, 1st edition.

Suleman, M. T. (2012). Stock market reaction to terrorist attacks; empirical evidence from a front line state. *Australasian Accounting Business and Finance Journal*, 6:97–110.

Sutton, P. C. and Costanza, R. (2002). Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. *Ecological Economics*, 41(3):509–527.

Sutton, P. C., Elvidge, C. D., and Ghosh, T. (2007). Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *International Journal of Ecological Economics & Statistics*, 8(S07):5–21.

Swanstrom, T. (2002). Are fear and urbanism at war? *Urban Affairs Review*, 38(1):135–140.

Takeyh, R. and Gvosdev, N. (2002). Do terrorist networks need a home? *Washington Quarterly*, 25(3):97–108.

Tavares, J. (2004). The open society assesses its enemies: shocks, disasters and terrorist attacks. *Journal of monetary economics*, 51(5):1039–1070.

Taylor, H. M. and Karlin, S. (2014). *An Introduction to Stochastic Modeling*. Academic Press, 3rd edition edition.

The Council of the European Union (2001). Council Common Position. 2001/931/CFSP.

The Fund for Peace (2015). The Fragile States Index Data. http://fsi.fundforpeace.org/data. [Accessed 29 June 2015].

The Washington Post (2013). $52.6 billion: The black budget. http://www.washingtonpost.com/wp-srv/special/national/black-budget. [Accessed 09 February 2015].

Thuraisingham, B. (2004). *Data Mining for Counterterrorism*, pages 157–183. MIT Press, Menlo Park, CA, USA.

Tijms, H. C. (2003). *A first course in stochastic models*. John Wiley and Sons.

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

Tollefsen, A. F., Strand, H., and Buhaug, H. (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49(2):363–374.

Tuathail, G. O. (2009). Placing blame: Making sense of Beslan. *Political Geography*, 28(1):4 – 15.

Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, 33(1):1–67.

Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, 34(1):23–25.

van der Linde, A. (2005). DIC in variable selection. *Statistica Neerlandica*, 59(1):45–56.

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432.

Vehtari, A., Ojanen, J., et al. (2012). A survey of Bayesian predictive methods for model assessment, selection and comparison. *Statistics Surveys*, 6:142–228.

Vinyard Sotfware (2008). International terrorism data center. http://vinyardsoftware.com/howtoorder.html. [Accessed 12 June 2014].

Wallin, J. and Bolin, D. (2015). Geostatistical Modelling Using Non-Gaussian Matern Fields. *Scandinavian Journal of Statistics*, 42(3):872–890.

Walsh, J. B. (1981). A stochastic model of neural response. *Advances in applied probability*, pages 231–281.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *The Journal of Machine Learning Research*, 11:3571–3594.

Weidmann, N. B. (2013). The higher the better? the limits of analytical resolution in conflict event datasets. *Cooperation and Conflict*, 48(4):567–576.

Weidmann, N. B., Rød, J. K., and Cederman, L.-E. (2010). Representing ethnic groups in space: A new dataset. *Journal of Peace Research*.

Weimann, G. (2005). The theater of terror: The psychology of terrorism and the mass media. *Journal of aggression, maltreatment & trauma*, 9(3-4):379–390.

Weimann, G. (2008). The psychology of mass-mediated terrorism. *American Behavioral Scientist*, 52(1):69–86.

Weimann, G. (2012). Chapter 9: The role of the media in propagating terrorism. In Kumar, U. and Mandal, M. K., editors, *Countering Terrorism. Psychological Strategies*, chapter Chapter 9: The Role of The Media in Propagating Terrorism, pages 182–202. SAGE Publications India.

Weimann, G. and Brosius, H.-B. (1988). The predictability of international terrorism: A time-series analysis. *Terrorism*, 11:491–502.

Weisburd, D., Bushway, S., Lum, C., and Yang, S.-M. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of seattle. *Criminology*, 42(2):283–322.

Weisburd, D. and Green, L. (1995). Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly*, 12(4):711–735.

Welsh, B. C. and Farrington, D. P. (2004). Surveillance for crime prevention in public space: Results and policy choices in Britain and America. *Criminology & Public Policy*, 3(3):497–526.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, pages 434–449.

Whittle, P. (1963). Stochastic-processes in several dimensions. *Bulletin of the International Statistical Institute*, 40(2):974–994.

Wikipedia (2015). Hyperparameter. http://en.wikipedia.org/wiki/Hyperparameter. [Accessed 3 June 2015].

Wilkinson, P. (1979). *Terrorism and the Liberal State*. The MacMillan Press Ltd, Hong Kong, China, 2nd edition.

Wilkinson, P. (1990). *The Sources of Terrorism: Terrorists' Ideologies and Beliefs*, pages 139–145. St. Martin's Press, Inc, New York, NY, USA.

World Bank (2015). World Development Indicators. http://data.worldbank.org/data-catalog/world-development-indicators. [Accessed 6 July 2015].

Yang, W., Fotheringham, A. S., and Harris, P. (2012). An extension of geographically weighted regression with flexible bandwidths. In *Proceedings GISRUK 20th Annual Conference*.

Ye, X. and Wu, L. (2011). Analyzing the dynamics of homicide patterns in chicago: ESDA and spatial panel approaches. *Applied Geography*, 31(2):800–807.

Yip, C. Y. (2010). *Bayesian Spatio-Temporal Modelling for Forecasting Ground Level Ozone Concentration Levels*. PhD thesis, University of Southampton, UK.

Zammit Mangion, A. (2011). *Modelling from spatiotemporal data: a dynamic systems approach*. PhD thesis, University of Sheffield, UK.

Zammit-Mangion, A., Dewar, M., Kadirkamanathan, V., Flesken, A., and Sanguinetti, G. (2013). *Modeling Conflict Dynamics with Spatio-temporal Data*. SpringerBriefs in Applied Sciences and Technology, Heidelberg, Germany, 1st edition.

Zammit-Mangion, A., Dewar, M., Kadirkamanathand, V., and Sanguinetti, G. (2012). Point Process Modelling of the Afghan war diary. *PNAS*, 109:12414–12419.

Zammit-Mangion, A., Rougier, J., Bamber, J., and Schön, N. (2014). Resolving the Antarctic contribution to sea-level rise: a hierarchical modelling framework. *Environmetrics*, 25(4):245–264.

Zartman, I. W. (1995). Introduction: Posing the Problem of State Collapse. In *Collapsed states: the disintegration and restoration of legitimate authority*, pages 1–13. Lynne Rienner Publishers.

Zhang, R., Czado, C., and Sigloch, K. (2015). Bayesian spatial modelling for high dimensional seismic inverse problems. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*.

Zhukov, Y. M. (2010). Applied spatial statistics in R, section 2. IQSS Harvard University Workshop.

Zuur, A. F., Ieno, E. N., and Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1):3–14.