Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

IDENTIFICATION OF LATENT STRUCTURE IN SPATIO-TEMPORAL MODELS OF VIOLENCE

Nicholas J Clark

Department of Mathematical Sciences
United States Military Academy
Thayer Hall
West Point, NY 10996

Krista Watts

Department of Mathematical Sciences
United States Military Academy
Thayer Hall
West Point, NY 10996

ABSTRACT

The modeling of violence, including terrorist activity, over space and time is often done using one of two broad classes of statistical models. Typically, the location of an event is modeled as a spatio-temporal point process and the latent structure is either modeled through a latent Gaussian process motivated by a log-Gaussian Cox process or through data dependency similar to a Hawkes process. The former is characterized through dependence in an unobserved latent Gaussian process, while the later assumes a data driven dependence in the data. While both techniques have been used successfully it remains unclear whether the processes are practically different from one another. In this manuscript, we demonstrate that in many situations, the most common statistic to characterize clustering in a process, Ripley's K function, cannot differentiate between the two processes and should not be used.

1 INTRODUCTION

The statistical modeling of crime, or violence, serves to both predict where a violent event will occur as well as explain what factors contribute to an area having a high propensity towards violence. The best practices for creating a statistical model, however, are not entirely straightforward. For example, in 2017, the National Institute of Justice offered a competition to predict where crimes would take place (Hunt 2019). This competition was won by two teams who took, on the surface, very different approaches. In Mohler and Porter 2018, the authors used a marked Hawekes process to generate a predictive model whereas in Flaxman, Chirico, Pereira, Loeffler, et al. 2019, the team employed a latent Gaussian process with a spatially structured covariance. However, both teams were very successful in their predictions.

In the military this problem occurs as well. For example, in modeling the spread of terrorism, Clark and Dixon 2018 demonstrated that temporal Hawkes processes could be combined with spatial latent Gaussian structures and that the choice of the latent yielded significantly different results when analyzing Iraq data. Whereas in Python, Illian, Jones-Todd, and Blangiardo 2019 the authors placed all spatio-temporal structure in a latent Gaussian process when analyzing terrorism on across the globe.

While these different models have been used in both fields, it remains unclear if the data patterns generated from the two approaches are practically different from each other. While Mohler et al. 2013 demonstrated that a temporal log Gaussian Cox process could be decoupled from a temporal Hawkes process, it remains unclear whether the same principles apply to spatio-temporal models. Further, while the data may be uncoupled, it remains unclear whether the patterns that result from the different data generating processes are practically different from one another. Though the above discussion appears academic on the surface, it has practical implications for military intelligence and police analysts. There are assumptions about the data-generating mechanism inherent in the choice of the model. When justifying a Hawkes process through a first principle approach as in Short, D'orsogna, Pasour, Tita, Brantingham, Bertozzi, and

Chayes 2008 for crime or Clark and Dixon 2018 for terrorism, the results may drive practitioners to make policy decisions based on conclusions that are only valid if the actual mechanism is reasonably modeled in the first place. Therefore, fitting of these models to data may drive practitioners to make conclusions on repeat-victimization that lead to policy decisions (Farrell 1995). Repeat-victimization is the theory that there are geographical locations or individuals that get repeatedly victimized more often than would be expected (Farrell 1995). Repeat-victimization is often modeled through the use of a self-exciting statistical model, or a statistical model where the expected number of events is a function of the previous number of events (Mohler, Short, Brantingham, Schoenberg, and Tita 2011). As demonstrated in Clark and Dixon 2018, the choice of different latent Gaussian structure may impact the parameter that models self-excitation, so rather than repeat-victimization, the model may, in fact, be misspecified leading to incorrect conclusions about repeat victimization. This is further demonstrated in Reinhart and Greenhouse 2018 where it was shown that omitting covariates from the background rate also biased the self-excitation parameters.

In this manuscript, we give examples of both self-exciting processes and latent Gaussian driven processes that could be used to model the spatio-temporal spread of violence. We further simulate data over a range of parameters for both models and demonstrate that the most common spatial clustering statistic, Ripley's K function, should not be used to differentiate the two processes. While it remains an open question how to do so, our current recommendation is to not use data patterns to select the modeling approach but rather rely on first principles from fields such as criminology or military intelligence .

2 Background

Throughout this manuscript, formally we assume that Y is a Poisson process defined on a space-time set $s_i \times t$ with an intensity function for all bounded $B \in S$ given by $\lambda(s_i,t) = \int_{s_i \times t} \rho(\zeta) d\zeta$. Informally, we can think about terrorism or criminal events arising at higher rates in regions of high intensity and lower rates in regions of lower intensity. The primary challenge is how to structure the intensity field.

In Short, D'orsogna, Pasour, Tita, Brantingham, Bertozzi, and Chayes 2008, a collection of criminologists, mathematicians, and statisticians formulated a statistical model for crime that matched current crimonological beliefs on how crime moves in space and time. In their formulation, the number of burglaries observed at a given spatio-temporal location was a function of both near-repeat victimization as well as broken windows effect. This was achieved through allowing the expected number of events to be a function of the expected events at neighboring locations as well as self-excitation, or a data driven process. In Mohler, Short, Brantingham, Schoenberg, and Tita 2011 the authors demonstrated how the formulation used in Short, D'orsogna, Pasour, Tita, Brantingham, Bertozzi, and Chayes 2008 leads to a Hawkes point process model. The subsequent manuscript has been cited over 700 times as of April 2021 and has been used in many applications.

2.1 Hawkes Point Process

An approach to modeling $\lambda(s_i,t)$ that is directly motivated from ideas of self-excitation and spatial spread is to use a self-exciting process similar to a Hawkes Process. As in Mohler, Short, Brantingham, Schoenberg, and Tita 2011 the intensity field could be modeled as

$$\lambda(s_i,t) = \mu(s_i,t) + \sum_{k:t_k<t} g(t-t_k, |s_i - s_k|). \tag{1}$$

Here, the intensity is a function of two components, a background process, $\mu(s_i,t)$ that describes the background intensity and $g(.)$, the triggering function, which describes how much an observed event at location (t_k, s_k) increases the intensity at (s_i,t). While the seismology literature has proposed a plethora of choices for $g(.)$ (Ogata 1998), the criminology and terrorism literature has mostly used the exponential kernel in time and the Gaussian kernel in space(Mohler et al. 2013), (Clark and Dixon 2018), (Mohler 2014),(Reinhart and Greenhouse 2018), which is written as

$$g(t-t_k,|s_i-s_k|)=\lambda\exp(-\lambda(t-t_k))\frac{1}{2\sigma^2}\exp(-||s_i-s_j||\frac{1}{2\sigma^2}). \quad (2)$$

Reinhart and Greenhouse 2018 demonstrated that proper parameterization of the background rate is essential to inference and should be structured as $\mu=\exp(X^T\beta)$.

The significance of the modeling choice is that the parameters within the Hawkes process are often given real-world meaning. For instance, in Short, D'orsogna, Pasour, Tita, Brantingham, Bertozzi, and Chayes 2008, the amount of temporal self-excitement is referred to as repeat victimization, and the amount of spatial self-excitement is referred to as near-repeat victimization and the broken windows effect.

Practically, this has policy implications as the broken windows effect has led to changes in policing in many major cities (Harcourt and Ludwig 2006). While policies based on the broken windows theory of deterrence have been criticized (Harcourt 1998), the concepts underlying this phenomenon remain a subject of research in not only criminology and the study of terrorism, but also tourism, (Liu, Wu, and Che 2019) epidemiology (O'Brien, Farrell, and Welsh 2019) and international corruption (Alford 2012).

2.2 Log-Gaussian Cox Process

An alternate modeling technique that is proposed in Cressie and Wikle 2015 is to place spatio-temporal structure in an unobserved latent Gaussian process. The analogue to this in continuous space and time is a log-Gaussian Cox Process (LGCP). An LGCP is often refered to as a double stochastic process (Moller and Waagepetersen 2003) and is structured through assuming in a given space time region, A, the number of observed events follows

$$Z(A)\sim \text{Po}\left(\int_A \lambda(s_i,t)d(s_i)dt\right). \quad (3)$$

Further structure is then placed on the log of the intensity field

$$\log(\lambda)\sim MVN(\mu,\Sigma(\theta)). \quad (4)$$

Spatio-temporal structure, then, can either come from structure placed on μ or through specifying a covariance (or precision) matrix that captures the relationship between spatio-temporal locations.

For a spatial only model (no temporal aspect), log-Gaussian Cox processes aggregated over spatial regions can be well approximated by the geostatistical model

$$Z(s_i)\sim \text{Po}(\lambda(s_i)) \quad (5)$$

and $\Sigma(\theta)$ has a Matern class covariance function (Cressie and Wikle 2015). The Matern class covariance function assumes that the covariance between locations s_i and s_j is

$$c(s_i,s_j)=\frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(\kappa|s_i-s_j|)^\nu K_\nu(\kappa|s_i-s_j|), \quad (6)$$

where K_ν is the modified Bessel function of the second kind, κ is a scaling parameter, σ^2 is the marginal variance and ν is typically fixed due to poor identifiability (Guttorp and Gneiting 2006). For simplicity we assume $\nu=1$ throughout this manuscript.

While the Matern covariance function is extremely flexible and relates to the log-Gaussian Cox process, in practice (6) is a bit unwieldy and is computationally difficult to fit. Therefore, we can use the result of Lindgren, Rue, and Lindström 2011 which states that a Gaussian field with a Matern covariance matrix can be well approximated by a Gaussian Markov random field.

In order to extend this to a spatio-temporal model, we can form a covariance matrix through the kronecker product of the spatio covariance matrix with an auto-regressive (1) or AR(1) covariance matrix, which has the form

$$\Sigma_{s,t}^{-1} = \Sigma_t(\theta)^{-1} \otimes \Sigma_s(\theta)^{-1}. \tag{7}$$

As both Σ_t^{-1} and Σ_s^{-1} are sparse, so is $\Sigma_{s,t}^{-1}$ allowing computationally inexpensive inference.

While the latent Gaussian process is less easily described than (1), it is immensely flexible and has also successfully been used to model the spatio-temporal spread of violence (Shirota and Gelfand 2017; Rodrigues and Diggle 2012; Python, Illian, Jones-Todd, and Blangiardo 2019). In order to mimic the spread assumed in (2), we can assume that $\Sigma_t(\theta)^{-1}$ captures a Random Walk structure and $\Sigma_s(\theta)^{-1}$ is the precision matrix for a Matern covariance matrix.

2.3 Posterior Predictive Checks

If we are given a set of data, $\mathbf{y}$, and we fit it to a model $f(\mathbf{y}|\theta)$, one natural question to ask is, is the model appropriate for our data. One typical way to answer this question is with a χ^2 goodness of fit test. However, for complex models, such as models outside the generalized linear model family, it's nearly impossible to figure out the asymptotic distribution of this test statistic (Berkhof, Van Mechelen, and Hoijtink 2000). Furthermore, in the analysis of spatial data or spatio-temporal data, our question of interest usually is not whether the model is appropriate, but rather are our spatial points uniformly distributed over space/time, and if not, can we quantify how the clusters are formed.

One way to figure out whether our statistical model is appropriate is to compute a posterior predictive check (Gelman, Meng, and Stern 1996). A posterior predictive check is done through comparing the distribution of a given statistic of our data, $T(y)$ with the same statistic calculated from the posterior distribution of data simulated from our fitted model, $T(y_{rep}|y)$, where y_{rep} is simulated via

$$p(y_{rep}|y) = \int f(y_{rep}|\theta)p(\theta|y)d\theta \tag{8}$$

A natural choice of $T(.)$ is Ripley's K function (Dixon 2014) which quantifies the expected number of additional points within a set distance from another point over what is expected. Here we explore Ripley's K function calculated using a border correction as implemented via the spatstat library in R (Baddeley and Turner 2005). In broad terms, Ripley's K captures the amount of extra events within a given distance of a chosen event. High values of Ripley's K mean there is substantial clustering in the data, whereas low values means the data are more dispersed than what would be expected. However, Ripley's K does not always uniquely define point processes and two processes that have different generating distributions may, in fact, have similar Ripley's K values (Baddeley and Silverman 1984). While we will not rigorously prove this is the case between Hawkes and LGCPs processes, and there may in fact be some differences, we will demonstrate below that over a wide range of σ values, a LGCP is able to replicate the Ripley's K value of the Hawkes process.

3 Simulation

In this simulation we assume that the underlying generative process is actually a Hawkes process. That is, we assume that the broken windows phenomena holds and that the intensity of newly observed events are directly caused by events in close spatial or temporal proximity. We assume that the background intensity, $\mu(s_i,t)$ is constant and the spatio-temporal excitation can be broken into two distinguishable components as in (2). We allow $\lambda = 1$ and range $\sigma \in \{0.3, 0.7, 1.5, 1.9, 2.3\}$. These values of σ were chosen to reflect what we felt were realistic dispersion on the 10×10 field. For each value of sigma, we generated 100 separate realizations of the Hawkes process.

For each selection of σ we simulate from a Hawkes process using the methodology discussed below, and then fit it to a Log-Gaussian Cox-Process as in Section 3.2. We then calculate the Ripley's K statistic using the original data, as well as Ripley's K using data simulated from the fitted LGCP model. For each σ value we then compute the posterior-predictive check. If the models are able to simulate data with the same spatio-temporal characteristics we would expect about half the time the Ripley's K statistic from the Hawkes process would be larger than the Ripley's K statistic from the LGCP. If the posterior-predictive check yields values close to zero or close to 1, then one of the models allows for more clustering than the other, which suggests that Ripley's K could be used to determine if the generative process was indeed driven by a Hawkes process or not.

This entire process is computationally expensive and was done on a 24-core (48 thread) AMD Epyc 7401P processor operating at 2.0 GHz (3.0 GHz boost) paired with 256 GB of ECC DDR4 memory and two Nvidia Titan RTX GPUs each with 24 GB of DDR4 memory. The simulations were done in parallel using the R software and functions from the doParallel library (Corporation and Weston 2019).

3.1 Simulating from Hawkes Process

We define the intensity field given in (1) with the background process, $\mu(s_i,t)$, constant and the triggering function, $g(t-t_k,|s_i-s_k|)$, the product of an exponential kernel in time and a Gaussian kernel in space as given in (2).

We simulate data from a self-exciting background process using the method described in (Reinhart 2018). First we generate events from the background process over a 10x10 field of arbitrary units and 15 time periods by simulating from a Poisson process with mean $m \cdot A \cdot T$, where m is given in (Reinhart 2018), A is the total area, 100, and T is the total time, 15. We normalized our triggering function so that $m = 1$.

For each parent event, i, in our background process, we generate $N^{(i)}$ offspring events, where $N^{(i)} \sim Po(m)$, with time and location generated from our triggering function and events truncated at our spatio-temporal boundaries. We continue this for 15 generations of offspring.

3.2 Inference from Log-Gaussian Cox Process

As mentioned above, a challenge with the LGCP is how to structure the spatio-temporal dependence in an efficient manner while retaining flexibility in the process to capture a wide range of spatial and temporal dependencies. By far the most flexible spatial structure that can be used is assuming the spatial data arise from a Matern covariance matrix, which has been shown to be robust to misspecification (Stein 1988).

To allow for this, we use the stochastic partial differential equation (SPDE) inferential methodology introduced in (Lindgren, Rue, and Lindström 2011). Here, the authors demonstrate that the Matern class covariance function is a solution to a specific SPDE which, in turn, can be approximated using a sparse Gaussian Markov Random Field. This is implemented using the Integrated Nested Laplace Approximation (INLA) package in R (Lindgren and Rue 2015).

INLA performs approximate Bayesian inference for Latent Gaussian Models. The critical assumption is that the observations are linked to Gaussian (latent) linear predictors through a known link function similar to a generalized linear model. Here we denote $\eta(s_i,t)$ as the linear predictor $\eta(s_i,t) = \mu(s_i,t) + f(s_i) + g(t) + \varepsilon(s_i,t)$. Where $f(s_i)$ and $g(t)$ denote the spatial and temporal random effects respectively. As in Martino and Riebler 2014, we let $x = (\eta s_i, t, \mu(s_i,t), f(s_i), g(t), \varepsilon(s_i,t)$ denote the full latent Gaussian field. The complete assumed model is given by

$$Y|x,\theta \sim \Pi f(z_i|\eta_i,\theta) \tag{9}$$

$$x|\theta \sim N(0,\Sigma_{s,t}^{-1}(\theta)) \tag{10}$$

$$\theta \sim \pi(\theta) \tag{11}$$

INLA, as opposed to MCMC style Bayesian inference, directly approximates the posterior through exploiting the relationship

$$\pi(\theta|Y) \propto \frac{\pi(y|x,\theta)\pi(y|\theta)\pi(\theta)}{\pi(x|\theta,y)} \tag{12}$$

To calculate this, INLA computes a Gaussian approximation for the posterior of the latent Gaussian state, $\pi(x|\theta, y)$, by matching the model and curvature at the mode.

In order to fit a LGCP using INLA it is necessary to determine how to structure $f(s_i)$ in this framework. To do this, we first project all simulated points from the Hawkes process and project them onto the plane in $\mathbb{R}^2$. We then use the inla.mesh.2d function from the r-INLA package to create a triangular mesh where each point sits on a vertex. Then, the SPDE that gives rise to the Matern covariance function is approximately solved by using the finite element method through the basis function representation on the mesh. The resulting approximation has a sparse precision matrix. We assume that the temporal structure of the LGCP arises from a Random Walk (1) model which also has a sparse precision matrix, therefore allowing the entire latent structure to be efficiently computed.

4 Results

In general, we found that as σ increased, the amount of clustering at a distance of 1 unit decreases, which is to be expected. When $\sigma = 0.9$, the Hawkes process yielded a Ripley's K value of approximately 3.8 on average, decreasing steadily as shown in Table 1. Here we give the mean Ripley's K value across the 100 simulations for several of the σ values that we considered.

Table 1: Here we give the mean Ripley's K value, calculated directly from 100 simulated data sets for several σ values at a distance of 1 unit.

$\sigma = 0.3$	$\sigma = 0.7$	$\sigma = 1.5$	$\sigma = 1.9$	$\sigma = 2.3$
3.75	3.74	3.72	3.69	3.64

As we estimated parameters from the LGCP using the data generated from the Hawkes process and then simulated data from the fit model, we saw that the LGCP process was, for the most part, able to replicate the Ripley's K statistics from the original data. As seen in Table 2, the posterior predictive checks yielded values that were not extreme.

Table 2: Posterior Predictive Check for Select σ values

$\sigma = 0.3$	$\sigma = 0.7$	$\sigma = 1.5$	$\sigma = 1.9$	$\sigma = 2.3$
0.63	0.61	0.67	0.72	0.57

As the posterior predictive checks yielded values that were larger than 0.50 we can conclude the LGCP produced values that were, on average, slightly higher than the Ripley's K value found from the Hawkes process. However, as this value is not near 1, and for the most part is not larger than 0.70, we can conclude that this is very slight and likely to not be detectable by users. Overall, throughout the range of σ values we struggled to notice any discernable differences in the two cluster processes.

5 Conclusions

Selecting an appropriate statistical model is often the first task a statistician or data scientist must perform in order to conduct inference or make predictions. Often it is tempting to rely on characteristics of the data in order to determine the functional form of the model. For instance, continuous data often leads practitioners to start with a linear regression model unless there is observed curvature in the data or truncation that is present. Unfortunately, though, in the modeling of spatio-temporal data there does not appear to be a nice clear rule or statistic to differentiate two common methodologies from one another. While second order statistics are often used in spatio or spatio-temporal models to characterize data, in this manuscript we demonstrate that they cannot be used to differentiate between Hawkes processes and log-Gaussian Cox Processes.

While this may feel less than satisfying for practitioners seeking advice on structuring their statistical model, we conclude that this is actually freeing for intelligence analysts or criminologists seeking to conduct statistical analysis. Instead of allowing the data to drive the model, practitioners should be encouraged to use first principles from their chosen discipline to create a model that is testable. For example, if, as in (Tench, Fry, and Gill 2016), an intelligence analyst hypothesizes that the emplacement of IEDs is often done in response to other violent attacks, it would make sense to use a self-exciting Hawkes process. To test this, then, inference should be made on the triggering function, $g(.)$, given in (1). However, caution should be used when assigning meaning to or basing policies off of estimated parameters based on fit of the model alone. As is true in causal inference, the assumptions necessary to infer a specific data-generating mechanism are not readily verified in the data alone.

While we only looked at the spatial version of Ripley's K, there has been some work done on spatio-temporal Ripley's K (Lynch and Moorcroft 2008) that may provide a method for differentiation, however the implimentation of Ripley's spatio-temporal K function remains a challenge (Wang, Gui, Wu, Peng, Wu, and Cui 2020). Overall, we assess that attempting to find a statistic to differentiate the two processes will remain fruitless.

REFERENCES

Alford, R. P. 2012. "A broken windows theory of international corruption". *Ohio St. LJ* 73:1253.

Baddeley, A., and B. W. Silverman. 1984. "A cautionary example on the use of second-order methods for analyzing point patterns". *Biometrics*:1089–1093.

Baddeley, A., and R. Turner. 2005. "spatstat: An R Package for Analyzing Spatial Point Patterns". *Journal of Statistical Software* 12(6):1–42.

Berkhof, J., I. Van Mechelen, and H. Hoijtink. 2000. "Posterior predictive checks: Principles and discussion". *Computational Statistics* 15(3):337–354.

Clark, N. J., and P. M. Dixon. 2018. "Modeling and estimation for self-exciting spatio-temporal models of terrorist activity". *The Annals of Applied Statistics* 12(1):633 – 653.

Corporation, M., and S. Weston. 2019. *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*. R package version 1.0.15.

Cressie, N., and C. K. Wikle. 2015. *Statistics for spatio-temporal data*. John Wiley & Sons.

Dixon, P. M. 2014. "Ripley's K function". *Wiley StatsRef: Statistics Reference Online*.

Farrell, G. 1995. "Preventing repeat victimization". *Crime and Justice* 19:469–534.

Flaxman, S., M. Chirico, P. Pereira, C. Loeffler et al. 2019. "Scalable high-resolution forecasting of sparse spatiotemporal events with kernel methods: a winning solution to the NIJ "real-time crime forecasting challenge"". *Annals of Applied Statistics* 13(4):2564–2585.

Gelman, A., X.-L. Meng, and H. Stern. 1996. "Posterior predictive assessment of model fitness via realized discrepancies". *Statistica sinica*:733–760.

Guttorp, P., and T. Gneiting. 2006. "Studies in the History of Probability and Statistics XLIX on the Matérn Correlation Family". *Biometrika* 93(4):989–995.

Harcourt, B. E. 1998. "Reflecting on the subject: A critique of the social influence conception of deterrence, the broken windows theory, and order-maintenance policing New York style". *Mich. L. Rev.* 97:291.

Harcourt, B. E., and J. Ludwig. 2006. "Broken windows: New evidence from New York City and a five-city social experiment". *U. Chi. L. Rev.* 73:271.

Hunt, J. 2019. "From Crime Mapping to Crime Forecasting: The Evolution of Place-Based Policing". *National Institute of Justice. Available online: https://www. ncjrs. gov/pdffiles1/nij/252036. pdf (accessed on 1 December 2019)*.

Lindgren, F., and H. Rue. 2015. "Bayesian Spatial Modelling with R-INLA". *Journal of Statistical Software* 63(19):1–25.

Lindgren, F., H. Rue, and J. Lindström. 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.

Lindgren, F., H. Rue, and J. Lindström. 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.

Liu, J., J. S. Wu, and T. Che. 2019. "Understanding perceived environment quality in affecting tourists' environmentally responsible behaviours: A broken windows theory perspective". *Tourism Management Perspectives* 31:236–244.

Lynch, H. J., and P. R. Moorcroft. 2008. "A spatiotemporal Ripley's K-function to analyze interactions between spruce budworm and fire in British Columbia, Canada". *Canadian Journal of Forest Research* 38(12):3112–3119.

Martino, S., and A. Riebler. 2014. "Integrated nested laplace approximations (inla)". *Wiley StatsRef: Statistics Reference Online*:1–19.

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. et al. 2013. "Modeling and estimation of multi-source clustering in crime and security data". *The Annals of Applied Statistics* 7(3):1525–1539.

Mohler, G., and M. D. Porter. 2018. "Rotational grid, PAI-maximizing crime forecasts". *Statistical Analysis and Data Mining: The ASA Data Science Journal* 11(5):227–236.

Mohler, G. O., M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. 2011. "Self-exciting point process modeling of crime". *Journal of the American Statistical Association* 106(493):100–108.

Moller, J., and R. P. Waagepetersen. 2003. *Statistical inference and simulation for spatial point processes*. CRC Press.

O'Brien, D. T., C. Farrell, and B. C. Welsh. 2019. "Broken (windows) theory: A meta-analysis of the evidence for the pathways from neighborhood disorder to resident health outcomes and behaviors". *Social Science & Medicine* 228:272–292.

Ogata, Y. 1998. "Space-time point-process models for earthquake occurrences". *Annals of the Institute of Statistical Mathematics* 50(2):379–402.

Python, A., J. B. Illian, C. M. Jones-Todd, and M. Blangiardo. 2019. "A Bayesian approach to modelling subnational spatial dynamics of worldwide non-state terrorism, 2010–2016". *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 182(1):323–344.

Reinhart, A. 2018. "A Review of Self-Exciting Spatio-Temporal Point Processes and Their Applications". *Statistical Science* 33(3):299 – 318.

Reinhart, A., and J. Greenhouse. 2018. "Self-exciting point processes with spatial covariates: modelling the dynamics of crime". *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 67(5):1305–1329.

Rodrigues, A., and P. J. Diggle. 2012. "Bayesian estimation and prediction for inhomogeneous spatiotemporal log-Gaussian Cox processes using low-rank models, with application to criminal surveillance". *Journal of the American Statistical Association* 107(497):93–101.

Shirota, S., and A. E. Gelfand. 2017. "Space and circular time log Gaussian Cox processes with application to crime event data". *The Annals of Applied Statistics*:481–503.

Short, M. B., M. R. D'orsogna, V. B. Pasour, G. E. Tita, P. J. Brantingham, A. L. Bertozzi, and L. B. Chayes. 2008. "A statistical model of criminal behavior". *Mathematical Models and Methods in Applied Sciences* 18(supp01):1249–1267.

Stein, M. L. 1988. "Asymptotically efficient prediction of a random field with a misspecified covariance function". *The Annals of Statistics*:55–63.

Tench, S., H. Fry, and P. Gill. 2016. "Spatio-temporal patterns of IED usage by the Provisional Irish Republican Army". *European Journal of Applied Mathematics* 27(3):377–402.

Wang, Y., Z. Gui, H. Wu, D. Peng, J. Wu, and Z. Cui. 2020. "Optimizing and accelerating space–time Ripley's K function based on Apache Spark for distributed spatiotemporal point pattern analysis". *Future Generation Computer Systems* 105:96–118.

AUTHOR BIOGRAPHIES

NICHOLAS CLARK is an assistant professor in the Department of Mathematical Sciences at the United States Military Academy. He earned his Ph.D. in Statistics from Iowa State University. His research interests include spatio-temporal modeling, Bayesian inference, and statistical computation. Particularly, he is focusing on determining how violence clusters in space or time. His email address is nicholas.clark@westpoint.edu.

KRISTA WATTS is an associate professor in the Department of Mathematical Sciences at the United States Military Academy. She earned his Ph.D. in Biostatistics from Harvard University. Her research interests include Bayesian inference and effect estimation in observational data. Particularly, she is focusing on the fields of obesity, nutrition and public health. Her email address is krista.watts@westpoint.edu.