Journal of Applied Science and Innovation Studies
ISSN 2998-3606
Vol. 2, No. 1, 2025
Published by STSL Press

Original Paper

Testing for causal clustering in point processes

Ian McGovern[1], P. Jeffrey Brantingham[2], and Frederic Schoenberg[1]

[1] Department of Statistics, University of California, Los Angeles, CA 90095-1554.

[2] Department of Anthropology, University of California, Los Angeles, CA 90095-1554.

Abstract

Discriminating causal clustering from inhomogeneity in point processes is of high interest for a variety of applications. We propose a simulation-based test based on the relative likelihood of Hawkes models, Poisson cluster models, and inhomogeneous Poisson models, and compare with the time reversal test of Cordi et al. (2017). Under general conditions, causal clustering can be distinguished from inhomogeneity with high accuracy using these tests, with the test proposed here exhibiting somewhat higher power in simulations. The methods are applied to crime data on reported shootings in Boston from 2015-2021, where strong evidence of retaliatory triggering of events is seen in certain areas.

1. Introduction

Discriminating causal clustering from inhomogeneity is one of the key problems in point process analysis. Indeed, Diggle (2014) describes this problem as one of the most important challenges currently facing point process analysts. Both causal clustering and inhomogeneity can lead to aggregation of points in certain locations, though the mechanisms for this aggregation have very different implications. Causal clustering, or triggering, refers to the situation where the random occurrence of a point causes other points to be more likely to occur in the near vicinity of space-time. Inhomogeneity refers to the case where, because of differences in the background environment, points are simply more likely to occur at certain locations of space-time than others. Discriminating between these two phenomena can be very difficult in practice.

For example, suppose one is analyzing catalogues of reported gang-related violent crimes, and many occurrences are present in one spatial-temporal area. Is this aggregation of points due to the socio-economic and geographical circumstances in the particular location, in which case the explanation is inhomogeneity? Or is the aggregation due at least partly to retaliation, where one such crime that just happens to occur in the region may spark several retaliations, each of which might yield further retaliations, and so on, in which case the explanation is triggering?

One approach suggested in Diggle (2014) is to observe the point process repeatedly and inspect whether the clustering of points appears to occur predominantly in the same spatial-temporal locations, in which case inhomogeneity is the dominant paradigm. However, very often in practice, obtaining repeatedly observed point process data is not feasible. Furthermore, it is possible for a process, such as a Poisson cluster process, to exhibit aggregation of points at certain random hot-spots, which may be different in each realization, yet the aggregation of points is still not causal in the sense of individual points triggering others as in a Hawkes process.

An approach taken in Park et al. (2021) is to attempt to fit a Hawkes model with both inhomogeneity and triggering, and where the spatially and temporally varying background rate is modeled as accurately as possible, for instance using kernel smoothing of previously occurring points as well as covariate information that may influence the background rate, such as demographic information on each census tract. The idea is that if the inhomogeneity is accurately modeled via the background rate, then any triggering estimated in the resulting Hawkes model may be attributed to causal clustering. While this idea is sensible, it can be difficult to assess whether the background rate indeed accurately models all the inhomogeneity in the spatial-temporal environment, and any inhomogeneity not adequately captured by the model will leak into the estimated triggering function and be incorrectly characterized as causal clustering.

An idea explored in Cordi et al. (2017) is to fit a Hawkes model to the data and another to the data with the times reversed. If the model fits significantly better to the forward time data, then this suggests the aggregation of points may indeed be causal, whereas if the model fits equally well with the times reversed, then the aggregation is most likely due to inhomogeneity. The idea is that typically in applications it would make no sense for points to trigger the occurrence of *prior* points, so the observation that a Hawkes model fits as well to the time-reversed data as it does to the forward-time data is incompatible with actual causal clustering but is consistent with the notion that the Hawkes model's background rate term is not accurately describing the inhomogeneity in the process and thus incorrectly classifying some of the inhomogeneity as triggering. While the idea of Cordi et al. (2017) is very clever, its extension to the case where the inhomogeneity may vary over space instead of (or in addition to) time is problematic. Also, in some cases, a process with truly causal clustering might have the property that a Hawkes model fits equally well with the times reversed.

Previous research on causality in point processes has focused on different elements of causal inference from the causal ideas that this paper explores. For instance, Xu et al. (2016) studied Granger causality within Hawkes processes. Here, we focus on the discrimination between causal clustering and inhomogeneity, in order to identify whether clustering identified by a fitted Hawkes model is real, or whether it may more aptly be explained merely as an artifact of an inhomogeneous environment.

The main idea explored here is to fit both a Hawkes model with causal clustering as well as non-causal models as similar as possible to the Hawkes model but without causal clustering, such as Poisson cluster and inhomogeneous Poisson models, and compare how these models fit. This allows one to quantify the degree of causal clustering in the data, and, if the Hawkes model fits significantly better than the alternative models without causal clustering, then this is potentially strong evidence that the causal triggering identified by the Hawkes model is real.

We consider formal hypothesis tests using the log-likelihood statistic to determine if a Hawkes model fits significantly better than Poisson cluster or inhomogeneous Poisson alternatives, and a similar test is performed using the time-reversal method. Both the time reversal test and the model comparison tests are applied to simulated spatial-temporal point processes to quantify the accuracy in identifying causal clustering models. Following this, these methods are applied to reported shooting data from Boston. The analysis for the simulated data show that accuracy in correctly identifying causal structures is generally very high under most conditions. The application to the reported Boston shootings data suggest that there is indeed causal triggering in certain locations, perhaps due to retaliatory crime activity.

2. Background on inhomogeoeus Poisson, Hawkes, and Cluster processes

A spatial-temporal point process N on $S = \mathrm{R}^2 \times R$ is a random collection of points such that for any bounded Borel subset B of S, the number of points that are within B is some finite number, which is denoted as $N(B)$ (van Lieshout, 2019; Daley and Vere-Jones, 2003). A point process is *simple* if with probability one, all its points are distinct. Such point processes can be defined by their conditional intensity,

$$\lambda(x, y, t|\mathrm{H}_t) = \lim_{\delta_x, \delta_y, \delta_t \to 0} \mathrm{E}[N((B_{x,y,\delta_{xy}} \times [t, t+\delta_t)|H_t)]/\delta_{xy}\delta_t ,$$

where $B_{x,y,\delta_{xy}}$ is a ball of area δ_{xy} around the location (x, y) and H_t is defined as the history of the process up to time t.

If N is a simple point process whose conditional intensity λ varies with x, y, and t but $\lambda(x, y, t)$ does not depend on what points have occurred previously, then N is an inhomogeneous Poisson process. Such processes embody the notion that aggregation of points is due to inhomogeneity only. In the context of crimes, an inhomogeneous Poisson model may allow the rate of points at any particular location and time to depend on the socio-economic features of the location in question, but would not incorporate retaliatory behavior in the model.

A Hawkes process is referred to as a "self-exciting" process, in that a point may trigger future points in its spatial-temporal vicinity. This type of model has often been used to describe clustered phenomena such as earthquakes and infectious diseases (Ogata, 1988; Reinhart, 2018; Meyer et al., 2012). According to the Hawkes model, parent points occur according to a background inhomogeneous Poisson process, $\mu(x, y, t)$. These parent points then produce offspring according to some triggering density h and some productivity value κ, the latter of which represents the expected number of points triggered by any given point. Once the parent points have produced offspring, those offspring trigger further offspring, and so on. The conditional intensity is thus given by

$$\lambda(x,y,t) = \mu(x,y,t)+\kappa \sum_{i:t_i<t} h(x-x_i,y-y_i,t-t_i) .$$

A Poisson cluster process is a clustering model defined in a two-part process. First, "parent" points are distributed throughout the spatial-temporal domain. Each parent point creates a random number with mean A of offspring points according to a specified triggering distribution, and the final process consists only of the offspring points (Neyman, 1939). Poisson cluster processes have been used to describe clustered spatial processes such as tree stands (Penttinen et al., 1992), rainfall (Guttorp, 1996), and galaxies (Snethlage et al., 2002), and typically the triggering density is symmetric so that offspring are distributed around their parents according to some isotropic density. Here, we consider the spatial-temporal context where the offspring points are distributed around their parents isotropically in space and time, meaning the parent points generate offspring occurring both before and after their parents. Thus the aggregation in such a Poisson cluster process is causal but is not physically sensible for applications where a point cannot trigger prior points, and we will be using such models not for their physical plausibility but purely for purposes of comparison with Hawkes processes. The conditional intensity of a Neyman-Scott process is difficult to write in condensed form (Møller and Waagepetersen, 2004; Zhuang, 2018), but can readily be estimated via maximum likelihood, minimum contrast, or other methods, despite occasional difficulties with convergence failure or numerical instability (Baddeley et al., 2022).

3. Methods

For a point process with conditional intensity $\lambda(x, y, t)$ and with points denoted as $\tau_1 = (x_1,y_1,t_1),...,\tau_n = (x_n,y_n,t_n)$, the likelihood can be expressed as

$$\prod_{1\le i\le n} \lambda(\tau_i) \exp\left[-\int_S \lambda(x,y,t)dxdydt\right] ,$$

where S is the spatial-temporal observation region. Therefore, estimating the likelihood becomes a process of calculating the intensity at each observed point, then calculating the integral of the conditional intensity over the observation region. In practice, calculation of the conditional intensities is quite straightforward, though approximation is often necessary to compute the integral of the conditional intensity (Harte, 2012).

For formal comparison of models, a hypothesis test method based on the expected information gain per trial is performed. The expected information gain per trial is a measure of the change in entropy scores from a null model and an alternate model (Daley and Vere-Jones, 2003). This information gain is a measure of the predictive performance of a model in terms of predicting the next occurring point within the point process, and is closely approximated by the mean log-likelihood ratio (Harte and Vere-Jones, 2005),

$$\hat{G}_N = \log (L_1 / L_0) / \mathrm{N}$$

where L_1 is the likelihood of the alternate model, L_0 is the likelihood for the null model, and N is the total number of points observed.

We consider a test with the following hypotheses.

H_0: The degree of causal clustering is 0.

H_a: The degree of causal clustering is greater than 0.

We propose considering the log likelihood ratio $\hat{G}_N$ of a fitted Hawkes model and a fitted Poisson cluster model as a test statistic, and refer to this in what follows simply as a likelihood ratio test. Suppose the significance level $\alpha = .05$. By design, if the data truly arise from a Poisson cluster model, then the test will reject the null hypothesis H_0 with probability 5%.

The idea, however, is that such a formal test may also be useful when the data may arise from an inhomogeneous Poisson model, since in such cases the Hawkes model would not be expected to fit significantly better than the non-causal Poisson cluster model and thus the test may be expected often to fail to reject the null hypothesis that the non-causal Poisson cluster process is the generating mechanism. In the next section, using simulations, we consider the case where the data are generated via a Hawkes process. In order to obtain an approximate sampling distribution for the information gain statistic under the null hypothesis, the following Monte Carlo technique is used.

For each test performed in what follows, a non-causal Poisson cluster model was fit by maximum likelihood estimation to to the corresponding data or simulated data, and then realizations of non-causal Poisson cluster models were simulated re-peatedly with parameters equal to these maximum likelihood estimates, to create a sampling distribution for the information gain statistic. For each simulation, the likelihood, L_1, for a Hawkes model, and the likelihood L_0, for a Poisson cluster model, were calculated, and used to calculate $\hat{G}_N$. This creates a sampling distribution for the value of the information gain statistic, and the value of $\hat{G}_N$ for the actual data is then compared to this sampling distribution. If the information gain statistic from the data is higher than the 95th percentile of the simulated sampling distribution, then the null hypothesis is rejected, and otherwise the null hypothesis is not rejected. For the time-reversal test this procedure was repeated, but with L_1 as the likelihood of the Hawkes model given the data and L_0 as the likelihood of the Hawkes model with the times reversed.

4. Simulations

Hawkes processes are simulated in order to determine the power of the test, i.e. the fraction of times the test correctly rejects the null hypothesis of a Poisson cluster process in favor of the Hawkes model. We use a two-dimensional Gaussian distribution for the spatial triggering density, a truncated Gaussian distribution with a lower bound of 0 for the temporal triggering density, and a constant background rate μ. The spatial region is $[0, 1] \times [0, 1]$ and the temporal region is $[0, 1]$.

For each simulation, three likelihood values were calculated: the likelihood for the standard Hawkes process model, the likelihood for a Poisson cluster model, and then all the times were reversed and the likelihood was found for the "backwards" or "reversed" Hawkes process.

A hypothesis test was performed on simulated point processes using the information gain statistic. By design, this test will fail to reject the null hypothesis with probability 95% when the simulated data come from a Poisson cluster model. When the simulated data come instead from an inhomogeneous Poisson model, the information gain test failed to reject the null hypothesis approximately 95% of the time as well. A variety of different Hawkes processes were simulated in order to investigate the power of the test. In particular, we investigated various different values for the background rate, μ, the productivity, κ, and the standard deviations, σ_t and σ_{xy}, of the temporal and spatial triggering densities, respectively. Each of these was allowed to vary. The range of tested values was chosen to be similar to the fitted Hawkes model for the application to crime data in Section 5.

The dependence of the power on the productivity, κ, is shown in Figure 1. The Gaussian clustering test has much higher power than the time reversal test for nearly all values of κ. The power for the time reversal method increases as κ increases, before again decreasing after a peak at $\kappa = 0.834$.

As shown in Figure 2, the Gaussian clustering test has very high power for all values of σ_t, though its power appears to decrease as the spatial triggering density gets more diffuse.

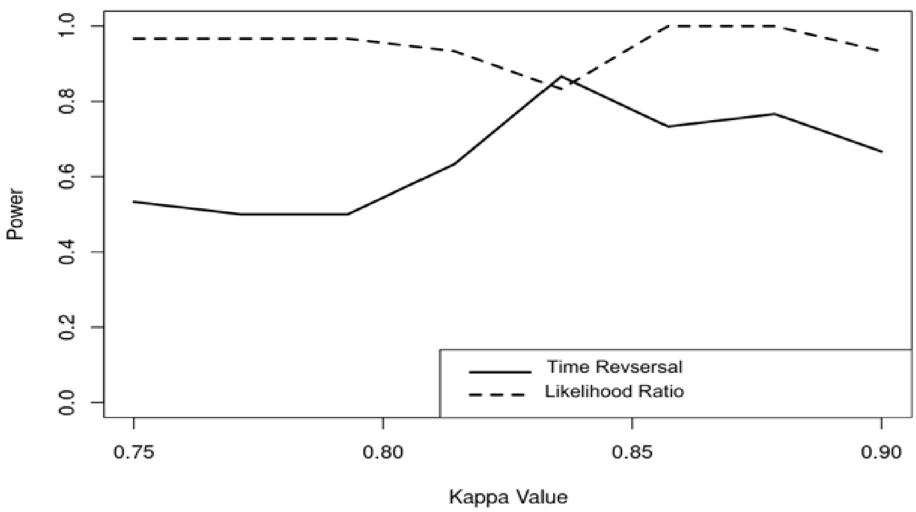

Figure 1: *Power of Poisson cluster and Time-Reversal tests as a function of κ, for 100 simulated Hawkes processes with triggering rate κ, each with $\mu = 18$, $\sigma_t = .0002$ and $\sigma_{xy} = .0002$, where for each simulated process, 100 Gaussian clustering processes were fit by MLE in order to obtain the sampling distribution*

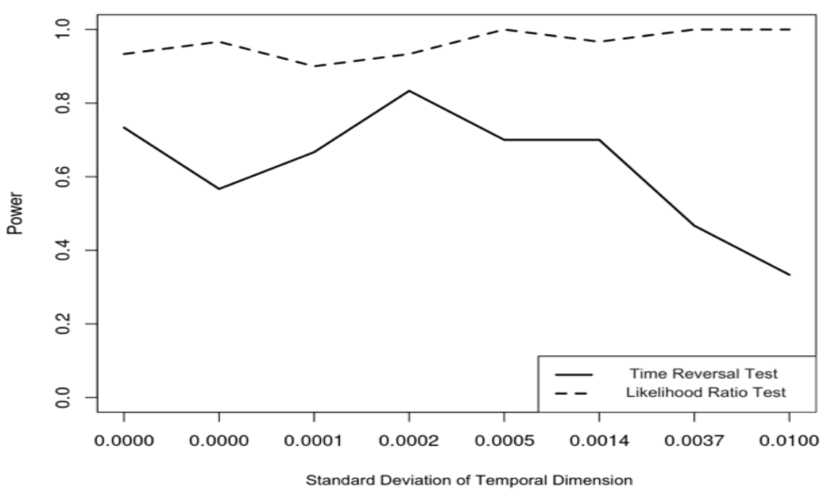

Figure 2: *Power of Poisson cluster and Time-Reversal tests as a function of σ_t, for 100 simulated Hawkes processes with temporal standard deviation σ_t, each with $\mu = 18$, $\kappa = .81$, and $\sigma_{xy} = .0002$, where for each simulated process, 100 Gaussian clustering processes were fit by MLE in order to obtain the sampling distribution.*

The Poisson cluster test appears to have very high power for all values of σ_t, while the Poisson cluster test has lower power for higher values of σ_{xy}. For the time reversal tests, increasing σ_t or σ_{xy} tends to reduce power, most likely because as the standard deviations of the triggering densities increase, it is increasingly difficult to discern any meaningful clusters, and the points simply appear nearly uniformly distributed within the observation region. This makes distinguishing between different types of clustering more difficult. As shown in Figure 4, the power of the tests considered here does not appear to change very significantly as μ varies.The two tests have nearly equal power over all values of μ, with the Poisson cluster test having slightly higher power overall. Overall, the power of the Poisson clustertest is 84.4% and the power of the reversal test is 74.4%.

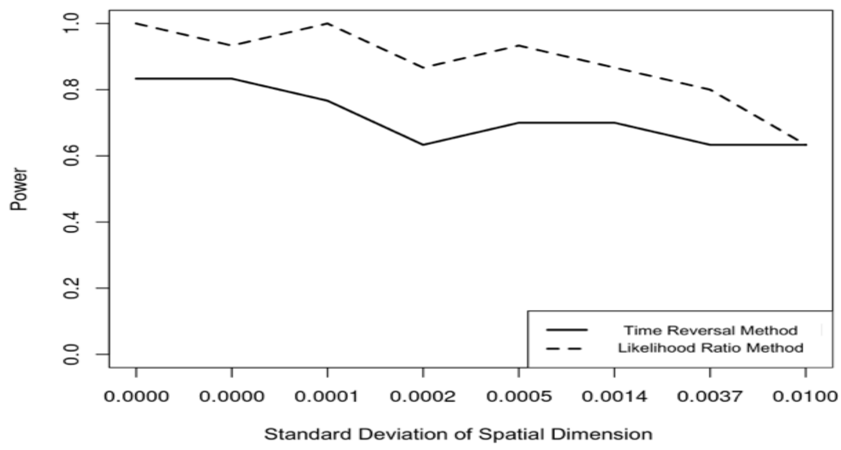

Figure 3: *Power of Poisson cluster and Time-Reversal tests as a function of σ_{xy}, for 100 simulated Hawkes processes with spatial standard deviation σ_{xy}, each with μ=18, κ=.81, and σ_t =.0002,wherefor each simulated process, 100 Gaussian clustering processes were fit by MLE in order to obtain the sampling distribution.*

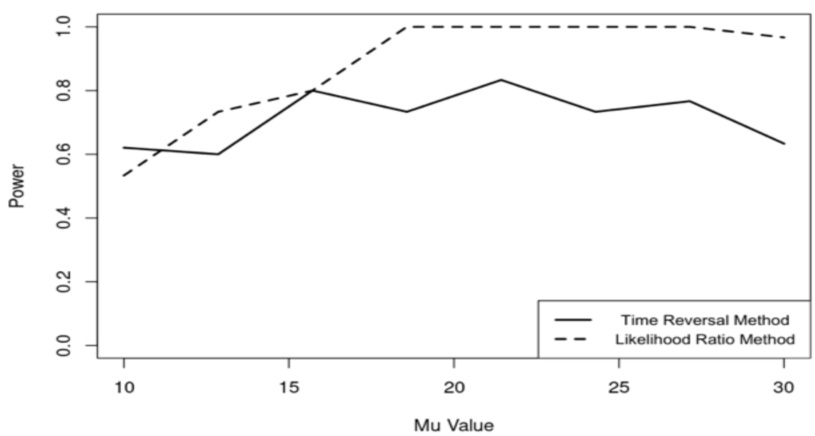

Figure 4: *Power of Poisson cluster and Time-Reversal tests as a function of μ, for 100 simulated Hawkes processes with background rate μ, each with κ = .81, σ_t = .0002 and σ_{xy} = .0002, where for each simulated process, 100 Gaussian clustering processes were fit by MLE in order to obtain the sampling distribution*

5. Application to Crime Data

5.1 Data

Recorded data on 8,862 reported illegal shootings in Boston between 2015 and 2021 were collected from the public data source for the Boston government (https://data.boston.gov/dataset/shootings). Figure 5 shows a kernel smoothing of the locations of these reported crimes.

The data were divided into uniform 10 × 10 grid cells, each analyzed individually using the tests described in Section 3. Grid cells including less than 5 points were excluded from the analysis as the tests have insufficient power in such cases.

5.2 Results

Figure 6 shows the results of the Poisson cluster hypothesis test. The results suggest that for the majority of locations in Boston, there is significant causal clustering present in the data on recorded shootings. Of the grid sections that were included within the analysis, 84.6% resulted in the test rejecting the null hypothesis, and these sections contained 83.7% of the total reported shootings. At the same time, there are several locations, especially on the Northwest borders of the data set, where the test fails to reject the null hypothesis and suggests that the local aggregation of points in these locations may be entirely due to inhomogeneity.

The time reversal test results, shown in Figure 7, have far more grid cells where the test fails to reject the null hypothesis. The time reversal test only rejected the null hypothesis in 44.2% of the grid cells, corresponding to a total of 46.8% of the reported shootings. The majority of grid cells where the Poisson cluster test failed to reject the null hypothesis also had the time reversal test fail to reject the null hypothesis, again suggesting inhomogeneity as the dominant cause of aggregation of points in these areas.

5.3 Analysis

The results of the Poisson cluster test indicate that, in the vast majority of locations within Boston, the reported shooting data from 2015-2021 are significantly better fit by a Hawkes model with causal clustering than by a Gaussian clustering model. Since the test fails to reject about 95% of the time when either a Gaussian clustering model or inhomogeneous Poisson model is the actual data generating mechanism, the results suggest that the clustering in these points is truly causal.

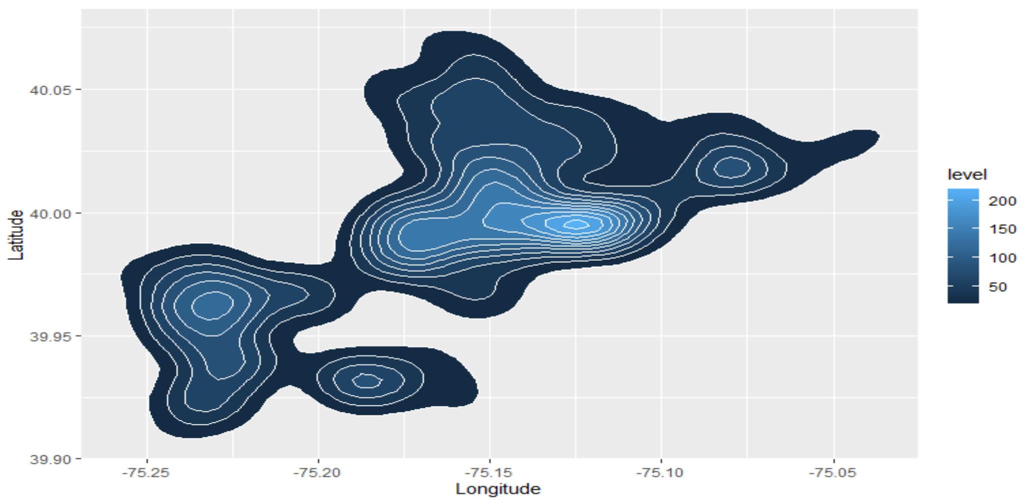

Figure 5: *Spatially smoothed density plot of Boston shooting data*

The results provide evidence that a Hawkes model with causal clustering may be appropriate for certain shooting data. However, there are still some areas, especially near the Northwestern borders of the observation region, where causal clustering is not indicated. This could possibly be due to spatially varying covariates, differences in gang territory, or other factors resulting in more causal clustering in certain locations rather than others.

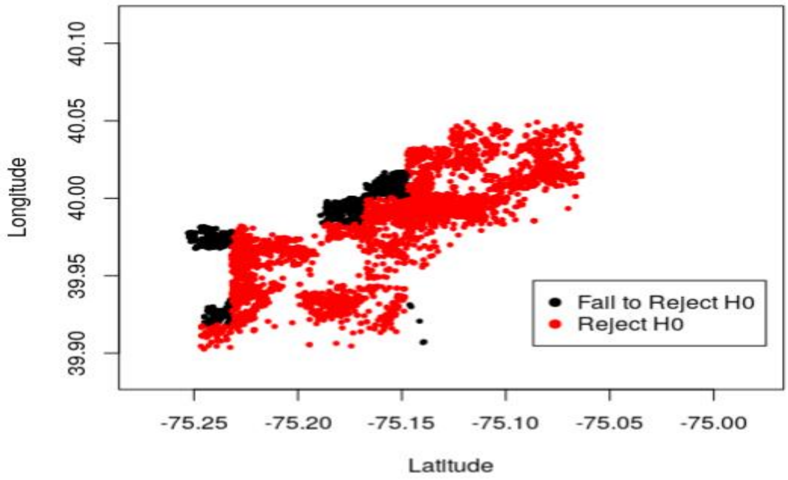

Figure 6: *Each dot represents a single shooting within a grid, and the data points within each grid section have been colored based on the result of the Poisson cluster hypothesis test within that grid section.*

The Poisson cluster and time reversal tests resulted in substantially different classifications. One possible explanation for this can be seen in the power analysis indicated by the simulations, since the Poisson cluster test had higher power than the time reversal test in most cases. Therefore, the reason that so many more sections failed to reject the null hypothesis using the time reversal test could be because the power of this test was too low.

6. Conclusion

Distinguishing between causal clustering and inhomogeneity in point processes is still a problem requiring much further study. Simulations show that under certain conditions, a simulated Hawkes model can be correctly distinguished from a Poisson cluster model using the information gain statistic, and furthermore, the test appears to have high power in distinguishing a Hawkes model from an inhomogeneous Poisson model as well. The time reversal test, by contrast, has somewhat lower power. This power is affected by the parameters of the simulation, with larger data sets and more intense clustering resulting in higher power for both tests.

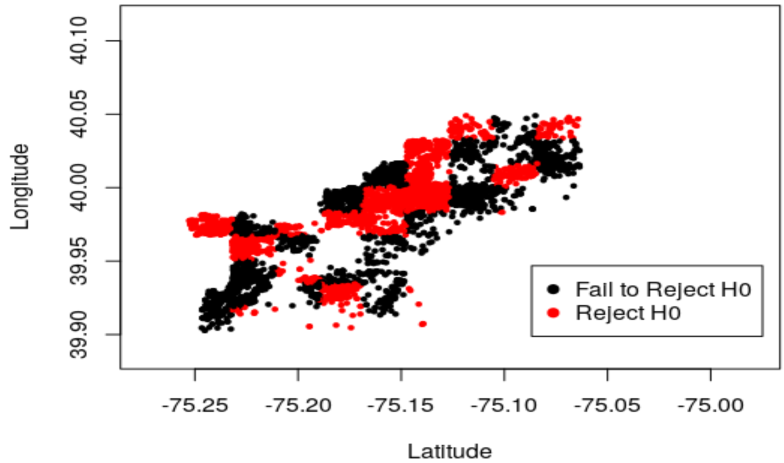

Figure 7: *Each dot represents a single shooting within a grid, and the data points within each grid section have been colored based on the result of the time reversal hypothesis test within that grid section.*

Hawkes models have been used extensively in crime data analysis, typically without much investigation into whether or not the assumption of causal clustering is indicated. Models without causal clustering, such as inhomogeneous Poisson models or Poisson cluster models, may fit just as well to the data in some situations. However, with regard to the application to the recorded shooting data in Boston, our results do suggest strong evidence of causal clustering in most areas of the city.

Future research should investigate this evidence of causal clustering further. Here, we considered Gaussian triggering functions for both the Poisson cluster and Hawkes model, but alternative triggering functions could be considered. In addition, we allowed each spatial grid cell to have its own background rate, to account for spatially varying covariates such as poverty levels or education levels. Future work could alternatively model the background crime rate more more explicitly as a function of such socio-economic covariates, as in Park et al. (2021). In addition, other types of reported crime data should be analyzed and the relationship between different types of crimes and the strength of evidence of causal clustering should be studied.

Acknowledgement

This material is based upon work supported by the National Science Foundation under grant number DMS-2124433.

References

Baddeley, A., Davies, T. M., Hazelton, M. L., Rakshit, S., and Turner, R. (2022). Fundamental problems in fitting spatial cluster process models. *Spatial Statistics*, 52:100709.

Cordi, M., Challet, D., and Toke, I.-M. (2017). Testing the causality of hawkes processes with time reversal. *Statistical Science*.

Daley, D. J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. I*. Probability and its Applications (New York). Springer-Verlag, New York, second edition. Elementary theory and methods.

Diggle, P. J. (2014). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. CRC Press.

Guttorp, P. (1996). Stochastic modeling of rainfall. In Wheeler, M. F., editor, *Environmental Studies*, pages 171–187, New York, NY. Springer New York.

Harte, D. and Vere-Jones, D. (2005). The entropy score and its uses in earthquake forecasting. *Pure Applied Geophysics*, 162.

Harte, D. S. (2012). Bias in fitting the ETAS model: a case study based on New Zealand seismicity. *Geophysical Journal International*, 192(1):390–412.

Junhyung Park, Frederic Paik Schoenberg, A. L. B. and Brantingham, P. J. (2021). Investigating clustering and violence interruption in gang-related violent crime data using spatial–temporal point processes with covariates. *Journal of the American Statistical Association*.

Meyer, S., Elias, J., and Höhle, M. (2012). A space– time conditional intensity model for invasive meningococcal disease occurrence. *Biometrics*, 68(2):607–616.

Møller, J. and Waagepetersen, R. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC.

Neyman, J. (1939). On a new class of contagious distributions, applicable in entomology and bacteriology. *The Annals of Mathematical Statistics*, 10(1).

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401):9–27.

Penttinen, A., Stoyan, D., and Henttonen, H. M. (1992). Marked Point Processes in Forest Statistics. *Forest Science*, 38(4):806–824.

Reinhart, A, (2018). A review of self-exciting spatio-temporal point processes and their applications. *Statistical Science*, 33(3):299–318.

Snethlage, M., Martínez, V., Stoyan, D., and Saar, E. (2002). Point field models for the galaxy point pattern : Modelling the singularity of the two-point correlation function.

Snethlage, M.; Martinez Garcia, Vicente; Stoyan, D.; Saar, E., 2002, Point field models for the galaxy point pattern : Modelling the singularity of the two-point correlation function, Astronomy and Astrophysics, vol. 388, no. 3, p. 758-765, 388.

van Lieshout, M. (2019). *Theory of Spatial Statistics: A Concise Introduction*. CRC Press.

Xu, H., Farajtabar, M., and Zha, H. (2016). Learning granger causality for hawkes processes.

Zhuang, J. (2018). Likelihood-based detection of cluster centers for neyman–scott point processes. *Journal of Environmental Statistics*, 8(3).