Model comparison for aftershock sequences following the 2005 Kashmir disaster

K. Türkyilmaz †

M.N.M. van Lieshout †

A. Stein ‡

†CWI, Amsterdam, The Netherlands.

‡Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands.

Abstract
In an earlier study (Van Lieshout and Stein, in press) we postulated the existence of two major earthquakes in Kashmir instead of a single one, based upon the pattern of aftershocks. In this note we explore this pattern further by fitting several spatial point pattern models. In particular we discuss the Hawkes and the trigger process models for earthquake aftershock sequences following the Kashmir catastrophe in 2005. The minimum contrast method is used for estimation of the parameters. The study shows that the trigger model fits better than the Hawkes model. The most likely number of main shocks is rounded to 2 generating the almost 200 aftershocks, whereas the Hawkes model would estimate a parent process of approximately 18 parents with on average 10 descendants. We conclude that the spatial pattern of aftershocks can best be understood as being generated by the trigger point process model.

Keywords & Phrases: Hawkes process, minimum contrast method, nearest-neighbour distance distribution function, pair correlation function, trigger process.

2010 Mathematics Subject Classification: 60D05, 60G55, 62M30.

1 Introduction

Due to the subduction of the Indo-Australian continental plate under the Eurasian plate, Pakistan is vulnerable to earthquakes. The risk varies across the country [11] with one hot spot in Pakistan-administered Kashmir, a region at which the two convergence zones associated with the subduction meet. A catastrophic earthquake occurred in October 2005 resulting in at least $86,000$ casualties. The Pakistan Meteorological Department estimated its magnitude by 5.2 on the Richter scale; the United States Geological Survey (USGS) measured at least 7.6 on the moment magnitude scale. The epicentre was at some 19 km north-east of Muzaffarabad with a hypocentre at 26 km below the surface.

The cluster of aftershocks following the main event on October 8 was studied in [2] who concluded that the seismicity decreased sharply and the number of earthquakes that occurred more than a month after the first shock was negligible. This picture was confirmed by the analysis carried out in [11] where it was noted that the spatial pattern formed by the earthquakes in the period October 8–November 7, 2005, could be well described by two clusters: one corresponding to the main earthquake, the other to the next largest earthquake with a magnitude equal to 6.4 occurring approximately seven hours later.

The goal of the present paper is to fit the two models most commonly used in seismology to the Kashmir data and to compare the results.

2 Data and models

The data consist of the locations and times of shallow earthquakes at a depth of less than 70 km of magnitude 4.5 or higher that happened between October 8, 2005, and November 7, 2005 in the Kashmir area. See [2] or [11] for further details of these data in a historical context. Figure 1 shows the pattern of the 176 spatial locations in the rectangular window $W = [72.65, 74.25] \times [33.70, 35.25]$; Figure 2 plots the frequency of earthquake occurrences against time (in hours counted from the 8th of October). The sharp decline in seismic activity is clearly visible; indeed 76 percent of shocks happen within the first 48 hours from the main shock. From now on, therefore, it is assumed that Figure 1 is an exhaustive map of the aftershocks with estimated spatial intensity $\hat{\lambda} = 176\ |W| = 70.97$ per square degree.

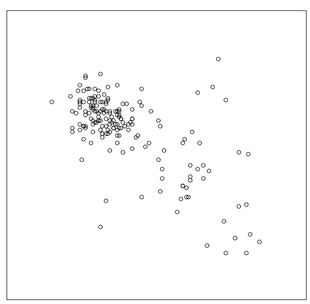

Figure 1: Locations of shallow earthquakes of magnitude 4.5 or higher recorded during the period October 8–November 7, 2005.

The natural model to describe data of the form described above is the Poisson cluster process X in the plane defined as the union of independent, identically distributed clusters Z_x centred around the points x of a stationary planar (marked) Poisson process Φ of 'parents',

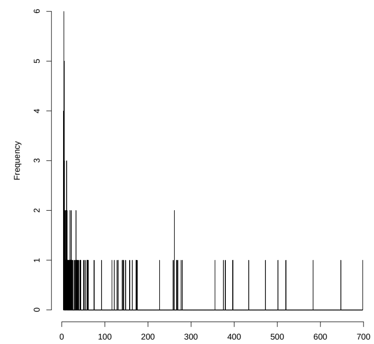

Figure 2: Frequency of occurrences of shallow earthquakes of magnitude 4.5 or higher recorded during the period October 8–November 7, 2005, plotted against time (in hours).

that is,

$$X = \bigcup_{x \in \Phi} (x + Z_x).$$

If the Z_x consist of a random number of (marked) points that are scattered independently with identical distribution around the parent x, X is a Neyman–Scott process and, under the further assumption that the cluster size distribution is Poisson, X is known as a trigger process in the seismological literature [1, 12, 18, 19].

Trigger processes are convenient to work with. Indeed, closed form expressions exist for many fundamental point process characteristics such as the intensity, the generating functional, Palm distribution, J-function and pair correlation function [6, 7, 9, 17]. This is not the case for the other sub-class of Poisson cluster processes that dominates the seismology literature, namely that of Hawkes processes.

In Hawkes processes, each parent independently generates a (marked) Poisson process of offspring with an intensity function that depends on the parent. The offspring in their turn also generate offspring independently of all others, and so on. Thus, Z_x is a branching process. An example is Ogata's Epidemic Type Aftershock-Sequences (*ETAS*) model which is now the standard first approximation for seismic catalogue data that come in the form of a list of earthquake locations marked by time. An excellent review including historical references is [14]. The temporal component is important in that a conditional intensity can be written down. Consequently, a likelihood function is available in closed form [14]. In the spatial case, on the other hand, only a series representation of the pair correlation function is available [13] and it seems to be hard to generalise this expression to marked point processes except for predictable marks [5].

3 Fitting a trigger process

The first model we consider is a trigger process X in the plane. The parent process Φ is assumed to be a stationary Poisson process of intensity $\kappa > 0$, which, in view of the small region under consideration, is a viable assumption to make. The number of offspring generated by each parent is taken to be Poisson with parameter $\nu > 0$. Conditional on the number of offspring, they are independently and identically distributed according to a bivariate normal distribution centred at the parent location with covariance matrix $\sigma^2 I$ where I is the 2×2 identity matrix. This trigger model is known in stochastic geometry as the modified Thomas process [17].

X is stationary with intensity $\lambda = \kappa\nu$ and has second order product density

$$\rho(x, y) = \lambda^2 + \frac{\kappa\nu^2}{4\pi\sigma^2} \exp\left[-||x - y||^2/(4\sigma^2)\right].$$

Heuristically, $\rho(x, y)dxdy$ can be interpreted as the probability of points falling in each of two infinitesimal areas dx and dy. Note that $\rho(x, y)$ is a function of the distance $r = ||x - y||$ alone. Upon standardisation, one obtains the pair correlation function

$$g(r) = \frac{\rho(x, y)}{\lambda^2} = 1 + \frac{1}{4\pi\kappa\sigma^2} \exp\left[-r^2/(4\sigma^2)\right].$$

Since $g(r) = g(r; \kappa, \sigma^2)$ does not depend on ν, we fix the theoretical intensity $\kappa\nu$ at its empirical level $\hat{\lambda} = 176/|W|$. The minimum contrast method can then be used to estimate κ and σ^2. More precisely, this method minimises

$$\int_{r_1}^{r_2} |\hat{g}(r)^q - g(r; \kappa, \sigma^2)^q|^p \, dr$$

with respect to the parameters. In words, the integrated L_p distance between the estimated pair correlation function $\hat{g}^q$ and its model counterpart is minimised. We follow the rule of thumb in [10] and set $q = 1/4$, $p = 2$. The range of integration is set to $[r_1, r_2] = [0.06, 0.35]$. To estimate the pair correlation based on an observation $\{x_1, \ldots, x_n\}$ of X in the window W, since X is isotropic, one may use

$$\hat{g}(r) = \frac{1}{(\hat{\lambda})^2} \frac{1}{2\pi r} \sum_{i=1}^{n} \sum_{j \neq i \in \{1,\ldots,n\}}^{n} \frac{k(r - ||x_i - x_j||)}{\alpha_{ij} |\{z : \partial b(z, ||x_i - x_j|| \cap W \neq \emptyset|}.$$

Here k is a kernel function, for example the Epanechnikov kernel, α_{ij} Ripley's edge correction [16], that is the proportion of the circle centred on x_i of radius $||x_i - x_j||$ that lies in W and $|\cdot|$ denotes area. See also [15].

For the data in Figure 1, the minimum contrast estimators are $\hat{\kappa} = 0.88$, $\hat{\sigma} = 0.08$ and $\hat{\nu} = 81$. In words, 2.17 parents are expected in the region with on average 81 offspring each scattered around their parent with standard deviation 0.08 for the displacements in langitude and longitude. The results should be compared to those of [11]. These authors applied Fisher's linear discriminant function to assign locations to two clusters and pooled the variance estimates with those of a diffuse swarm of aftershocks in another earthquake year, leading to a larger estimate (0.19) for the displacement standard deviation. The fitted and estimated pair correlation functions are shown in Figure 3.

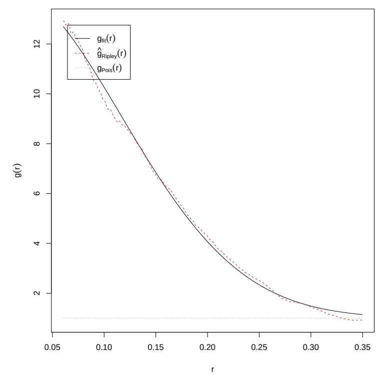

Figure 3: Fitted (solid line) and estimated (dashed line) pair correlation functions for the spatial pattern of shallow earthquakes of magnitude 4.5 or higher recorded during the period October 8–November 7, 2005 using the modified Thomas model.

4 Fitting a Hawkes process

The second model to be considered is a planar Hawkes process. Again, use the generic notation X for the point process that consists of all offspring and assume the parent process Φ is a stationary Poisson process of intensity $\kappa > 0$. In contrast to the trigger processes of Section 3, the points of Φ form a subset of X, called generation zero. As for the Thomas process, each point of Φ generates a Poisson number of offspring with parameter $\nu > 0$. Conditional on the number of offspring, they are independently and identically distributed according to a bivariate normal distribution centred at the parent location with covariance matrix $\sigma^2 I$. The combined offspring of all parents in generation zero forms generation one. This construction is iterated: conditional on generation j, each of its points acts as parent and produces a Poisson number of offspring with parameter $\nu > 0$; conditional on the number of offspring, they are independently and identically distributed according to a bivariate normal distribution centred at the parent location with covariance matrix $\sigma^2 I$ and the combined offspring of all parents in generation j forms generation $j + 1$. The iteration terminates at the first empty generation. In order to ensure this is the case, assume $\nu < 1$.

The point process X is stationary with intensity $\lambda = \kappa/(1 - \nu)$ and pair correlation function

$$\rho(x, y) = \frac{\kappa^2}{(1-\nu)^2} + \frac{\kappa}{1-\nu} \sum_{n=1}^{\infty} (n+1)\nu^n \frac{1}{2\pi n\sigma^2} \exp\left[-||x-y||^2/(2n\sigma^2)\right].$$

Note that $\rho(x, y)$ is a function of the distance $r = ||x - y||$ alone. Upon standardisation, the pair correlation function is given by

$$g(r) = 1 + \frac{1-\nu}{\kappa} \sum_{n=1}^{\infty} (n+1)\nu^n \frac{1}{2\pi n\sigma^2} \exp\left[-r^2/(2n\sigma^2)\right].$$

To estimate the parameters, equate the intensity $\kappa/(1-\nu)$ with its empirical counterpart $\hat{\lambda} = 176/|W|$ and use the minimum contrast method with $q = 1/4$, $p = 2$ and $[r_1, r_2] = [0.06, 0.35]$ as before.

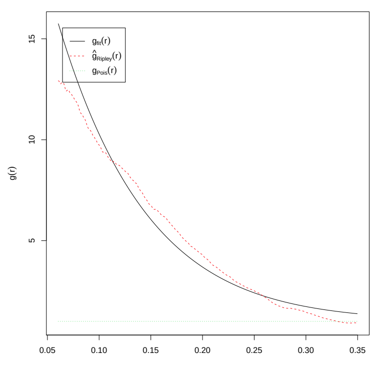

Figure 4: Fitted (solid line) and estimated (dashed line) pair correlation functions for the spatial pattern of shallow earthquakes of magnitude 4.5 or higher recorded during the period October 8–November 7, 2005 using the Hawkes model.

The parameter estimates are $\hat{\kappa} = 7.44$, $\hat{\sigma} = 0.03$ and $\hat{\nu} = 0.9$. In words, 18.45 parents are expected in the region with on average $1/(1-\nu) = 9.54$ descendants each. Note that due to the branching process nature of a Hawkes process, the estimate of the scatter standard deviation is smaller than that of the modified Thomas process of Section 3, the number of parents larger. The fitted and estimated pair correlation functions are shown in Figure 4. Note that the estimated line is too high in the middle range and too low beyond, indicating that the branching structure of the Hawkes clusters may not be appropriate.

5 Model validation

Comparing Figure 3 to Figure 4 , it is clear that the empirical pair correlation function is better matched by a trigger than by a Hawkes process. To validate the former model, we consider the nearest neighbour distance distribution function

$$G(r) = \mathbb{P}(d(x, X \setminus \{x\}) \leq r | x \in X).$$

Since X is stationary, the definition does not depend on the choice of x. Figure 5 plots the empirical nearest neighbour distance distribution together with the upper and lower envelopes based on a hundred independent samples from the fitted model. The point-wise average of the sample estimates $\hat{G}_i(r)$, $i = 1, \ldots, 100$, is the dashed line.

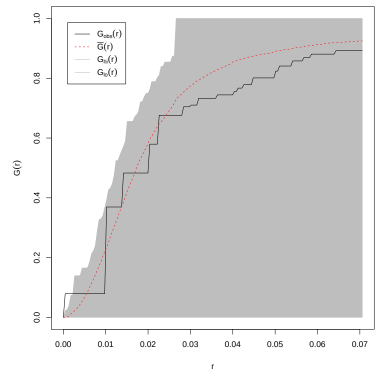

Figure 5: Empirical (solid line) nearest neighbour distance distribution function with upper (G_{hi})and lower (G_{lo}) envelopes over 100 simulations of the fitted modified Thomas process for the spatial pattern of shallow earthquakes of magnitude 4.5 or higher recorded during the period October 8–November 7, 2005.

It can be seen that the empirical G-function lies almost entirely within the grey region bounded by the upper and lower envelopes, indicating a good fit. Note, though, that the lower envelope takes the constant value zero due to the fact that the probability

$$v(W) = \exp\left[-\kappa \int \left[1 - e^{-\nu P(W_{-x};\sigma^2)}\right] dx\right]$$

of finding no points in the rectangular region W is non-negligible. Here $P(\cdot;\sigma^2)$ denotes the bivariate normal distribution measure centred at the origin with covariance matrix $\sigma^2 I$ and W_{-x} is a translation of the observation window W. Tighter envelopes can be obtained by using the information that an earthquake did happen on October 8th with epicentre at $e = (73.59, 34.54)$ and sample conditional on having a parent at the epicentre while adapting κ in such a way that the expected number of parents in W remains constant, i.e. $\tilde{\kappa} = 0.47$. The result is shown in Figure 6. It should be stressed, though, that the plot is not a model validation in the strict sense as conditioning affects the pair correlation function. Indeed, the conditioned model is neither stationary nor isotropic.

Additionally, κ and σ^2 were estimated for a range of settings of the parameters r_1, r_2, p and q in the minimum contrast method. These proved to have hardly any effect on the outcome.

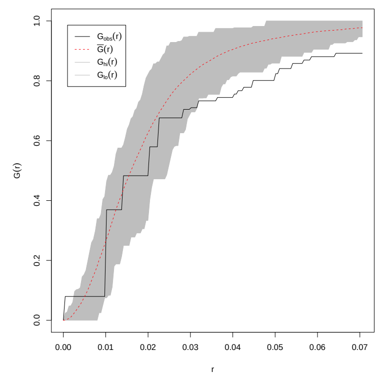

Figure 6: Empirical (solid line) nearest neighbour distance distribution function with upper (G_{hi})and lower (G_{lo}) envelopes over 100 simulations of the superposition of a cluster with normally distributed displacements centred at the epicentre of the initial Kashmir shock with a fitted modified Thomas process for the spatial pattern of shallow earthquakes of magnitude 4.5 or higher recorded during the period October 8–November 7, 2005.

6 Discussion

In this paper, trigger and Hawkes process models were fit to data on aftershocks following the Kashmir earthquake on 8 October, 2005. Although both Poisson cluster processes, these models differ in their offspring generating mechanism. In a trigger process, an unobserved parent generates offspring in an i.i.d. fashion, whereas in the Hawkes process offspring is generated according to a branching process.

The study shows that the trigger model fits the pattern of aftershocks better than the Hawkes process. The observed pattern can thus best be understood as being generated by a parent process of 2.17, i.e. rounded to 2, parents, followed by 176 aftershocks. The Hawkes process model provides entirely different values and identifies a pattern of 18.45 main shocks with on average 9.54 aftershocks. The regular shape of the pattern of aftershocks after two main shocks is clearly visible from the spatial pattern of the aftershocks as well. We note that an appropriate definition of an aftershock does not exist. The existence of the second main shock, in particular, could as well be (and also is most likely) an aftershock of the first earthquake. However, this second major shock almost independently generates a clearly recognizable pattern of aftershocks. Increasing the number of main shocks to higher values such as 18.45 is possible, and the definition of aftershocks does not prohibit this. This combination of the numbers of main shocks and aftershocks though is less likely and intuitively less appealing. Note that in order to describe the pattern in terms of a Hawkes process with an expected number of two main shocks, one would need 0.99 for the mean number of offspring which is dangerously close to the critical value of 1 at which the process explodes.

The trigger process is based upon an entirely different assumption than the Hawkes process. Such a model thus leads to a different interpretation of the mechanism generating earthquakes. This study shows that for the Kashmir data aftershocks are most likely generated by a Poisson cluster process with identically distributed clusters of aftershocks centred around the points of a stationary planar (marked) Poisson process of main shocks. The Hawkes model, i.e. a branching process in which each main shock independently generates a (marked) Poisson process of aftershocks with an intensity function that depends on the parent, each of which again generates aftershocks independently of all others and so on, is less likely. This is remarkable, as the ETAS model as the commonly used model for aftershocks is a Hawkes process.

Acknowledgements

This research was supported by The Netherlands Organisation for Scientific Research NWO (613.000.809). Calculations were done using the `R` library `spatstat` [3].

References

[1] Adamopoulos, L. (1976) Cluster models for earthquakes: Regional comparisons. *Math. Geol.*, **8(4)**, 463–475.

[2] Anwar, S., Stein, A. and Genderen, J. van (2011) Implementation of the marked Strauss point process model to the epicenters of earthquake aftershocks. In *Advances in Geo-Spatial Information Science* (ed W. Shi), ISPRS Book Series. London: Taylor & Francis, to appear.

[3] Baddeley, A. and Turner, R. (2005) Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, **12(6)**, 1–42.

[4] Besag, J. and Diggle, P. J. (1977) Simple Monte Carlo tests for spatial pattern. *Appl. Statist.*, **26(3)**, 327–333.

[5] Brémaud, P., Massoulié, L. and Ridolfi, A. (2005) Power spectra of random spike field and related processes. *Adv. Appl. Probab.*, **37(4)**, 1116–1146.

[6] Daley, D. J. and Vere–Jones, D. (2003) *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*, 2nd edn. New York: Springer–Verlag.

[7] Daley, D. J. and Vere–Jones, D. (2008) *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*, 2nd edn. New York: Springer–Verlag.

[8] Diggle, P. J. (1979) On parameter estimation and goodness-of-fit testing for spatial point patterns. *Biometrics*, **35(1)**, 87–101.

[9] Gelfand, A. E., Diggle, P. J., Fuentes, M. and Guttorp, P. (2010) *Handbook of Spatial Statistics*, Part IV. Boca Raton: CRC.

[10] Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008) *Statistical analysis and modelling of spatial point patterns*, Chichester: Wiley.

[11] Lieshout, M. N. M. van and Stein, A. (2011) Earthquake modelling at the country level using aggregated spatio-temporal point processes. Research Report PNA-1102. Amsterdam: CWI.

[12] Lomnitz, C. and Hax, A. (1966) Clustering in aftershock sequences. In The Earth beneath the Continents (eds J. S. Steinhart and T. Jefferson Smith), pp. 502–508. Washington: American Geophysics Union.

[13] Møller, J. and Torrisi, G. L. (2007) The pair correlation function of spatial Hawkes processes. *Statist. Probab. Lett.*, **77(10)**, 995–1003.

[14] Ogata, Y. (1998) Space-time point-process models for earthquake occurrences. *Ann. Inst. Statist. Math.*, **50(2)**, 379–402.

[15] Ohser, J. (1983) On estimators for the reduced second moment measure of point processes. *Math. Operationsforsch. Statist. Ser. Statist.*, **14(1)**, 63–71.

[16] Ripley, B. D. (1976) The second-order analysis of stationary point processes. *J. Appl. Probability*, **13(2)**, 255-266.

[17] Stoyan, D., Kendall, W. S. and Mecke, J. (1995) *Stochastic Geometry and its Applications*, 2nd edn. Chichester: Wiley.

[18] Vere–Jones, D. and Davies, R. D. (1966) A statistical survey of earthquakes in the main seismic region of New Zealand. *N. Z. J. Geol. Geophys.*, **9**, 251–284.

[19] Vere–Jones, D. (1970) Stochastic models for earthquake occurrence. *J. Roy. Statist. Soc. Ser. B*, **32(1)**, 1–62.