Self-Exciting Point Process Models of Civilian Deaths in Iraq

Erik Lewis   George Mohler   P. Jeffrey Brantingham   Andrea Bertozzi

Abstract

Our goal in this paper is to analyze temporal patterns of civilian death reports in Iraq. For this purpose we employ a branching point process model similar to those used in earthquake analysis. Here the rate of events is partitioned into the sum of a Poisson background rate and a self-exciting component in which events trigger an increase in the rate of the process. More specifically, each event generated by the process in turn generates a sequence of offspring events according to a Poisson distribution. Whereas the background rate is typically assumed to be stationary for seismic activity, such an assumption is not valid in the context of civilian deaths in Iraq. We propose three simple adjustments to account for background rate variation and compare the effectiveness of each model using Iraq Body Count data from 2003 to 2007. Our results indicate that branching point processes are well suited for modeling the temporal dynamics of violence in Iraq.

1   Introduction

Since Operation Iraqi Freedom began on March 20, 2003, over 100,000 Iraqi civilians have died. These deaths have occurred as a result of garden-variety and organized crime, sectarian civil strife, local insurgent actions, and transnational terrorist attacks. As a result of the war, coalition forces and Iraqi security forces have also inadvertently contributed to the total. Given so many violent deaths and so many complex forces at play, one wonders whether there is any regular temporal (and spatial) structure to the violence, or whether it is purely random. Understanding whether there

is such structure is critical to developing approaches to countering heterogenous threats arising in insurgent and active terrorist movements.

In general, temporal events may be either statistically independent, following a Poisson process, or temporally correlated. In the latter case, the term event dependence captures the idea that an initial event can increase or decrease the likelihood of subsequent events in the future. Event dependency for insurgent activity in Iraq was established by Townsley, Johnson, and Ratcliffe (2008). In that paper, the authors analyze three months of IED attacks in Iraq starting in February of 2004. They use a Monte-Carlo method to decide if a pair of events is unlikely to have occurred independently. They conclude that event dependence does occur locally in space-time and explain this phenomenon in terms of a communication of risk similar to the spread of an infectious disease.

Drawing on concepts developed in the study of crime patterns (Short, D'Orsogna, Brantingham, & Tita, 2009), we propose that violence in Iraq arises from a combination of exogenous and endogenous effects. Spatial heterogeneity in background rates is conditioned by fixed environmental characteristics, a fact well-documented in criminology (Andresen, Brantingham, & Kinney, 2010). Event inter-dependence, by contrast, may be understood in the context of rational choice and routine activity theory of crime (Cohen & Felson, 1979; Cornish & Clarke, 1986). In some instances, individuals committing an initial act of violence may later return to the same or a nearby place, within a short period of time, to replicate the successes of the previous event. In other instances, an act of violence by any individual or group may incite reprisals, and counter reprisals leading to a cycle of violence. We attempt to capture these two mechanisms using self-exciting point process models, evaluating our results against the simplest case where violence follows a Poisson process. We find that violence in different spatial locations within Iraq may indeed be partitioned into a background rate and self-excited component. However, we also find that the background rate of violence is non-stationary, with shifts corresponding to major changes in the nature of the Iraqi conflict.

Our analysis will proceed similarly to Mohler, Short, Brantingham, Schoenberg, and Tita (2010) where they analyze burgarly and robbery data in Los Angeles. Related work has also been done

in Short et al. (2009) where the authors show that clusters of residential burglaries can not be explained by risk heterogeneity alone. If we assume that correlations are present between nearby events and that these correlations are positive, then we arrive at what is called a self-exciting point process. Commonly used in earthquake analysis, these models capture both the global Poisson, or independent, nature between clusters as well as the local causal dependence within clusters.

Unfortunately, existing models like those in Egesdal, Fathauer, Louie, and Neuman (2010), Mohler et al. (2010), and Ogata (1998) are not directly applicable for the purpose of modeling civilian deaths in Iraq. One of the central assumptions in previous work is that the background rate driving events is stationary. Because exogenous factors such as political decisions, changes in troop levels, etc. have a significant impact on the rate of events, a stationary background rate is inappropriate. One focus of our work is the development of parametric and non-parametric methodologies for incorporating these exogenous factors into the self-exciting point process framework.

In Section 2, we will introduce point process models, as well as three modifications to standard self-exciting point processes, in order to account for variation in the background rate. In Section 3, we include a description of the data and the assumptions necessary to proceed with analysis. In Section 4, we analyze four different regions of Iraq comparing the effectiveness of each model. In Section 5, we will discuss the implications of our findings and possible directions for future work.

2 Model

A point process N is a random measure on a complete separable metric space S that takes values on $\mathbb{N} \cup \{\infty\}$ (Brillinger, Guttorp, & Schoenberg, 2002). In our case, a convenient way to view a realization of N is that of a list of times $t_1, t_2, ..., t_n$ at which events $1, 2, ...n$ occur. We say that a point process N is orderly if for any time t,

$$\lim_{\Delta t \to 0} \frac{P(N((t, t + \Delta t]) > 1)}{\Delta t} = 0$$

We say a point process is simple if $P(t_i = t_j) = 0$ for all $i \neq j$.

A point process is typically characterized by prescribing its conditional intensity $\lambda(t)$, which represents the infinitesimal rate at which events are expected to occur around a particular time t, given the history of the process up to t, $\mathcal{H}_t = \{t_i : t_i < t\}$ (Ogata, 1988),

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{E[N([t, t + \Delta t))|\mathcal{H}_t]}{\Delta t}.$$

Notice that since the right hand side is a conditional expectation, $\lambda(t)$ is a random variable.

An important example of a point process is the Poisson process. We introduce a Poisson process as a useful benchmark for evaluating self-excitation, since it represents complete randomness. If (t_1, t_2) is an interval of time, then $N(t_1, t_2)$ represents the number of events occurring between time t_1 and t_2. Given disjoint sets $(t_1, t_2), (t_3, t_4)..., (t_{k-1}, t_k)$ where $t_1 < t_2 \leq t_3 < t_4 \leq ... \leq t_{k-1} < t_k$, N is a Poisson process if the finite dimensional distributions $N(t_1, t_2), N(t_3, t_4)..., N(t_{k-1}, t_k)$ each have a Poisson distribution and are independent. Notice that a Poisson process always has a deterministic conditional intensity $\lambda(t)$. If the process is stationary then $\lambda(t)$ is a constant.

We say that a point process N is self-exciting if

$$\text{Cov}[N(t_1, t_2), N(t_2, t_3)] > 0$$

for any $t_1 < t_2 < t_3$. This means that if an event occurs, another event becomes more likely to occur locally in time and space. This is not true for a Poisson process. It has independent increments so $\text{Cov}[N(t_1, t_2), N(t_2, t_3)] = 0$. For more on point processes, see Daley and Vere-Jones (2003) and Daley and Vere-Jones (2008).

We investigate a specific class of point processes termed a Hawkes Process (Hawkes, 1971), which can be represented by the conditional intensity function:

$$\lambda(t) = \mu + k_0 \int_{-\infty}^{t} g(t - t_k) dZ(u)$$

$$= \mu + k_0 \sum_{t_k < t} g(t - t_k) \tag{1}$$

where Z is the normal counting measure (Hawkes & Oakes, 1974). In Equation (1), μ represents the background rate of events, which in most applications is assumed to be constant in time (Zhuang, Ogata, & Vere-Jones, 2002). The second half of the sum describes the self-exciting part of the process and has components k_0 and g, which reflect the magnitude of self-excitation and density at which self-excitation is triggered, respectively. Many choices for the triggering density g have been used (Hawkes, 1971; Ogata, 1988). We use an exponential distribution similar to that in Egesdal et al. (2010)

$$\lambda(t) = \mu + k_0 \sum_{t_k < t} w e^{-w(t - t_k)}. \tag{2}$$

Here, w is a rate of decay constant for the triggering kernel controlling how quickly the overall rate λ returns to its baseline level μ after an event occurs. Such time-limited self-excitation is well documented in the criminological literature (Short et al., 2008; Johnson, 2008). In behavioral terms, k_0 corresponds to the strength of the incentive to replicate a past success, or the strength of the drive to seek retribution for a previous attack, while w^{-1} represents the average time until a repeat event occurs.

Inspection of our data on violent death in Iraq indicates that a stationary background rate μ is unrealistic (Figure 1). An attempt to fit the model in (2) to the data would require that the upward trend be driven exclusively by self-excitation, which on behavioral grounds seems unjustified. Consequently we consider a non-stationary background rate μ as an alternative to the models in Egesdal et al. (2010), Mohler et al. (2010), and Ogata (1998).

The simplest choice for a non-stationary μ is a step function. Since fitting more than five

parameters is difficult, we choose a step function parameterized by three values μ_1, μ_2 and μ_3. Our baseline model becomes:

$$\lambda(t) = \mu_{step}(t) + k_0 \sum_{t_k < t} w e^{-w(t-t_k)} \tag{3}$$

where

$$\mu_{step}(t) = \begin{cases} \mu_1 & \text{for } 0 \leq t \leq t_1 \\ \mu_2 & \text{for } t_1 < t \leq t_2 \\ \mu_3 & \text{for } t_2 < t \leq T \end{cases}$$

We choose t_1 and t_2 based on visual inspection of where the largest jumps in activity occur. Values of μ_1, μ_2 and μ_3 are held constant while fitting the other model parameters.

We consider a second model with a linear increase in the background rate beginning between 400 and 1000 days into the data. The exact onset of this change depends on which city we are considering, but the dates do suggest, in each case, a steady increase in the number of events. The second model becomes:

$$\lambda(t) = \mu_l(t) + k_0 \sum_{t_k < t} w e^{-w(t-t_k)} \tag{4}$$

where

$$\mu_l(t) = \begin{cases} \mu_c & \text{for } 0 \leq t \leq t_c \\ \mu_s(t - t_c) & \text{for } t_c < t \leq T \end{cases}$$

We choose t_c based on inspection.

Finally, we propose a third model based on a non-parametric estimation of μ (Silverman, 1986). We use variable bandwidth kernel smoothing to construct a smoothed version of the data:

$$\hat{\mu}_{sm} = \frac{1}{n} \sum_{i=1}^{n} K\left(\frac{t - t_i}{h_i}\right)$$

6

where

$$K\left(\frac{t - t_i}{h_i}\right) = \frac{1}{\sqrt{2\pi}h_i}e^{-\frac{(t-t_i)^2}{2h_i^2}}$$

and each h_i is the maximum of the k-th nearest neighbor and b_{min} is the minimum bandwidth. Since we need the rate function $\lambda(t)$ to integrate to the total number of events, we introduce a third parameter p to allow for this. Our final model becomes:

$$\lambda(t) = p\mu_{sm}(t) + (1 - p)k_0 \sum_{t_k < t} we^{-w(t-t_k)} \tag{5}$$

The choices of the k-th nearest neighbor and bandwidth b_{min} directly affect the shape of $\hat{\mu}_{sm}(t)$ and thus $\hat{\lambda}(t)$.

To estimate parameters, we use maximum likelihood estimation (Fisher, 1922; Rubin, 1972; Ozaki, 1979). For example, if we want to estimate parameters for the linear model in (4), we get the log likelihood function:

$$\log L(\mu_c, \mu_s, k_0, w | t_1, ..., t_n) = \sum_{t_i : 1 \leq i \leq n} \log(\lambda(t_i)) - \int_0^T \lambda(t)dt. \tag{6}$$

We want to find the maximum of L over all possible values of μ_c, μ_s, k_0 and w subject to the constraint that all four parameters are positive.

To compare models, we use Akaike's Information Criterion (AIC) (Akaike, 1973, 1974). For a given model, the AIC is equal to $2k - 2\log(L)$ where k is the number of parameters in the model and L is the maximum value of the likelihood function. The criterion penalizes a model with more parameters, so a smaller AIC value implies a better model (Bozdogan, 1987). We note that the AIC is a relative scale used to compare different models, and is not a test for goodness of fit. In Egesdal et al. (2010), a self-exciting model for gang violence is compared to a stationary Poisson process with rate equal to the average number of events over the time interval in consideration. In the context of gang retaliations, it is reasonable to assume the background rate is stationary, so it is more natural to compare that process to a stationary Poisson process through the AIC. However

7

due to the secular trend in the Iraq data, a stationary Poisson process is unlikely to give rise to our observed sequence of events. Instead we use the AIC to evaluate each self-exciting model against a corresponding non-stationary Poisson model with self-excitation removed (i.e. $k_0 = w = 0$).

3 Data

We evaluate data on violent deaths in Iraq derived from Iraq Body Count, an organization dedicated to accurately recording all civilian deaths in Iraq (Iraq Body Count, 2008). The number of fatalities linked to any event is not an estimate by the organization, but a count corroborated by at least two reliable news sources. In the data we consider, from March 20, 2003 to December 31, 2007, there are 15,977 events. Each entry in the data contains a start date, end date, minimum number of deaths, maximum number of deaths, town and possibly a district of where the event occurs.

Several simplifying assumptions are necessary to make the temporal analysis straightforward. First, we consider each event as an entry in the data. For example, on December 28, 2007, fourteen people were killed by a car bomb in Al-Tayaran Square in Baghdad. We consider this one event, not fourteen.

Second, we only consider whether an event occurs on a specific day, not the type of event nor how many people were killed. This means no distinction is made between different types of events, nor are any events discarded from the data set because they are seen as unrelated to our analysis. For example, we do not distinguish between IED attacks and gunfire, nor do we ignore all events in the data that are unrelated to IED attacks.

Third, we only consider the start date and not the end date. The entries for which the time of occurrence is a range of days correspond to uncertainty as to when the event happened. For example, if a mass grave is discovered, there is uncertainty as to how and when the deaths occurred. The effect from this simplification depends on the region in consideration. Overall, 93.45% of the events have the same start and end date.

Fourth, we group data by location according to the smallest known region available. For some

cities, like Baghdad, we have events down to each district, but this is not true of most other cities in the data.

Finally, we assume that events recorded on the same day are statistically independent. Since we only have events recorded to the day, we do not know the order in which such events occurred. This means we cannot determine whether these events are correlated with one another. In our model, when we compute the sum for $\lambda(t)$ in (2) we only consider events occurring on dates strictly prior to the day on which a focal event occurs. Technically this means the process is no longer simple. However, we reason that we can analyze the data without allowing events occurring on the same day to effect one another.

To evaluate the nature of this final assumption, we provide a histogram of the frequencey of days with a given number of events (Figure 2). Out of a possible 1,748 days, there are 133 days with no events, while there are 169 days with one event. This means there are 1,446 days with more than one event. We present this histogram to highlight the coarseness of timescale. These events did not presumably happen at the exact same moment in time which means that they could be affecting each other, but we are unable to detect that.

4   Analysis and Results

We examine temporal patterns of violent deaths for four different regions in Iraq including Karkh, Najaf, Mosul, and Fallujah. There is considerable uncertainty about the demographic and sectarian characteristics in each of these regions over the time period. Our choices of these areas are therefore based on inspection of the data above. Karkh is a district at the heart of Baghdad with the most events of any spatial region in Iraq. Najaf is a mid-sized city in central Iraq. It displays obvious clustering, but far fewer events than other regions we analyze. Mosul, in the North, and Fallujah, in the west, have intermediate numbers of events.

Karkh

The first region we consider is Karkh, a district of Baghdad. Over the entire data set there are 2,278 events spanning 1,742 days. Table 1 shows that the AIC values for models with self-excitation are smaller than those for models without self-excitation. For example, the AIC value for the linear model in equation (4) is 905.7, while for the same model with no self-exciting portion the AIC is 1,624.7. The step function and the linear models perform almost identically. This similarity could be explained by the drastic jump in the number of events around 1,385 days into the data, allowing the step function to pick up this feature better than the linear model. However, the smoothed background rate model outperforms the other models with an AIC value of 801.1. This heavily favors the self-exciting point process, which is visually confirmed in Figure 3.

Estimating the parameters for the smoothed background rate model using equation (6), we get $\hat{k}_0 = 1.003, \hat{w} = 0.064$ and $\hat{\mu}_{sm} = 0.6368$. These values all have behavioral interpretations. First, k_0 signifies the average number of direct offspring events that are caused by an event. All four models suggest that every event causes between 0.9 and 1 more event on average. Second, we can interpret w^{-1} as an average time over which we expect an excited event to happen following a background event. Here we have that $w^{-1} = 15.08$ days. The briefness of this interval could correspond to the amount of time a hostile actor needs to prepare for another attack. Finally, since $\int_0^T \lambda(t)dt$ is equal to the number of events in $[0, T]$, we have that $\int_0^T \hat{\mu}(t)dt$ is an estimate for the number of background events in our data set. For the first row of Table 1, where the background rate is stationary (i.e. $\hat{\mu}(t) = 0.060$), the number of background events is estimated at 137, which represents only 6% of the total number of events. Self-excited events would therefore make up 94% of the sample or 2,141. In addition to the higher AIC value, this result points toward a poor fit for the stationary background rate. In contrast, if we consider the smoothed background rate model we have $\int_0^T \hat{\mu}_{sm}(t)dt = 1,451$ events, which is 63.68% of the total number of events. Here, 36.32% or 827 events are the product of self-excitation.

We note the consistency of the estimates for k_0 and w^{-1} across all of the models. This adds weight to the accuracy of the estimates for these two parameters. This consistency is not the case for

other areas in Iraq. We expect that sample size is at least partially responsible for the consistency across models.

Najaf

The second region we consider is Najaf, a medium-sized city 180 miles south of Baghdad. A total of 149 events were recorded in Najaf spanning 1,718 days. AIC values for each model with and without self-excitation are quite similar to one another compared with Karkh (Table 2), due to the smaller sample size. The AIC value of 949.8 for the linear model with self-excitation is noticeably better than that of the other three models. This is different from the other districts, where the smoothed background rate outperforms all other models. In Figure 4, the three large clusters of events around 450 days, 1,200 days and 1,600 days for $\hat{\lambda}_l(t)$ force the estimate for k_0 to be larger than is acceptable for the rest of the data, leading to the spike around 500 days into the data.

Following the same intuition from the previous example, we see that $k_0 = 0.497$ (Table 2) and thus for every two events that occur one offspring event is generated on average. However, we expect to wait only about 8 days for an offspring event to occur. The timescale here is about half as long as the timescale for Karkh. The estimated number of background events from the linear model is $\int_0^T \hat{\mu}_l(t)dt = 94$ which is 63% of all the events, very similar to Karkh.

The consistency of the estimates for w^{-1} is again remarkable. However, the range of possible values for k_0, .497 to 1.012, is much larger. The variablity in k_0 could again be due to the fact that there are far fewer data points in this sample.

Mosul

The third region we consider is Mosul, the second largest city in Iraq. There are 1,300 events in Mosul occurring over 1,718 days. The AIC value of 2,545.1 corresponding to the smoothed background rate model with self-excitation performs better than all other models (Table 3). The step function model performs next best. Figure 5 shows what looks like a steady increase in events that does not drop

off to the same degree as the other areas. One might expect such a pattern to be a perfect candidate for a linear background rate, yet it performs marginally better than a stationary background rate.

The estimates in the fourth row of Table 3 yield $k_0 = 1.05$. Values for w^{-1} vary much more widely from model to model in Mosul. The value of w^{-1} for the smoothed background model is 41.08 days, suggesting a much longer time scale for self-excited events compared to Karkh and Najaf. Note that we are estimating w not w^{-1}, which means that smalls differences in w are amplified when looking at w^{-1}. The estimate for the number of background events in the smoothed background model is $\int_0^T \hat{\mu}_{sm}(t)dt = 657$ events which is 50.5% of the total number of events, slightly less than in Karkh and Najaf. Accordingly, 643 events are attributable to self-excitation.

Fallujah

The last region we consider is Falljuah. There are 501 events in this region over 1,748 days. Here too the AIC for the smoothed background rate model with self-excitation is the smallest at 1,929.8 (Figure 6). The step function model out performs the linear model. This latter result is most likely due to the drop off in events near the end of the data set, favoring a step function as opposed to a linear increase through the end of the time interval. Allowing for two knots instead of one would give the linear model enough freedom to track this change in the background rate. The close fit of the smoothed background model is corroborated in Figure 6.

Similar to Mosul and Najaf, the estimates for k_0 vary from 0.67 to 1.01 between the four models. The estimates for w^{-1} are a little more consistent between models, with 17 days for the smoothed background rate being the shortest. This timescale is fairly close to the estimate for Karkh. The estimated number of background events is $\int_0^T \hat{\mu}_{sm}(t)dt = 302$ which is about 60% of all the events in Fallujah. The corresponding number of self-excited events is 199.

5 Discussion and Conclusion

Self-excited point process models partition the rate of events occurring in time and space into background and self-excited components. Background events are statistically independent of one another, whereas offspring events are triggered by prior events. We have developed self-exciting point process models to describe the time-course of violent deaths in four regions of Iraq dating between March 20, 2003 and December 31, 2007. These models are unique in the statistics literature in that they consider a special class of point-processes where the background rate is non-stationary. Most applications including those in seismology (Ogata, 1988) and in previous studies of crime patterns (Egesdal et al., 2010) have assumed that the background rate is stationary and equivalent to a Poisson Process. We consider four models including: (1) a stationary background rate; (2) a step function background rate; (3) a linear increasing background rate; and (4) a smoothed background rate. The models were fit to observed violent events derived from Iraq Body Count using maximum likelihood estimation and models were compared using Akaikes information criterion (AIC).

In no instance did a stationary background rate model outperform a model with a non-stationary background rate. Moreover, AIC values suggest that models which include a self-exciting component outperform the equivalent model where self-excitation has been removed. These results confirm that self-excitation is an important component of the dynamics of violence in Iraq. Furthermore, integrating the background rate component $\int_0^T \hat{\mu}(t)dt$ for models favored by AIC estimates that background events make up between 50-63% of all observed violent events. Accordingly, self-excited events are estimated to comprise 37-50% of all events. Clearly, self-excitation drives a substantial proportion of the total violence in Iraq.

Our comparisons of different regions within Iraq suggest both commonalities and differences in the specific dynamics of violence. In three of the four regions examined here, including Karkh, Mosul and Fallujah, each initiating event is expected to generate slightly more than one self-excited daughter or offspring event (i.e., $\hat{k}_0 = 1.0194$). In Najaf, every two initiating events are expected to generate approximately one self-excited daughter event (i.e., $\hat{k}_0 = 0.497$). It is possible therefore

that the number of daughter events generated by an initiating event is characteristic across all regions of Iraq. By contrast, there is considerable variability in the time scales over which self-excited events occur. Examining the models favored by AIC suggests that self-excited daughter events may happen on average within one week of an initiating event (Najaf), within approximately two weeks (Karkh, Fallujah), or within slightly more than one month (Mosul). These regional differences may stem from the unique characteristics of crime, insurgent and terrorist violence, as well as the nature of the war, in these areas.

Self-exciting point process models offer a tractable approach to describing the statistical structure of violence in Iraq. There are also good behavioral reasons to suppose that the violence on-the-ground is driven by a combination of background and self-excited components. It is well-known in criminology that environments influence the temporal and spatial structure of crime patterns (Andresen et al., 2010). Areas with higher traffic volume provide greater numbers of potential targets and those that lack natural security designs are more vulnerable to attack. Background rates of crime will vary from area to area based on such environmental heterogeneity. We suppose that environmental heterogeneity plays a similar role in generating variability in background rates of violent death in Iraq, even in instances of insurgent and terrorist suicide attacks. Self-excitation is also well-documented in criminal event patterning (Short et al., 2008; Townsley et al., 2008). Criminal offenders are known to prefer to return to locations that they know and where they can be reasonably certain that they will be successful in their repeated victimization of targets at that site. It is also known that violent retribution is frequently meted out in the same location as and close in time to the initiating offense (Jacobs & Wright, 2006). Swift retribution is necessary to restore ones reputation. We suppose that violent crime in Iraq, including blood feuds and sectarian violence, is also similarly disposed to cycles of violence whereby an initial act of violence brings quick reprisal and counter reprisals. The extent to which the quantitative characteristics of our models such as the number of daughter events and the time scale of their occurrence are dependent upon the Iraqi setting warrants further investigation.

Our results also raise the possibility that intervention strategies can be designed to counteract

self-excitation in patterns of Iraqi violence. If it is know that a large fraction of events generate daughter events, then it may be possible to strategically deflect this mechanism. For instance, if daughter events are generated out of a desire to replicate recent successes, then recognizing and altering the environmental or situational characteristics the facilitated success in the first place may help to decrease the chance of self-excitation. Alternatively, if daughter events are driven by cycles of reprisals, then intervening with the impacted parties may decrease the chance of self-excitation. While there may be general strategies that are applicable across both types of self-excitation, such events are inherently situational and will require a situational response.

Future work could include adding a conditional magnitude to the intensity function, where magnitude would be measured by the severity of the attack. This is routine in the use of models to study earthquake dependencies where large magnitude earthquakes are more likely to generate aftershocks than smaller magnitude earthquakes. A similar dynamic may characterize violent acts, where large magnitude attacks cause more combatants to "pile in." To our knowledge, this has not been done with respect to criminal and other violent activity. It may be difficult to determine the magnitude of a violent attack, but in the present case the number of deaths occurring in each event may serve as a proxy for the magnitude. Another possible direction for the future work involves estimating the background intensity more accurately. Considering a semi-parametric or non-parametric estimation of the background rate, or possibly using other data sets like troop levels could prove effective. In earthquake research, different choices of the triggering function g in (1) have been analyzed and compared (Ogata, 1998). Similar work could be done here to determine accuracy of triggering densities.

6   Acknowledgements

15

References

Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle.* Budapest: Akadémiai Kiadó.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automatic Control*, *AC-19*, 716–723. (System identification and time-series analysis)

Andresen, M. A., Brantingham, P. J., & Kinney, J. B. (2010). *Classics in environmental criminology.* Boca Raton, New York, London: Simon Fraser University Publications and CRC Press.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (aic): the general theory and its analytical extensions. *Psychometrika*, *52*, 345–370.

Brillinger, D. R., Guttorp, P. M., & Schoenberg, F. P. (2002). Point processes, spatial-temporal. *Encyclopedia of Envioronmetrics*, *3*, 1577-1581.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, *44*(4), 588–608.

Cornish, D. B., & Clarke, R. V. (1986). *The reasoning criminal: Rational choice perspectives on offending.* New York: Springer-Verlag.

Daley, D. J., & Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. I* (Second ed.). New York: Springer-Verlag.

Daley, D. J., & Vere-Jones, D. (2008). *An introduction to the theory of point processes. Vol. II* (Second ed.). New York: Springer.

Egesdal, M., Fathauer, C., Louie, K., & Neuman, J. (2010). Statistical and stochastic modeling of gang rivalries in Los Angeles. *SIAM Undergraduate Research Online*.

Fisher, R. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A*, *222*, 309-368.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, *58*, 83–90.

Hawkes, A. G., & Oakes, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Probability*, *11*, 493–503.

Iraq Body Count. (2008). *Iraq body count.* http://www.iraqbodycount.net/ (accessed February, 2008).

Jacobs, B. A., & Wright, R. (2006). *Stret justice: Retaliation in the criminal underworld.* New York: Cambridge University Press.

Johnson, S. (2008). Repeat burglary victimisation: a tale of two theories. *IEEE Trans. Automatic Control*, *4*, 215–240.

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P., & Tita, G. E. (2010). Self-exciting point process modeling of crime. *submitted*.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Amer. Statist. Assoc.*, *83*(401), 9–27.

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

Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Ann. Inst. Statist. Math.*, *31*(1), 145–155.

Rubin, I. (1972). Regular point processes and their detection. *IEEE Trans. Information Theory*, *IT-18*, 547–557.

Short, M., D'Orsogna, M., Brantingham, P., & Tita, G. (2009). Measuring and modeling repeat and near-repeat burglary effects. *Journal of Quantitative Criminology*, *25*(3), 325–339.

Short, M., D'Orsogna, M., Pasour, V., Tita, G., Brantingham, P., Bertozzi, A., et al. (2008). A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences*, *18*, 1249–1267.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis.* London: Chapman & Hall.

Townsley, M., Johnson, S. D., & Ratcliffe, J. H. (2008). Space time dynamics of insurgent activity in Iraq. *Security Journal*, *21*, 139–146.

Zhuang, J., Ogata, Y., & Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *J. Amer. Statist. Assoc.*, *97*(458), 369–380.
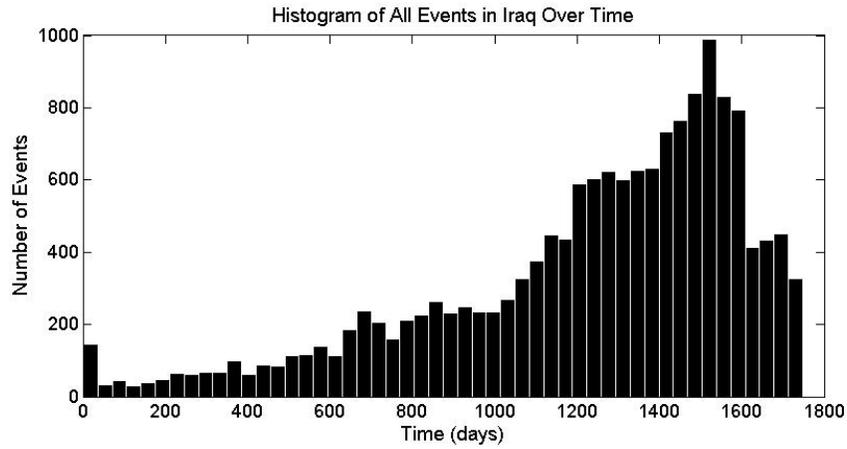
Figure 1: A histogram of all events occurring in Iraq with time on the horizontal axis and number of events on the vertical axis. Over 1,747 days, between March 20, 2003 and December 31, 2007, there are a total of 15,977 events in 50 bins.
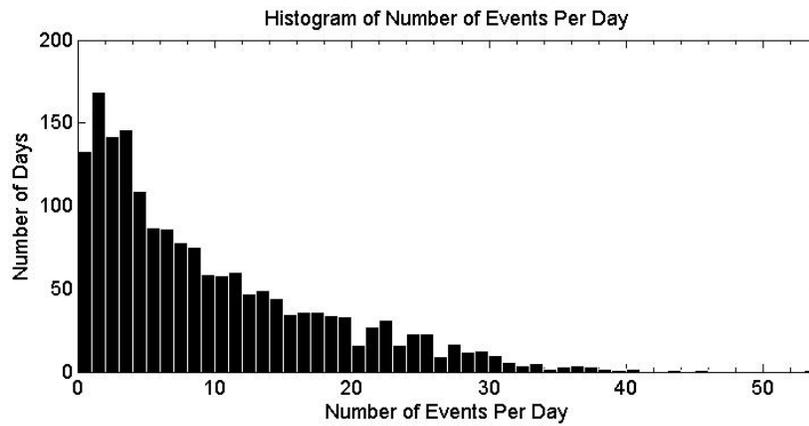

Figure 2: A histogram of the number of events per day over all of Iraq between March 20, 2003 and December 31, 2007. The most events occuring on a single day is 53. The mean number of events on a single day is 9.15.
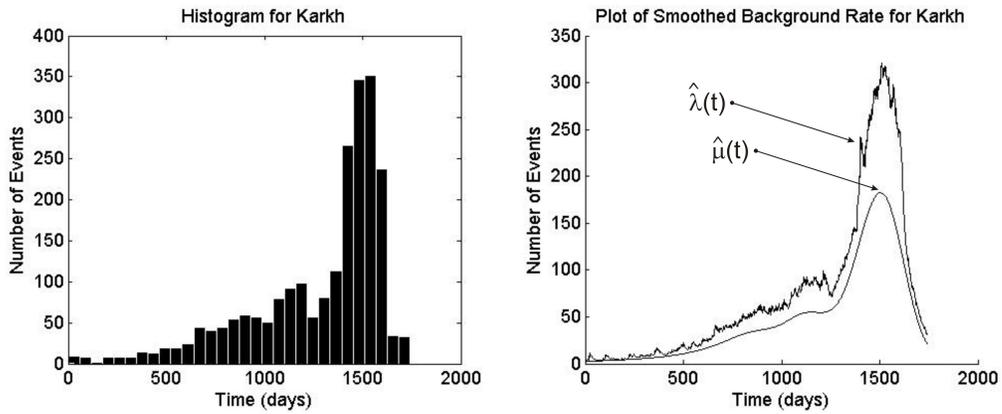
18

Figure 3: A histogram of all events in Karkh (left). The estimated fit of the data for the smoothed background rate model $\hat{\lambda}(t)$ (right). The smoothed background rate $\hat{\mu}_{sm}(t)$ is plotted on the right as well.
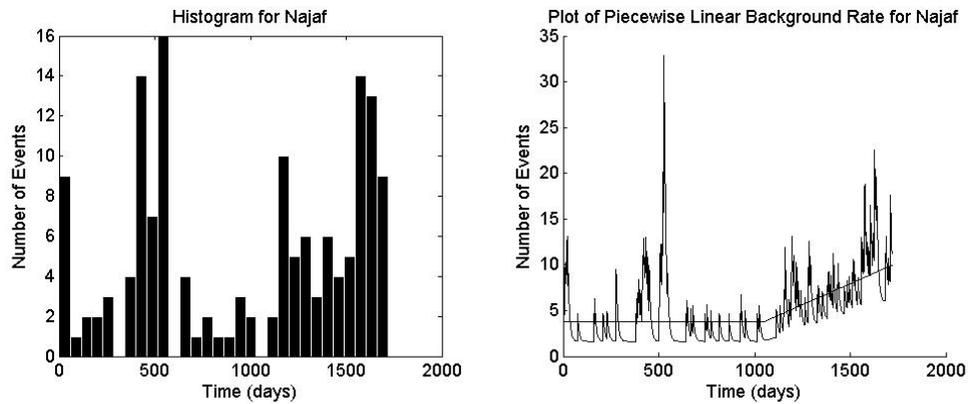

Figure 4: A histogram of all 149 events in Najaf with 30 bins is plotted on the left. The estimated fit for $\hat{\lambda}(t)$ with a linear background rate is plotted on the right (the jagged curve). The fit for the data without self excitation is plotted on the right as well.
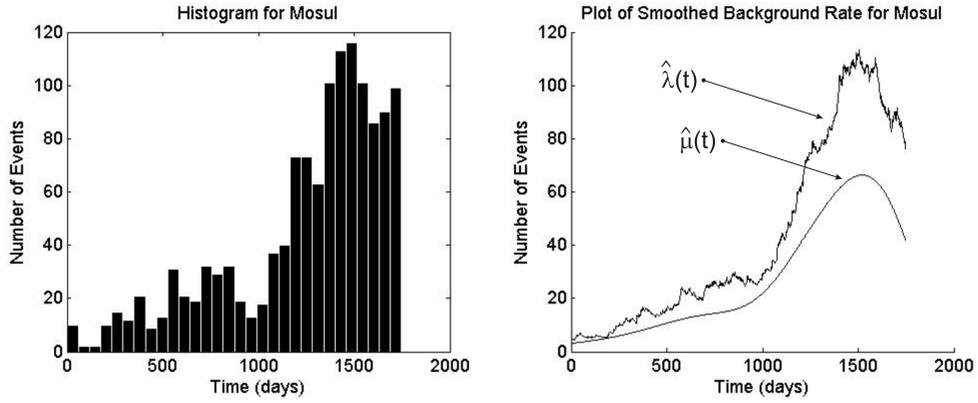
Figure 5: A histogram of the all the events in Mosul with 30 bins is plotted on the left. The estimated fit of the data for the smoothed background rate model $\hat{\lambda}(t)$ is plotted on the right. The smoothed background rata $\hat{\mu}_{sm}(t)$ is plotted on the right as well for reference.
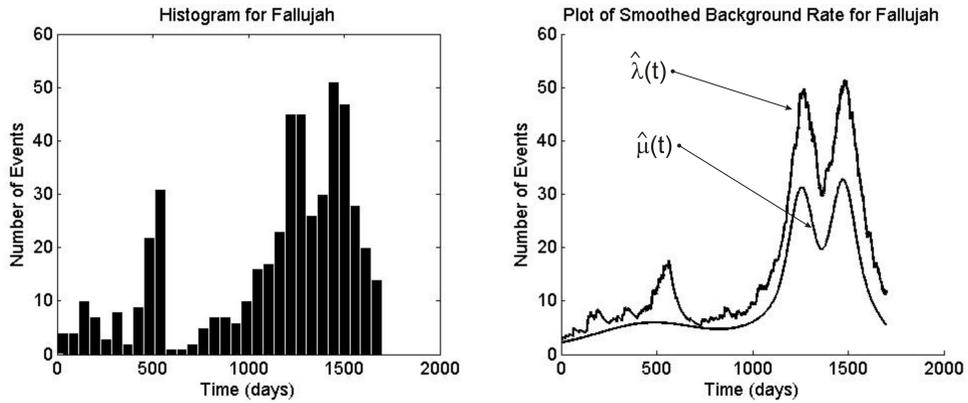

Figure 6: A histogram of the all the events in Fallujah with 30 bins is plotted on the left. The estimated fit of the data for the smoothed background rate model $\hat{\lambda}(t)$ is plotted on the right. The smoothed background rata $\hat{\mu}_{sm}(t)$ is plotted on the right as well for reference.

Table of Parameters for Karkh						
City	$\hat{\mu}$	$\hat{k}_0$	$\hat{w}^{-1}$	Hawkes AIC	No SE AIC	Best Fit
Karkh (μ)	0.060	0.959	19.23	906.0	3,330.6	Hawkes
Karkh (μ_{step})	0.053, 0.1430, 0.053	0.935	19.23	905.6	3,331.1	Hawkes
Karkh (μ_l)	0.0552, 0.0001	0.915	18.59	905.7	1,624.7	Hawkes
Karkh (μ_{sm})	0.6368	1.003	15.08	829.6	855.0	Hawkes

Table 1: The four rows contain estimates for every event in Karkh using models (2), (3), (4), and (5) respectively. For the second row, $t_1 = 661$ and $t_2 = 1,385$ while for the third row $t_c = 400$. For the fourth row, we use the 200-th nearest neighbors and $b_{min} = 80$. The "No SE AIC" (No Self-Excitation AIC) for each row corresponds to estimating the model with k_0 and w set to zero. For example, in the first row this corresponds to a Poisson proess with a rate equal to the average number of crimes per day.

Table of Parameters for Najaf						
City	$\hat{\mu}$	$\hat{k}_0$	$\hat{w}^{-1}$	Hawkes AIC	No SE AIC	Best Fit
Najaf (μ)	0.037	0.590	9.709	963.7	1,028.4	Hawkes
Najaf (μ_{step})	0.032, 0.034, 0.078	0.521	8.772	958.2	1,007.5	Hawkes
Najaf (μ_l)	0.0286, 0.0001	0.497	8.354	949.8	1,004.8	Hawkes
Najaf (μ_{sm})	0.5054	1.012	8.606	959.2	1,004.9	Hawkes

Table 2: Parameter estimates for Najaf using self-excited and non-self-excited models (2), (3), (4), and (5). For the second row, $t_1 = 661$ and $t_2 = 1,385$ while for the third row $t_c = 1,050$. For the fourth row, we use the 50-th nearest neighbors and $b_{min} = 30$.

Table of Parameters for Mosul						
City	$\hat{\mu}$	$\hat{k}_0$	$\hat{w}^{-1}$	Hawkes AIC	No SE AIC	Best Fit
Mosul (μ)	0.0533	1.0024	58.82	2,570.5	3,370.4	Hawkes
Mosul (μ_{step})	0.0969, 0.4169, 0.5639	0.7123	68.28	2,558.4	2,611.2	Hawkes
Mosul (μ_l)	0.0950, 0.0008	0.7354	49.02	2,570.4	2,626.9	Hawkes
Mosul (μ_{sm})	0.6344	1.0548	41.08	2,545.1	2,551.8	Hawkes

Table 3: Parameter estimates for Mosul using self-excited and non-self-excited models (2), (3), (4), and (5). For the second row, $t_1 = 1,050$ and $t_2 = 1,350$ while for the third row $t_c = 975$. For the fourth row, we use the 200-th nearest neighbors and $b_{min} = 150$.

Table of Parameters for Fallujah						
City	$\hat{\mu}$	$\hat{k}_0$	$\hat{w}^{-1}$	Hawkes AIC	No SE AIC	Best Fit
Fallujah (μ)	0.0394	0.8788	23.52	1,952.0	2,277.0	Hawkes
Fallujah (μ_{step})	0.0605, 0.0350, 0.1717	0.6739	19.85	1,944.5	2,011.5	Hawkes
Fallujah (μ_l)	0.0447, 0.0002	0.7758	20.33	1,949.8	2,054.8	Hawkes
Fallujah (μ_{sm})	.6020	1.0094	17.20	1,929.8	1,946.2	Hawkes

Table 4: The four rows contain estimates for every event in Fallujah using models (2), (3), (4), and (5) respectively. For the second row, $t_1 = 575$ and $t_2 = 1,025$ while for the third row $t_c = 400$. For the fourth row, we use the 80-th nearest neighbors and $b_{min} = 30$.