Supplementary Material for "(No) Spillovers in reporting domestic abuse to police"

Lara Vomfell, Jan Povala

1 Likelihood

We did not write out the full log likelihood in Equation (21) in the main text. We do this here, but for simplicity, we simplify the background to be denoted by $m_0\mu(t,s)$ and the trigger by $\theta_M f(t,s)$. Altogether, our model in Equation (5) of the main text with parameter vector $\Theta = \{m_0, \theta_0, \theta_1\}$ has the following log likelihood:

$$
\begin{aligned}
\ell(\Theta) &= \sum_i \log \lambda(t_i, s_i) - \int_0^T \int_X \lambda(t,s)\mathrm{d}t\mathrm{d}s \qquad (1)\\
&= \sum_i \log\left(m_0\mu(t_i,s_i) + \sum_{j:t_j<t_i} \theta_{M_j} f(t_i - t_j, s_i - s_j)\right)\\
&\quad - \int_0^T\int_X \left(m_0\mu(t,s) + \sum_{j:t_j<t} \theta_{M_j} f(t-t_j, s-s_j)\right)\mathrm{d}t\mathrm{d}s\\
&= \sum_i \log\left(m_0\mu(t_i,s_i)\right) + \sum_i \log\left(\sum_{j:t_j<t_i} \mathbb{I}(M_j=0)\theta_0 f(t_i-t_j, s_i-s_j)\right)\\
&\quad + \sum_i \log\left(\sum_{j:t_j<t_i} \mathbb{I}(M_j=1)\theta_1 f(t_i-t_j, s_i-s_j)\right) - \int_0^T\int_X m_0\mu(t,s)\mathrm{d}t\mathrm{d}s\\
&\quad - \int_0^T\int_X \sum_{j:t_j<t} \mathbb{I}(M_j=0)\theta_0 f(t-t_j, s-s_j)\mathrm{d}t\mathrm{d}s\\
&\quad - \int_0^T\int_X \sum_{j:t_j<t} \mathbb{I}(M_j=1)\theta_1 f(t-t_j, s-s_j)\mathrm{d}t\mathrm{d}s.
\end{aligned}
$$

We can now take the derivative of Equation (1) with respect to m_0, using the chain rule:

$$
\begin{aligned}
\partial\ell(\Theta)/\partial m_0 = 0& \\
&= \sum_i \frac{\mu_{\text{trend}}(t_i)\mu_{\text{weekly}}(t_i)\mu_{\text{daily}}(t_i)\mu_{\text{area}}(s_i)}{\lambda(t_i, s_i)} \\
&\quad - \int_0^T \int_X \mu_{\text{trend}}(t)\mu_{\text{weekly}}(t)\mu_{\text{daily}}(t)\mu_{\text{area}}(s)\mathrm{d}t\mathrm{d}s. \qquad (2)
\end{aligned}
$$

Again, using the chain rule we can now take the derivative of Equation (1) with respect to θ_0 (the derivative with respect to θ_1 is analogous) and obtain:

$$
\begin{aligned}
\partial\ell(\Theta)/\partial\theta_0 = 0& \\
&= \sum_i \frac{\sum_{j:t_j<t_i} \mathbb{I}(M_j = 0)g(t_i - t_j)h(s_i - s_j)}{\lambda(t_i, s_i)} \\
&\quad - \int_0^T \int_X \sum_{j:t_j<t} \mathbb{I}(M_j = 0)g(t - t_j)h(s - s_j)\mathrm{d}s\mathrm{d}t. \qquad (3)
\end{aligned}
$$

2 Inference algorithm

We can write out the inference procedure for the model explained in the main text in Section 4.3 in algorithmic form. The procedure consists of two main steps: The initialisation and the inference loop.

In the initialisation stage, we obtain initial values for the daily, weekly, trend, area and triggering components. We then use those to calculate the entire background component $\mu(t, s)$. We need to do this calculation twice: Once to obtain $\mu(t_i, s_i)$, that is the background value at all events i and once more to obtain $\int \mu(t, s)\mathrm{d}t\mathrm{d}s$, that is the background integrated over the entire study area. We then repeat this step to obtain the trigger at all events i $f(t_i, s_i)$ and integrated over the study area $\int f(t, s)\mathrm{d}t\mathrm{d}s$.

With those quantities in hand, we then update m_0 and θ_M from some initial guesses and then calculate the intensity λ, again at all events i and integrated over the study area.

We then enter the inference loop where essentially the procedure repeats: We obtain updated values for the daily, weekly, trend, area and triggering components; we calculate the background and triggering components at the events and integrated over the study area. We update m_0 and θ_M, and calculate λ at all events i and integrated over the study area. When m_0 and θ converge, we break the inference loop.

More formally, we write:

Algorithm 1 Inference algorithm

Input: n_p, b_{daily}, b_{weekly}, b_{trend}, b_{area}, b_g, b_h, m_0 and θ_M
Initialisation
Initialise components μ_{daily}, μ_{weekly}, μ_{trend}, μ_{area}, $g(t)$, $h(s)$,
Calculate background $\mu(s_i, t_i)$ and $\int_0^T \int_X \mu(s,t)\mathrm{d}t\mathrm{d}s$
Calculate trigger $g(t-t_i)h(s-s_i)$ and $\sum_i \int_{t_i}^T \int_X g(t-t_i)h(s-s_i)\mathrm{d}t\mathrm{d}s$
Update m_0 and θ_M
Calculate intensity $\lambda(t_i, s_i)$ and $\int_0^T \int_X \lambda(t,s)\mathrm{d}t\mathrm{d}$
while not convergence **do**
  Update components μ_{daily}, μ_{weekly}, μ_{trend}, μ_{area}, $g(t)$, $h(s)$
  Calculate background $\mu(s_i, t_i)$ and $\int_0^T \int_X \mu(s,t)\mathrm{d}t\mathrm{d}s$
  Calculate trigger $g(t-t_i)h(s-s_i)$ and $\sum_i \int_{t_i}^T \int_X (t-t_i)h(s-s_i)\mathrm{d}t\mathrm{d}s$
  Update m_0 and θ_M
  Calculate intensity $\lambda(t_i, s_i)$ and $\int_0^T \int_X \lambda(t,s)$
  Check convergence of m_0 and θ_M

3 Initialization of triggering functions

Although our estimation of $g(\cdot)$ and $h(\cdot)$ is non-parametric and solely driven by the data, we consider different initializations of the estimates to assess the sensitivity of the inference. Firstly, we consider a monotonically decreasing function with peak at zero for both $g(\cdot), h(\cdot)$:

$$g(t_i - t_j) = \frac{1}{(t_i - t_j)/24 + 1/24}, \qquad t_i > t_j$$

$$h(s_i - s_j) = \frac{1}{1 + (s_i - s_j)^\top (s_i - s_j)}.$$

Secondly, we consider forms which introduce a delayed peak for $g(\cdot)$. This is motivated by the delayed criminal reaction model, where the intensity of the triggering is set to peak after some delay instead of reaching the peak at zero and decaying monotonically afterwards (Gilmour, 2019). There are criminological theories that do explain such delays in homicide and burglary near-repeat occurrences. In the context of domestic abuse, we expect that the spillover effect is delayed by the time it takes to process the event and the time for information to propagate. Following Gilmour (2019), we consider the following form:

$$g(t_i - t_j) = \omega^2 (t_i - t_j) \exp(-\omega(t_i - t_j)), \qquad t_i > t_j,$$

where $\omega = 1/15$ is the delay parameter which makes the function $g(\cdot)$ peak at approximately $t_i - t_j = 15$ days.

4 Model Selection Output

Table 1 shows the grid search for optimal model parameters. For each combination of parameters we compute the prediction error (loss) according to (26) in Section 5.1 of the main text.

5 Voronoi residuals

5.1 Approximating Voronoi Residuals

To compute the resisudal for cell c, R_c, we uniformly sample S locations, $\{s^{(i)}\}_{i=1}^{S}$ in the cell c and compute

$$R_c \approx |A_c|\frac{1}{S}\sum_{i=1}^{S}\lambda(s^{(i)}), \tag{4}$$

where $|A_c|$ is the area of cell c and

$$\begin{aligned}
\lambda(s) &= \int_0^T \lambda(t, s|\mathcal{H}_t)\mathrm{d}t && (5)\\
&= \int_0^T \big\{\mu(t,s) + \sum_{j:t_j<t} g(t-t_j)h(s-s_j)\big\}\mathrm{d}t && (6)\\
&= \int_0^T \mu(s)\mu(t)\mathrm{d}t + \int_0^T \sum_{j:t_j<t} g(t-t_j)h(s-s_j)\mathrm{d}t && (7)\\
&= \mu(s)\int_0^T \mu(t)\mathrm{d}t + \int_0^T \sum_{j:t_j<t} g(t-t_j)h(s-s_j)\mathrm{d}t, && (8)
\end{aligned}$$

in which the integration is approximated using equally-spaced evaluation points, $\{t_i\}_{i=1}^{N_T}$, with interval size $\frac{T}{N_T}$:

$$\lambda(s) \approx \mu(s)\frac{T}{N_T}\sum_{i=1}^{N_T}\mu(t_i) + \frac{T}{N_T}\sum_{i=1}^{N_T}\sum_{j:t_j<t_i} g(t_i-t_j)h(s-s_j). \tag{9}$$

5.2 Analysis of Voronoi Residuals

As discussed in Section 5.3.2 in the main text, the residuals should follow a gamma distribution: $R_c \sim 1-\Gamma(3.569, 3.569)$. Figure 1 shows the histogram (normalised to sum to 1) of the residuals and the density function of the theoretical probability distribution. Evidently, the histogram does not match the expected distribution. The large spike at the value of one means that our model predicts 0 events, when in fact there was an event. Upon further investigation, our model predicts zero events for very small Voronoi tessellation polygons. The

Table 1: Caption

b_{daily}	b_{weekly}	b_{trend}	loss
0.021	0.17	5.0	184.532889
0.021	0.17	10.0	184.366152
0.021	0.17	15.0	184.299209
0.021	0.17	20.0	184.263409
0.021	0.33	5.0	184.493149
0.021	0.33	10.0	184.328862
0.021	0.33	15.0	184.276286
0.021	0.33	20.0	184.249867
0.021	0.5	5.0	184.483746
0.021	0.5	10.0	184.329928
0.021	0.5	15.0	184.270184
0.021	0.5	20.0	184.250233
0.042	0.17	5.0	184.491920
0.042	0.17	10.0	184.347753
0.042	0.17	15.0	184.290113
0.042	0.17	20.0	184.258421
0.042	0.33	5.0	184.450288
0.042	0.33	10.0	184.312629
0.042	0.33	15.0	184.267339
0.042	0.33	20.0	184.242055
0.042	0.5	5.0	184.442933
0.042	0.5	10.0	184.306565
0.042	0.5	15.0	184.263047
0.042	0.5	20.0	184.239526
0.083	0.17	5.0	184.439228
0.083	0.17	10.0	184.316948
0.083	0.17	15.0	184.274103
0.083	0.17	20.0	184.249890
0.083	0.33	5.0	184.406666
0.083	0.33	10.0	184.290781
0.083	0.33	15.0	184.253614
0.083	0.33	20.0	184.240578
0.083	0.5	5.0	184.401316
0.083	0.5	10.0	184.289823
0.083	0.5	15.0	184.251150
0.083	0.5	20.0	184.240045
0.125	**0.67**	**30.0**	**184.223818**
0.12	0.67	40.0	184.230693
0.12	0.67	50.0	184.235441
0.12	1.0	30.0	184.230291
0.12	1.0	40.0	184.230829
0.12	1.0	50.0	184.235758
0.21	0.67	30.0	184.226855
0.21	0.67	40.0	184.224510
0.21	0.67	50.0	184.236617
0.21	1.0	30.0	184.229135
0.21	1.0	40.0	184.230260
0.21	1.0	50.0	184.235262
0.29	0.67	30.0	184.226865
0.29	0.67	40.0	184.228421
0.29	0.67	50.0	184.233843
0.29	1.0	30.0	184.228178
0.29	1.0	40.0	184.229987
0.29	1.0	50.0	184.235493

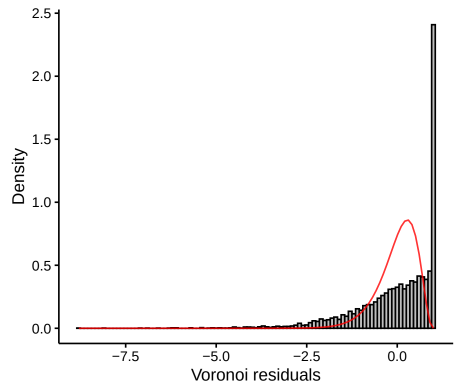


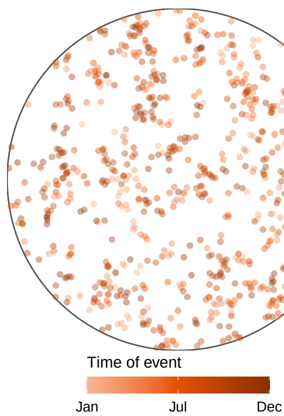
Figure 1: Histogram of residuals from the model fit to the domestic abuse data for the period of 2018, and approximate theoretical density (continuous line).

small polygons are a result of many events occurring within an immediate vicinity of each other. In our dataset, this is caused by the mapping of the spatial locations of events into pre-defined 'snap points' to preserve privacy. To prevent duplicate points, we added a small random perturbation to the location of each event as part of the data preparation step. Even with this intervention, the distributional skew of the residuals persists.

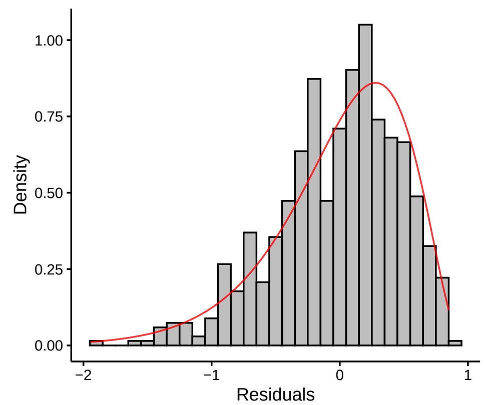
To further confirm this effect, we generated two sets of synthetic data for a one-year period on a circular domain as shown in Figure 2. The top rows show a dataset where 540 original events are spread evenly throughout the space and 210 follow-up events occur in a spatial vicinity (point are separable visually on the map on the left). The bottom row shows 405 original events spread uniformly across the space and 245 follow-up events occurring within the *immediate* vicinity of the original events, making the points visually inseparable. The right panels of both rows show Voronoi residuals after the model fit and their theoretical distributions. It is clear that in the setting where the follow-up events are near-duplicates of the original events (Figure 2c), we observe a similar pattern in the Voronoi residuals as for the real domestic abuse data: an unexpected spike at the value of one. For the more regular dataset, shown in Figure 2a, the empirical distribution of the residuals closely matches the theoretical distribution.

References

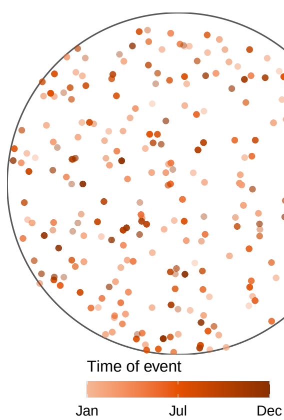
Gilmour, C. (2019) *Self-Exciting Point Processes and Their Applications to Crime Data.* Ph.D. thesis, University of Strathclyde, Glasgow.



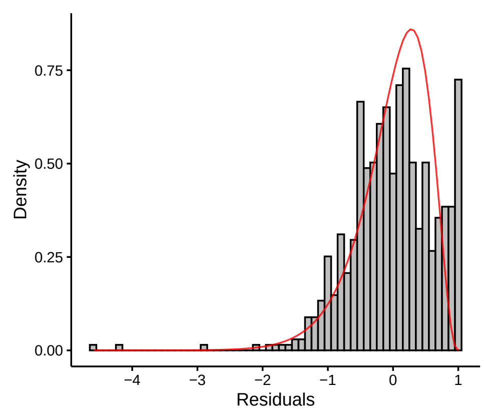
(a) Synthetically generated events over a one-year period. Out of the total 750 events, 540 are triggered by and located in a spatio-temporal vicinity of a past event.



(b) Histogram of residuals from the model fit to the synthetic data shown in Figure 2a. The plot also shows the theoretical distribution of the residuals (red).



(c) Synthetically generated events over a one-year period. Out of the total 750 events, 405 are triggered by and located in a *very immediate* spatio-temporal vicinity of a past event. Note that some points are so close to other points that they appear as a single point in this plot.



(d) Histogram of residuals from the model fit to the synthetic data shown in Figure 2c. The plot also shows the theoretical distribution of the residuals (red).

Figure 2: Synthetic experiments for a circular domain over a one-year period to demonstrate the sensitivity of the distribution of Voronoi residuals.