

Forecasting social conflicts in Africa using an Epidemic Type Aftershock Sequence model

Gilian van den Hengel

Philip Hans Franses

EI2018-31

Econometric Institute

Erasmus School of Economics

Abstract

We propose to view social conflicts in Africa as having strong similarities with earthquake occurrences and hence to consider the spatial-temporal Epidemic Type Aftershock Sequence (ETAS) model. The parameters of this highly parameterized model are estimated through simulated annealing. We consider data for 2012 to 2016 to calibrate the model for four African regions separately, and we consider the data for 2017 to evaluate the forecasts. These forecasts concern the amount of future large events as well as their locations. Examples of our findings are that the model predicts a cluster of large events in the Central Africa region, which was not expected based on past events, and that in particular for East Africa it apparently holds that small conflicts can trigger a larger number of conflicts.

Key words: ETAS model; social conflicts; Africa; Forecasting

JEL codes: D74, F51, O55

This version: August 2018

Correspondence: PH Franses, Econometric Institute, Erasmus School of Economics, POB 1738, 3000 DR Rotterdam, the Netherlands, phone: +31104081273, email: franses@ese.eur.nl

Introduction and motivation

Social conflicts data and earthquakes data show strong similarities. First, they seem to correspond to self-exciting processes. Long periods of non-activity correspond with building up tension, and once an event happens, the tension is released, perhaps causing more events. Such a triggering increases with the size of the event and may decrease over time and distance. A second similarity concerns clustering. That is, earthquakes often take place at the edge of tectonic plates and for conflicts it often occurs that conflicts are witnessed around country or region borders. Triggered events seem to occur more frequently near their previous events than further away.

To exploit these similarities, we propose to analyze social conflicts data using a popular model in the earthquake literature, and this is the so-called Epidemic Type Aftershock Sequence (ETAS) model. To illustrate, we use social conflicts data in Africa. ETAS models are a special type of a self-exciting point processes and are widely used in studying earthquake occurrences. In this model, events can occur spontaneously and events can be triggered by other events. Each event can trigger new events, which in turn can trigger new events. The model describes the intensity and locations of spontaneous background events and the triggering dynamics.

The outline of our paper is as follows. In Section 2 we describe our sample with social conflicts data for Africa, and we suggest that indeed such data bear similarities with earthquake data. In Section 3 we discuss the model and a method to estimate its parameters. Next, we discuss ways to evaluate the quality of the model, and a method to create forecasts. Section 4 brings the ETAS model to the data for Africa where we take the data for 2012 to 2016 as the in-sample data, and we consider the data for 2017 as our forecast sample. We see that the model fits the data well, and also that out-of-sample forecasts can be informative and accurate. Section 5 concludes with limitations, as there are several, and with various avenues for further research, of which we think there are many too.

Social conflicts in Africa

We obtain the conflict data from the Armed Conflict Location & Event Data (ACLED) Project¹. The project collects the dates, actors, types of violence, locations, and fatalities of all reported political violence and protest events across Africa ever since 1997. Before we can analyze the data, we need to compile a proper database. We define a social conflict (in this database) as an event which consists of four numbers, that is, the event date, location (latitude and longitude) and the number of fatalities. As exact time stamps are unavailable, we assume that each conflict happened at 12:00 noon².

The numbers of fatalities are estimated as the numbers reported by any source material and are not verified by the ACLED project itself. In order to be most reliable, the ACLED project uses the most conservative estimates. For example, if a news article mentions *several*, *many* or *few* fatalities, it is recorded as 10. *Dozens* is recorded as 12. If a report mentions *hundreds* of casualties or a *massacre*, it is recorded as 100. The location of a conflict is recorded as the location of the nearest town or city and if a larger region is described, the provincial capital of that region is used. Furthermore, if a conflict lasts several days, the event is recorded as a new event each day with the average number of fatalities over those days. We take the data as they are and assume that the dates, locations and numbers of fatalities are accurate.

We analyze conflicts between 2012 and 2017 with at least two fatalities. This time window and fatality range results in 17334 events in total with around 2500 to 7000 events in each of the major African regions. This number of events should be large enough to accurately estimate the model in each region³. Using more years of data or including events with one fatality would make model estimation computationally infeasible.

In Figure 1 we present the cumulative number of events and fatalities over time, using dual axes. We see that the average number of events per day increases slightly after 2014 but remains relatively constant over the period. The number of fatalities per event also remains

¹ <https://www.acleddata.com/data/>

² The complete description of methodological accuracies and definitions can be found on the ACLED website: <https://www.acleddata.com/resources/methodology/>.

³ Our estimation routine will be implemented in the SEDA software package, and this number of observations is regarded as properly sized, see Lombardi (2017).

relatively constant, except for an increase between 2014 and 2015, which could be addressed to the start of the Second Libyan Civil War in 2014.

To explore the spatial distribution of conflicts, we display all events in this period in a map of Africa, see Figure 2. The size of the dots represents the number of fatalities. Events with more than one hundred fatalities are in red. Figure 2 shows that there are several large clusters of events. For example, we see many events with more than a hundred fatalities in the top right corner of Nigeria, which is the area where Boko Haram is active. In the top right corner of Egypt, conflicts resulting from the Arab Spring can be seen. Furthermore, there are large clusters of conflicts around the border of Sudan with South-Sudan which associates with the South Sudanese Civil War. Other notable clusters can be found in the eastern part of the Democratic Republic of the Congo and in the south of Somalia. Overall, and as expected, we see that conflicts tend to be located more around country borders than elsewhere.

To explore the spatial distribution of the events over time, the latitude and longitude of each event are presented in Figure 3. The red lines represent the latitude and longitude of the geographical center of Africa. We see that there are more events in the northern half of Africa than in the southern half, and more events in the eastern half than in the western half. In the top panel of the graph, we see that most conflicts are clustered around the equator and that there are relatively fewer events in the south of Africa. In the bottom panel of the graph, we see that events are more evenly distributed across the longitude. Over time, however, events tend to occur around the same latitudes and longitudes. Note the appearance of events around 30 degrees latitude from 2013 onwards, which can be associated with the many conflicts resulting from the Arab Spring.

Next, we discuss the distribution of the sizes, or magnitudes, of the conflict events. In Table 1, the fatality quantiles are given. It can be seen that the distribution of fatalities is quite skewed, that is, most events have a small number of fatalities and relatively few events have a large number of fatalities. Over half of the events have 4 fatalities or less, 99% of the events has up to 100 fatalities and the largest event has 600 fatalities.

The ETAS model we will consider for our data originates from Ogata (1988, 1998) and it assumes that the magnitudes of events follow the Gutenberg-Richter law. So, before we can apply our model, we have to align the data. This law is given by

$$N = 10^{a-bM} \quad (1)$$

where N is the number of events with magnitude M or larger and a and b are constants. If one were to estimate these parameters for actual data, the value of b is typically close to 1 and in that case, events with one larger magnitude are ten times less likely to happen. To make the conflicts follow this law, we group the events based on their fatalities and assign magnitudes to each group in such a way that the number of events at each increasing magnitude decreases by a factor of ten. The derivation of the correspondence between fatalities and magnitudes is given in the Appendix.

In the end, we obtain the magnitude distribution for the conflict events as given in the Table 2, where we adopt a where the magnitude range is 3.0, 3.1, 3.2, ... , 7. The full table is given in the Appendix. This table only shows the integer magnitudes. In the left two columns, the correspondence between fatalities and magnitude is shown. The next two columns show the percentage of events expected by the Gutenberg-Richter law at each magnitude, together with the observed percentage. The last two columns show the expected and observed absolute numbers of events. We see that with each magnitude step, the percentage and number of events decreases roughly by a factor ten.

The expected and observed number of events for each magnitude is presented in Figure 4, together with the log 10 of these numbers (bottom panel). If the magnitude distribution follows the Gutenberg-Richter law, the blue bars in the figure should correspond with the red line. Estimating the b value in (1) using Maximum Likelihood results in a value of 0.95 with standard error 0.01. Hence, this way of assigning magnitudes to fatalities gives a magnitude distribution that resembles the distribution of an earthquake catalog (which is the way that data are called in the earthquake literature), which is assumed by the model.

The ETAS model

In this section we will discuss the ETAS model used to study the social conflicts. The model parameters will be estimated using the simulated annealing algorithm proposed in Lombardi (2015). Furthermore, we will discuss several ways of testing the model and describe the way to make forecasts.

Representation

Earthquake occurrences have been modeled using a variety of point process models and the most popular is the Epidemic Type Aftershock Sequence (ETAS) model. In this model, events are divided into two categories. There are background events and triggered events. Background events are events that occur spontaneously. Triggered events occur as a result of other events. Each event can trigger new events, which in turn can again trigger new events. Zhuang et al. (2002) provide a list of common features of ETAS models, and these are (a) The occurrence rate of background events depends on location and magnitude, but is independent of time, (b) The magnitude of a background event is independent of its location, (c) Each event produces offspring events independently and the number of offspring events produced depends on the magnitude of the event, (d) The occurrence time of an offspring event depends only on the time difference from its ancestor and is independent of the magnitude, (e) The location of an offspring event depends on the location of its ancestor, and (f) The magnitude of an offspring event is independent of the magnitude of its ancestor. These six features explain the design of the ETAS model.

One of the first models is the temporal ETAS model proposed by Ogata (1988) to describe the origin times and magnitudes of earthquakes. This model has been extended to also include the spatial aspect of earthquake occurrences in Ogata (1998) and Zhuang et al. (2002). Below, we will use this spatial-temporal ETAS model for the social conflicts.

The occurrences of events can often be completely described by the conditional intensity function of the process, which gives the probability of an event with a certain magnitude occurring at a specific place at a specific time. This function often consists of two parts, that is, a

part describing the occurrence rate of background events and another part describing the way events trigger new events.

In our study we use the model as defined in Lombardi (2017), which is similar to the one in Ogata (1998). The conditional intensity function in this model is given by

$$\lambda(t, \mu, x, y | H_t) = (\mu \cdot u(x, y) + \sum_{T_i < t} g(t, x, y; T_i, M_i, X_i, Y_i)) f(m) \quad (2)$$

where event i is described by a time T_i , location (X_i, Y_i) and magnitude M_i . This intensity function consists of three parts, that is, the spatial background distribution $u(x, y)$, the decay function $g(t, x, y; H_i)$ and the magnitude probability density function $f(m)$. Each of these components will be discussed next.

The magnitude probability density function is based on the Gutenberg-Richter law (hence the reason to scale that data as in the appendix) and is given by

$$f(m) = \frac{\beta \cdot \exp(-\beta(m - M_c))}{1 - \exp(-\beta(M_{max} - M_c))} \quad (3)$$

where $\beta = b \log(10)$ and M_{max} is the magnitude of the largest event. The M_c refers to the minimum magnitude for which the catalog is complete, that is, the magnitude for which all events with magnitude $M \geq M_c$ are recorded.

The rate at which event i can trigger new events decreases with distance and time from that event. This is described by the decay function $g(t, x, y; H_i)$ which is given by

$$g(t, x, y; H_i) = \frac{k \cdot \exp(\alpha(M_i - M_c))}{(t - T_i + c)^p} \cdot \frac{c_{d,q,\gamma}^i}{(r_i^2 + d^2 \cdot \exp(2\gamma(M_i - M_c)))^q} \quad (4)$$

Here, H_i is the history of events before event i . The decay function in (4) consists of two parts. The first part describes the temporal decay rate and the second part the spatial distribution. The temporal decay is governed by the parameters $\{\alpha, c, k, p\}$. The parameter α measures the influence of excess magnitude $M_i - M_c$ in generating offspring events and k is a normalizing constant for the number of offspring events. A small value for α implies that the effect of the

magnitude on triggering new events is small, and thus that smaller events can trigger a larger amount of events more easily, see Kumazawa et al. (2014). As usual, we assume that larger events cause more aftershocks than smaller events, that is, $\alpha > 0$. As the number of offspring events cannot be negative, we have that also $k > 0$. It is hard to register all aftershocks after a large event and this incompleteness is measured by the parameter c . Finally, the parameter p is the rate of decay over time. We would expect $p > 1$, as this implies that each event can only generate a finite number of offspring events.

The second part of (4) determines the spatial probability distribution. The term r_i is the distance to event i and $c_{d,q,\gamma}^i$ is a normalization constant such that the integral over the entire region equals 1. The term $d^2 \cdot \exp(2\gamma(M_i - M_c))$ measures the influence of magnitude on the spatial extent of the aftershock region for event i . Similar to p , the parameter q measures the spatial decay rate. If we ignore any spatial components in the model and integrate over the entire region, the right-hand term of (4) becomes 1 and we obtain the decay function for the temporal model in Ogata (1988).

The first term in (2) consists of a parameter μ and the spatial background distribution $u(x,y)$. The parameter μ is the Poisson rate of background events. The background distribution determines how these events are distributed across the region. We will estimate the spatial background distribution following the procedure in Lombardi (2017), which is largely based on the iterative kernel method proposed by Zhuang et al. (2002). For this, the region of interest is first divided into N_c cells C_i ($i = 1, 2, \dots, N_c$) to make a discrete approximation. The cells are of size 0.5×0.5 degrees latitude/longitude, which corresponds to an area of approximately $55 \times 55 = 3025$ square kilometers. The background rate can vary among cells but is assumed to be homogeneous within each cell. We can then write $u(x, y) = u_i/A_i$ where u_i is the probability of a background event occurring in cell i with area A_i .

The probabilities u_i are estimated using the iterative kernel method proposed by Zhuang et al. (2002), which uses the following function for the background distribution, that is,

$$\mu(x, y) = \frac{1}{T} \sum_{j=1}^N (1 - \rho_j) \frac{1}{2\pi d_j} \exp\left(-\frac{(x-x_j)^2 + (y-y_j)^2}{2d_j^2}\right) \quad (5)$$

where T is the length of the study period, N the total number of events, ρ_j is the probability that event j is a triggered event and d_j is a variable bandwidth. The probability ρ_j is given by the ratio between the intensity produced by all previous events and the total intensity at that time and place, that is,

$$\rho_j = \frac{\sum_{i=1}^{j-1} g(t_j - t_i, x_j - x_i, y_j - y_i | m_i)}{\lambda(t_j, x_j, y_j | H_{t_j})} \quad (6)$$

Finally, d_j is a variable bandwidth defined as the smallest disk around event j which includes at least $n_p = 10$ events, as in Zhuang et al. (2002). This way, events happening in rural areas have a further reaching effect than events that happen in densely populated areas, which makes regions with high and low event densities more comparable. Following Lombardi (2015), we can use (5) to estimate the probabilities u_i by

$$u_i = \frac{\mu(X_i, Y_i)}{\sum_{j=1}^{N_c} \mu(X_j, Y_j)}, \quad i = 1, 2, \dots, N_c \quad (7)$$

These probabilities then give the spatial background distribution, where the sum over all cells equals 1. Together with the average background rate μ , the decay function $g(t, x, y; H_i)$ and the magnitude distribution $f(m)$, we now have described the conditional intensity function of the process. This function will be used for calculating the log likelihood to estimate the parameters of the model.

The log-likelihood function is given by Vere-Jones (1970), and it reads as

$$\log L(\theta | H_t) = \sum_{i=1}^N \ln[\lambda(T_i, X_i, Y_i | H_{T_i})] - \iiint_{T,R} \lambda(t, x, y | H_t) dt dx dy \quad (8)$$

For the calculation of the log likelihood, the background probabilities u_i from (7) are needed for the cells in which an event has taken place.

Hence, the parameters of the model are $\{\mu, k, p, c, \alpha, d, q, \gamma; u_i, i = 1, 2, \dots, N'_c\}$ where N'_c is the subset of nonempty cells. These parameters can be estimated by maximizing the log likelihood function using the simulated annealing algorithm, which will be discussed below.

Due to the high dimensionality of the likelihood function, there is a chance that a parameter is not identified and becomes arbitrarily large or small during the estimation. Therefore, we will impose certain parameter restrictions on the eight parameters in the model. These restrictions are based on the restrictions used for modeling earthquake occurrences and on values that are physically desirable, see Ogata (1998). These restrictions are given in Table 3. The branching ratio of the process gives the average number of events triggered by an event and is given by Lombardi (2017) and equals

$$\phi = \int_0^\infty \int_{M_c}^{M_{max}} \frac{k \cdot \exp(\alpha(M_i - M_c))}{(t+c)^p} f(m) dm dt \quad (9)$$

If $\phi > 1$ due to $p < 1$, events are generated faster than they die out and the process becomes explosive. The four parameters in the temporal component, α , p , c , and k influence the number of events in the simulated catalogs. Furthermore, the parameter μ has a positive linear effect on the number of events and the parameters that determine the spatial decay, d , γ and q have no effect on the number of events. The spatial parameters describe the distribution of the events over the region. A higher value for α and k lead to more events and a higher value for c leads to fewer events. For the temporal decay rate p there is a small region in which the process is not explosive. However, the size of this parameter region varies depending on the values of the other parameters.

Estimation

The parameters will be estimated using a simulated annealing algorithm, which is a method for approximating the global optimum of a given function. It originates from annealing in metallurgy, describing the physical process of reducing defects by heating and then slowly cooling the material. Even though this method is unlikely to find the optimal solution, it can often find a very good approximation of the optimum. In particular, it can be preferable over (quasi) Newton methods in situations with a large number of independent variables like in our

case. As the algorithm finds a good solution in a relatively short amount of time, we can run it many times, providing a probability distribution for the parameters and the background distribution. These can be used to evaluate model uncertainties. We will run the algorithm $N_{SIM} = 100$ times.

The algorithm starts by generating a random solution, after which it searches for a new point in that neighborhood. This search is based on a probability distribution that depends on the so called temperature of the process. The method accepts all points that raise the objective function (in case of maximization) and also accepts points that lower the objective, with a certain probability. As such, it avoids being trapped in local maxima. By decreasing the temperature, the probability of accepting a worse solution is lowered when the solution space is explored.

More formally, we adopt the simulated annealing procedure used to estimate the ETAS models as described in Lombardi (2015). It consists of an initialization, a loop and a cooling scheme.

To initialize set the count $k = 0$ and generate a random starting point $\vec{\theta}_{opt}$ and set the initial temperature $T_0 > 0$. The starting temperature must be high enough that any solution can be selected and is calculated based on the function to be optimized, see Lin et al (1993).

The loop is a random search for a better solution around the local maximum. First, set $\vec{\theta}_k = \vec{\theta}_{opt}$ and generate the next candidate \vec{y} from a multi-dimensional Cauchy distribution G . We move to this point with a certain probability. Sample $z \in [0,1]$ uniformly and move to the newly generated point if $z \leq A(\vec{\theta}_{opt}, \vec{y}, T_k)$, where the acceptance function A is the Metropolis criterion in Sen and Stoffa (2013), that reads as

$$A(\vec{\theta}_{opt}, \vec{y}, T_k) = \min \left\{ 1, \exp \left(\frac{f(\vec{y}) - f(\vec{\theta}_{opt})}{T_k} \right) \right\} \quad (10)$$

If the log-likelihood at this new point is larger, we set this new point as the current optimum, that is,

$$\text{LogL}(\vec{y}|H_t) > \text{LogL}(\vec{\theta}_{opt}|H_t) \rightarrow \vec{\theta}_{opt} = \vec{y}$$

Once a new optimum is found, the temperature is lowered according to the cooling schedule proposed in Ingber (2000) and Lombardi (2017), that is,

$$T_k = T_0 \exp \left[-13.8 \cdot \exp \left(-\frac{3.4}{D} \right) \cdot k^{\frac{1}{D}} \right]$$

where $D = 8$ is the number of parameters in the model. The two constants -13.8 and 3.4 were found to provide a fast algorithm based on a simulation study in Lombardi (2017). After decreasing the temperature, the loop is repeated until convergence. Events that took place before the study period can still have an effect on events inside the study period. We will therefore use a one year burn-in period when estimating the model. Hence, the estimation period consists of five years from 2012 to 2016, where 2012 will be used as a learning period and the four other years, 2013-2016, will be used as the study period. The data of the last year 2017 will be used for evaluating out-of-sample forecast performance.

Inference

If the model describes the dynamics of the conflicts well, the residuals are expected to follow a stationary Poisson process with a unit rate. The residuals are obtained by a transformation of the time axis, as in Ogata (1998), that is,

$$\Lambda(t) = \int_0^t \iint_R \int_{M_c}^{M_{max}} \lambda(t', x, y, m | H_{t'}) dt' dx dy dm \quad (11)$$

These residuals give the expected number of events with magnitude larger than M_c up to time t and in the region R . If the residuals follow a stationary Poisson process with rate one, the inter-event times follow an exponential distribution. This can be tested using a one-sided Kolmogorov-Smirnov test (KS-test). We will also perform a Wald-Wolfowitz runs test, which tests the null hypothesis that the inter-event times are stationary and not autocorrelated.

Furthermore, we can perform a visual check by plotting the number of events expected by the model against the observed number of events. This can be done for the transformed times, all events, background events and triggered events. The expected numbers of events are

calculated by integrating the intensities over time, magnitude and space. The expected number of total events is given by

$$N_{exp}^{All}[\theta|H_t] = \int_T \iint_R \lambda(t, x, y|H_t) dt dx dy \quad (12)$$

The expected number of background events is given by

$$N_{exp}^{Back}[\theta|H_t] = \int_T \iint_R \mu \cdot u(x, y) dt = \mu(T_2 - T_1) \quad (13)$$

and the expected number of triggered events is

$$\begin{aligned} N_{exp}^{Trig}[\theta|H_t] &= N_{exp}^{All}[\theta|H_t] - N_{exp}^{Back}[\theta|H_t] \\ &= \int_T \iint_R \frac{k \cdot \exp(\alpha(M_i - M_c))}{(t - T_i + c)^p} \cdot \frac{c_{d,q,\gamma}^i}{(r_i^2 + d^2 \cdot \exp(2\gamma(M_i - M_c)))^q} dt dx dy \end{aligned} \quad (14)$$

where T_1 and T_2 are the start and end times of the sample period. Note that we also integrate over the magnitude distribution, but as the magnitudes are independent of the parameters, this integral equals one.

The observed number of all events (background + triggered) can be obtained directly. As we cannot know for certain whether an event is a background or a triggered event, we cannot count them. Therefore, the number of observed background and triggered events are calculated as the sum of the probabilities of being a background or triggered event. The probability that event i is a background event is given in Lombardi (2017), and it reads as

$$Prob_i^B = \frac{\mu \cdot u(X_i, Y_i)}{\lambda(T_i, X_i, Y_i|H_{T_i})} \quad (15)$$

and the probability that event i is a triggered event is given by $Prob_i^T = 1 - Prob_i^B$. The sum of these probabilities over all events determine the 'observed' number of background and triggered events.

The number-of-events test compares the number of events in the conflict catalog (the dataset) with the number of events expected by the model. The expected number of events and its distribution are obtained from simulating a large number of event catalogs and calculating the number of events in each of them. Specifically, this is done in the following way. Simulate $N_{SIM} = 100$ catalogs based on the estimated parameters in an ETAS model. Then, fit a normal distribution to the number of events in each simulated catalog. Next, calculate the median and the 95% confidence bounds, and with these we compute the probability that we observe more events than in our dataset.

Finally, we will test the performance of the model by making a forecast for the number of events and their location in the hold-out sample. We focus on events with magnitude $M_f > 4$, which for the African countries data corresponds to 22 or more fatalities. For this, we simulate $N_{SIM} = 100$ catalogs using the empirical ETAS model and calculate the expected number of events with magnitude $M_f > 4$ in each cell of the regions in the hold-out sample. The expected number of events in simulated catalog i in cell C_j is given in Lombardi (2017), and it reads as

$$f_{ij} = \int_T \iint_{C_j} \int_{M_f}^{M_{max}} \lambda(t, x, y, m | H_t^i) dm dx dy dt \quad (16)$$

where H_t^i is the history of the process up to time t for catalog i . For each cell we then take the median number of events of all simulated catalogs as the forecast. The total number of events can then be calculated as the sum of the median number of events over all cells. The forecasted number of events and locations will then be compared with their observed number and locations to examine accuracy.

In Van den Hengel (2018), the above described estimation and evaluation methodology is examined using extensive simulations. There it is concluded that the ETAS model can be reliably implemented for simulated data. It can happen though that parameters are estimated at their boundary value, and this is not unexpected given that the ETAS model is heavily parameterized.

Modeling and forecasting social conflicts in Africa

The continent of Africa is often divided in five major geographical regions. These are North-, West-, Central-, East- and South Africa. These regions are displayed in Figure 5. The numbers of events in each region for the learning period, study period, forecasting period and total, are given in Table 4. As the number of events in the region South Africa is too small to reasonably estimate the model, we will restrict the study to the other four major regions. Some summary statistics regarding the number of fatalities and magnitudes are given in Table 5. This last table shows that the size of events is quite similarly distributed in each region.

As the model assumes the magnitude probability density function based on the Gutenberg-Richter law in (1), we check whether this assumption more or less holds for the catalogs in each region. If it deviates too much, we might assume the wrong magnitude distribution for the data. For this, we estimate the parameter b in (1) for each region, and the estimation results are given in Table 6. As the value of b is close to one for each region, we continue with the assumed probability density function in the conditional intensity function in (2)⁴.

For each region, the study area is determined by the smallest rectangular area covering all events. The spatial background distribution is estimated for this area. As this means that it can cover areas that are not in the study region, such as other countries or parts of the ocean, the cells outside the study region are removed. The effect on the total probability should be minimal, as these areas have a very low probability due to a lack of events. The estimated spatial background distributions are given in Figure 6 to 9 for North, West, Central and East Africa, respectively. Furthermore, a map with the actual conflicts in each region is given adjacent to the map of the background distribution. We see that the estimated distributions are in line with what we would expect from the catalogs, as the areas with high probability correspond with areas with high event density.

⁴ Due to an unfortunate bug in the software, events with a negative longitude cause an error. We will therefore restrict our study region to the right side of the prime meridian. This means that out of a total of 17829 events, 495 events will be omitted in North- and West Africa. This should not affect the estimation results in a large way as this number of events is relatively small.

Parameter estimates

For each region we have estimated the eight parameters of the spatial-temporal ETAS model $N_{SIM} = 100$ times. The results for the North, West, Central and East Africa catalog are given in Tables 7, 8, 9, and, respectively. In these tables, the parameter estimates of the best model, their median and distributions are given. Estimates with an * are estimates that are on, or are very close to, the boundary of their constraint. We see that this is the case for the parameters c , d , q and γ in each region, where parameters c , d , and γ go to their lower bound and q goes to its upper bound. The parameter c measures the incompleteness of the catalogs after a large event. As the conflicts in our database are reported on a daily basis, the next triggered event occurs the next day or later. It therefore makes sense that there are no undetected aftershocks right after an event. Furthermore, a value of $c = 0$ would not much change the shape of the likelihood function. The term $d^2 \cdot \exp[2\gamma(M_i - M_c)]$ measures the influence of the magnitude on the spatial extent of the aftershock region for event i . As d goes to its lower bound and γ tends to become arbitrarily small, it suggests that the magnitude has little influence on the region of aftershocks. If we decrease the lower bound of the constraint for parameter d it also tends to become arbitrarily small. In that case, the only influence on the spatial decay is the distance from the event.

The other parameters, μ , k , p , and α vary across the regions. For easy comparison, they are displayed together in Table 11. What is most noticeable is that $p < 1$ in the four regions. This implies that the process is explosive: each event generates an infinite amount of triggered events when $t \rightarrow \infty$. However, this effect is not obvious from the data due to the short time span.

The value for the background rate μ in each region corresponds to the number of background events, which is given in Table 12. The numbers of expected and observed events are very close, which indicates that the model describes the data well. These expected and observed numbers are given over time in Figures 10 for North Africa. Results for the other three regions are very similar. When we divide the number of background events by the number of days, we obtain an average of 0.47, 0.48, 0.36 and 0.86 background events per day for the regions North, West, Central and East, respectively. Hence, the estimates for μ resemble the number of background events in the catalog.

The parameter k governs the expected number of direct aftershocks caused by an event. The estimation results in Table 11 suggest that events cause the most aftershocks in the North region ($k = 0.0376$) and the least in the West region ($k = 0.0226$). The parameter α measures the effect of magnitude in generating aftershocks. A small value indicates that the triggering rate is less dependent on magnitude. Hence, smaller events can trigger a larger amount of events more easily if α is small, see Kumazawa and Ogata. (2014). This suggests that in the East region ($\alpha = 0.0893$) small conflicts can trigger larger amount of conflicts easier as compared to the other three regions.

Residuals

Figure 10 shows that the transformed event times, using (11), seem to follow a Poisson process with unit rate. In the bottom two figures, we see that this is also the case for background and triggered events. This suggests that the model is able to capture the temporal behavior of the process.

Whether the model captures the temporal behavior of the process can be tested more formally using the KS-test and the runs test. For each region, the null hypothesis that the inter-event times are exponentially distributed is rejected at the 1% level. This suggests that the model does not perfectly capture the temporal behavior of the process. In contrast, the runs test fails to reject the null hypothesis at the 1% level that the inter-event times are stationary and not autocorrelated.

Further results

The results of the Number of Events tests are given in Table 13. For each region, the median number of events and 95% confidence interval are given for the $N_{SIM} = 100$ simulated catalogs based on the estimated ETAS model, as described above. Furthermore, the number of events in the conflict catalog is given, together with the probability that we observe a number of events larger than the number in the conflict catalog. We see that for the regions North, Central and East, the number of events in the conflict catalog lies within the 95% confidence bounds. For the West region the actually observed number of events is larger than expected by the model.

Forecasting for 2017

For the forecasts we simulated $N_{SIM} = 100$ catalogs and calculated the median expected number of events in each cell in the year 2017. The total forecast can then be calculated as the sum over all cells. We only forecast large events, which are events with magnitude $M \geq 4$. This corresponds to events with 22 fatalities or more. The forecasts for each region are given in Table 14. The forecasts are compared with two naive forecasts and the observed number of events. For the first naive forecast we use the average number of events in the period 2012-2016 and for the second naive forecast we use the number of large events of last year. We see that for the regions North and West, our ETAS based forecasts are much more accurate than the naive forecasts. In the Central and East region, our prediction underestimates the number of large events.

The results for the forecast locations of events is given in Figures 11 to 14. For forecasting of event locations, we consider three other predictions for comparison. Two of them are naive predictions. We predict that future events happen at places where past events have happened and we predict that future events happen in highly populated areas or around country borders. For this, a map with the events is displayed, together with a map that shows every city with a population larger than 1000. The country borders can also be seen on this map. The third other prediction is based on our model and is the estimated spatial background distribution. This predicts that future events will happen at places where there is a high probability of background events. These three maps are displayed together with our forecasts map. The forecast map also includes the actual locations of the large events, indicated by an *.

First, we notice for the various graphs that there is little connection between population density and conflict areas. This can for example be seen in the West catalog where most conflicts take place in the north-east part of Nigeria or in the East catalog where most conflicts occur in South Sudan. The locations of previous conflict do seem to be a good predictor for future events.

For our forecasts map, the similarity with the spatial background distribution is noticeable. A more detailed look reveals that each cell in the forecasting map is highly correlated with its equivalent cell in the background distribution. One explanation could be that, as the parameter $p < 1$, the process becomes explosive and the simulations for the forecasting period are not accurate. Nonetheless, we believe that our forecasts maps are good predictors of the actual locations of future large events.

On average, the forecasts coincide with the locations in the event catalogs. There are also some areas in which our model is extra informative. These are regions with relatively few or small events in the past, but still show a high expected number in the future. For example, in the North catalog, we would not expect a future event in the center of Libya, as the previous events at that place were relatively small. Other such regions can be found in the West catalog, where two events in the west of Nigeria occurred. Finally, one of the clearest examples is in the Central catalog. In the forecast map, there is a cluster of large conflicts in an area of high intensity in the south of the Democratic Republic of the Congo. This was not expected from the event map, where only a few small events can be seen. Hence, even though the forecasting map is highly correlated with the spatial background distribution, it does produce useful predictions for the locations of future large events.

Conclusion and further research

We have successfully estimated a spatial-temporal ETAS model on social conflict data in Africa. Using this model we could make an out-of-sample forecast for the number of large future events and their locations. The estimated spatial background distribution showed that the probability distribution of background events was in line with the locations in the conflict catalogs, that is, high probability areas were associated with high event density areas. The Number of Events test showed that for three out of four regions, the number of events in the conflict catalog were close to the number of events expected by the model. Only for the West region the number of events was slightly underestimated. The KS-test showed that the inter-event times were not exactly exponentially distributed and this indicated that the model does not perfectly capture the temporal behavior of the data. The runs test however suggested that the inter-event times are stationary and not autocorrelated.

We were able to use the estimated model to make predictions for the number and locations of future large events, out-of-sample. The forecast locations were very much in line with the locations of actual events.

Limitations and future research

The ETAS model and estimation procedures have been optimized specifically for estimating earthquake catalogs, and our study is the first to bring it to social conflict data. We used the constants for the cooling procedure and it seemed to work well. This, however, does not mean that these are the optimal constants for estimating conflict catalogs. Moreover, the parameter constraints in estimating the model are obtained from the earthquake literature and these might be too strict for modeling conflicts. However, when we relaxed these constraints we ran estimation problems, as some parameters became arbitrarily small or large. This problem still persisted after much trial and error with different constraint values. Hence, further research in the estimation procedure and parameter constraints regarding social conflicts might produce better results.

The model estimated in this study is the stationary ETAS model, which assumes that the background rate does not change over time. This is not always the case for conflict data. For example, the rise of the Arab Spring in North Africa caused a stream of events. Our study can perhaps be extended using a non-stationary ETAS model as in Kumazawa and Ogata (2014), where the background rate varies over time. Furthermore, the models for earthquakes work with continuous time, while the conflicts are only reported with a daily precision. Another possible extension might therefore be a discrete time implementation. Finally, the model used in this study assumed the Gutenberg-Richter law for the magnitude distribution, as it closely resembles the magnitude distribution of earthquakes. One could consider a different distribution that is more natural to the size distribution of conflicts.

Regarding the data, it would be interesting to consider our model to smaller regions. Regions of interest could be the north-east of Nigeria, where the Boko Haram insurgency is active, the conflicts around Lake Chad or around the Sudan South-Sudan border. Our study focused more on aggregate conflict dynamics and it would be interesting to see if the findings hold at a smaller scale.

Appendix: Magnitude distribution of social conflicts

Let M_{min} and M_{max} be the minimum and maximum of our magnitude range. For a bin size of 0.1, that is, magnitudes are reported with one decimal accuracy as in our data, we have a total of

$$N_{steps} = 10(M_{max} - M_{min}) + 1$$

magnitude steps. One extra step is included as M_{min} is in the range. Given that a percentage x of events has magnitude M_{min} we want a percentage $x/10$ to have magnitude $M_{min}+1$, a percentage $x/100$ to have magnitude $M_{min} + 2$, and so on. In smaller magnitude steps, an increase of 0.1 corresponds to a frequency decrease of $\omega = 10^{1/10} \cong 1.259$. Then, given that a percentage x of events has magnitude M_{min} , we want a percentage x/ω to have magnitude $M_{min}+0.1$, a percentage x/ω^2 to have magnitude $M_{min} + 0.2$, and so on.

To determine the percentage of events with a magnitude of M_{min} , we solve the following equation

$$\sum_{n=0}^{N_{steps}} \frac{x}{\omega^n} = 100\%$$

which gives us $x \cong 20\%$ for $10 < N_{steps} < 100$.

Next, we have to estimate the number of magnitude steps. The largest event is one event with 600 fatalities in a dataset of about 20000 events in total. Hence, the largest event, which will be given the largest magnitude M_{max} has an occurrence frequency of 1 in 20,000 or 0.005%. If the percentage of events with magnitude M_{min} is roughly 20%, then the percentage of events with magnitude $M_{min} + 4$ will be $20\%/10^4 = 0.002\%$, which is about the frequency of the largest event. Hence, if we use 4 full magnitude steps, or 40 smaller steps, the frequency of the smallest and largest events follow the Gutenberg-Richter law. This means we will use a magnitude range of $(M_{min}, M_{max} = M_{min} + 4)$. Lastly, we have to determine M_{min} , which can be chosen rather arbitrarily. In earthquake studies, the minimum magnitude is often set at $M_{min} = 3$, as this associates with an earthquake considered of reasonable size. We will also set $M_{min} = 3$ which

gives a magnitude range of $\{M_{min}, M_{max}\} = \{3, 7\}$. For this magnitude range, $x \cong 20.56\%$ and the set $\{x/\omega^n\}_{n=0}^{n=40}$ gives the percentages of events for each magnitude $\{3.0, 3.1, \dots, 7.0\}$.

Next, we have to assign to numbers of fatalities to a certain magnitude, in such a way that the resulting magnitude distribution matches the above distribution as close as possible. For this, we first calculate the cumulative percentages $\{\sum_{k=0}^n \frac{x}{\omega^k}\}_{n=0}^{n=40}$ for each magnitude, which we use as fatality quantiles. If a number of fatalities corresponds to multiple quantiles, it will be assigned to the magnitude of the lowest quantile. For each magnitude, we have the number of fatalities, expected and observed percentage of events, and the expected and observed number of events. The next table presents the final results.

Table: The magnitude distribution of the social conflict data. For each magnitude, there is the corresponding number of fatalities, the expected percentage of events, the observed percentage of events, the expected number of events and the observed number of events

Magnitude	Fatalities	% expected	% observed	Expected	Observed
3	2	20.57	26.94	3565	4669
3.1	3	16.34	14.49	2832	2511
3.2	4	12.98	17.29	2250	2997
3.3	6	10.31	8.64	1787	1498
3.4	8	8.19	4.76	1419	825
3.5	10	6.50	10.00	1127	1733
3.6	10	5.17	0.00	896	0
3.7	12	4.10	3.91	711	677
3.8	15	3.26	3.64	565	631
3.9	19	2.59	2.02	449	351
4	22	2.06	1.77	357	306
4.1	27	1.63	1.29	283	224
4.2	31	1.30	1.07	225	186
4.3	38	1.03	0.93	179	162
4.4	45	0.82	0.33	142	58
4.5	50	0.65	0.85	113	147
4.6	58	0.52	0.44	90	76
4.7	67	0.41	0.34	71	59
4.8	78	0.33	0.27	57	47
4.9	97	0.26	0.05	45	9
5	100	0.21	0.33	36	57
5.1	107	0.16	0.13	28	22
5.2	129	0.13	0.07	22	13
5.3	150	0.10	0.00	18	0
5.4	183	0.08	0.07	14	12
5.5	204	0.07	0.05	11	8
5.6	248	0.05	0.05	9	8
5.7	310	0.04	0.03	7	6
5.8	349	0.03	0.00	6	0
5.9	400	0.03	0.05	4	8
6	400	0.02	0.00	4	0
6.1	409	0.02	0.02	3	3
6.2	458	0.01	0.01	2	2
6.3	590	0.01	0.00	2	0
6.4	597	0.01	0.02	1	4
6.5	597	0.01	0.00	1	0
6.6	597	0.01	0.00	1	0
6.7	598	0.00	0.01	1	1
6.8	598	0.00	0.00	1	0
6.9	599	0.00	0.00	0	0
7	600	0.00	0.01	0	1

Figures

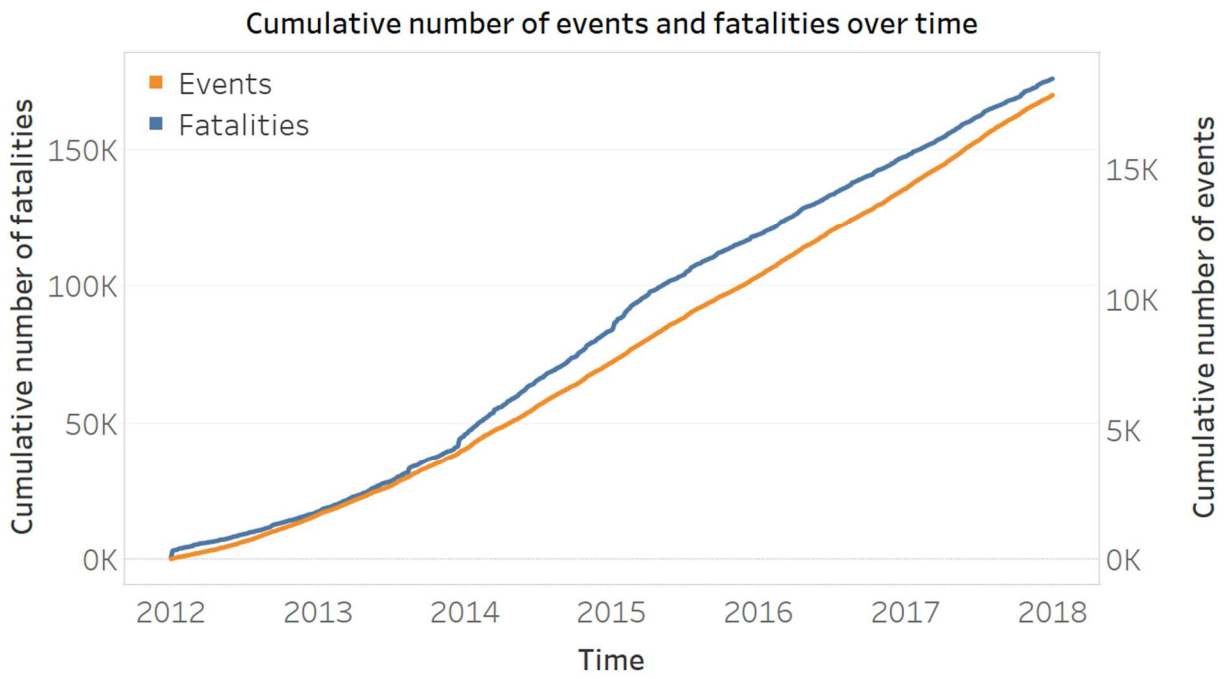


Figure 1: Cumulative number of events and fatalities over time

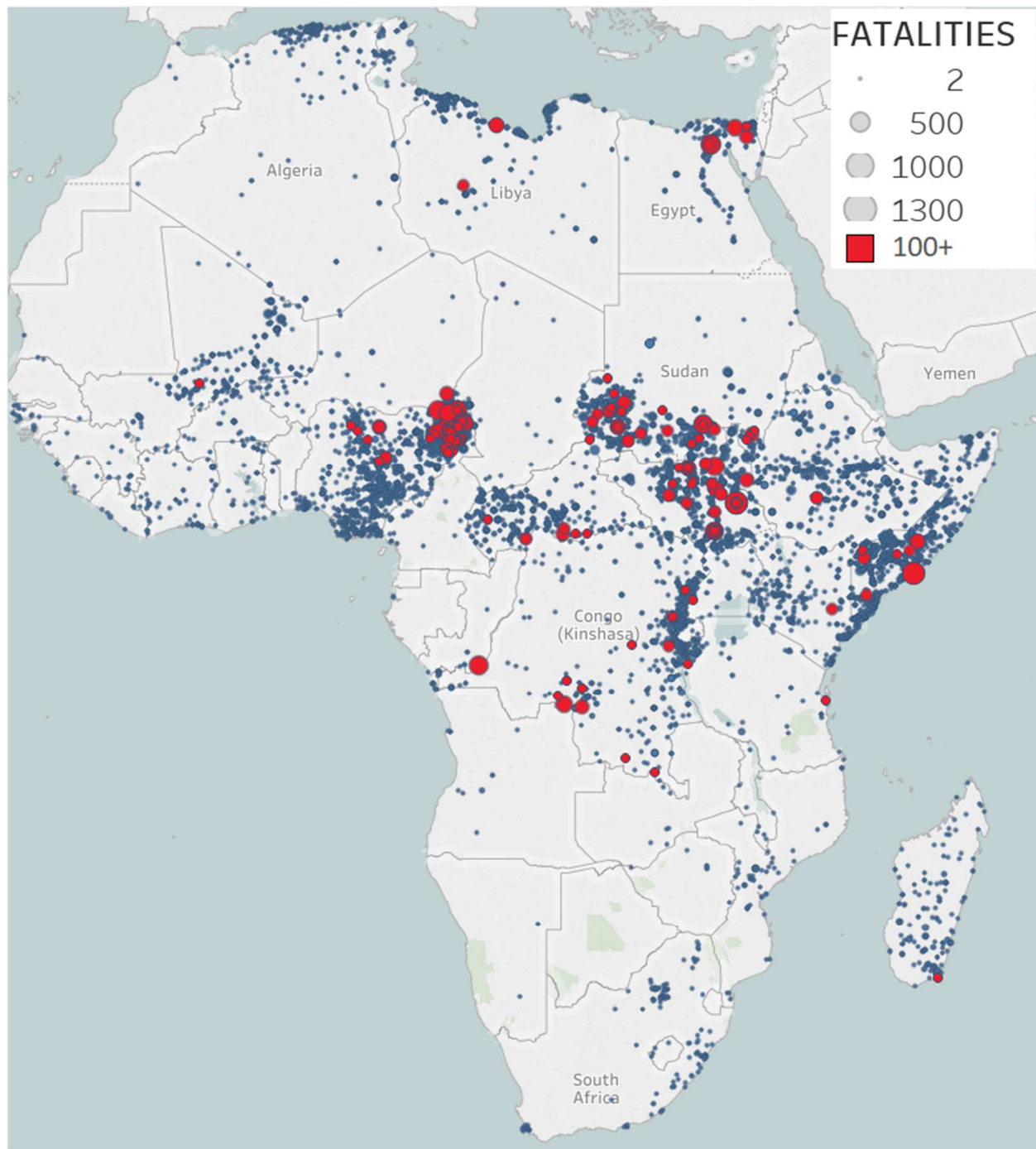


Figure 2: Map of social conflicts in Africa in the period 2012-2017 with at least two fatalities. Events with more than 100 fatalities are colored red.

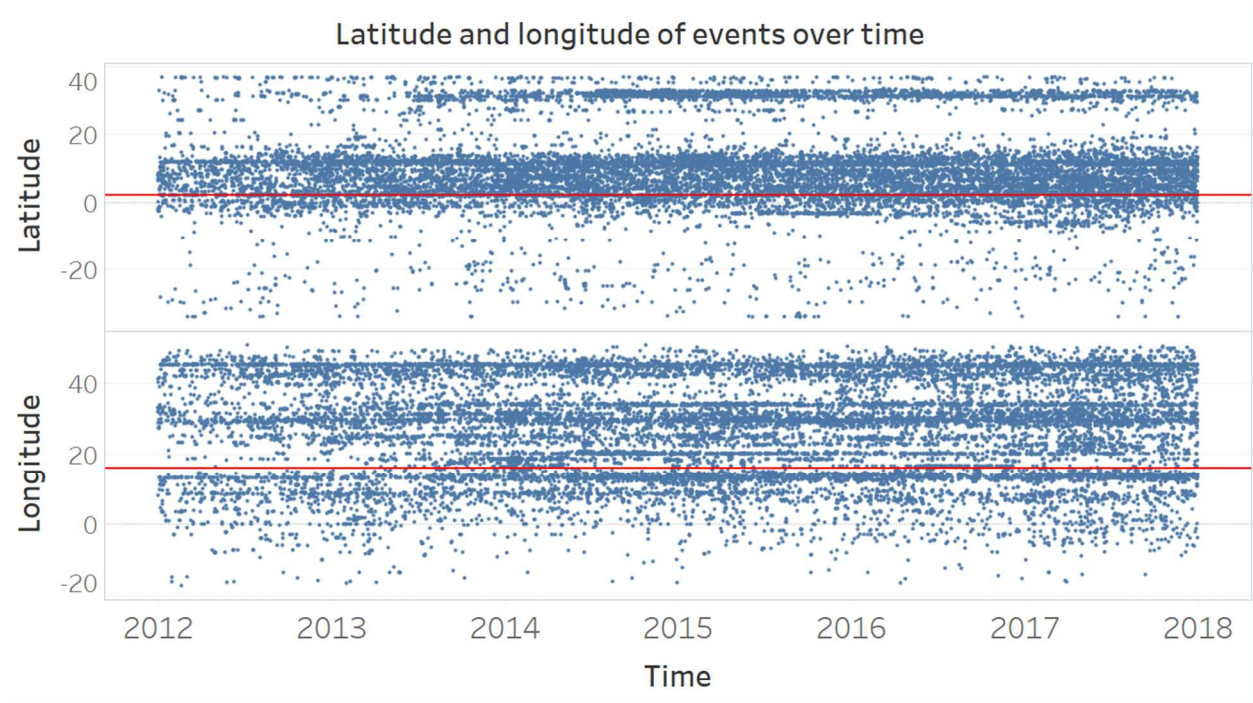


Figure 3: The latitude and longitude of events over time. The red lines represent the latitude and longitude of the geographical center of Africa. In the latitude plot, points above the red line are more north and points below are more south. In the longitude plot, points above the red line are more east and points below the red line are more west.

Distribution of number of events

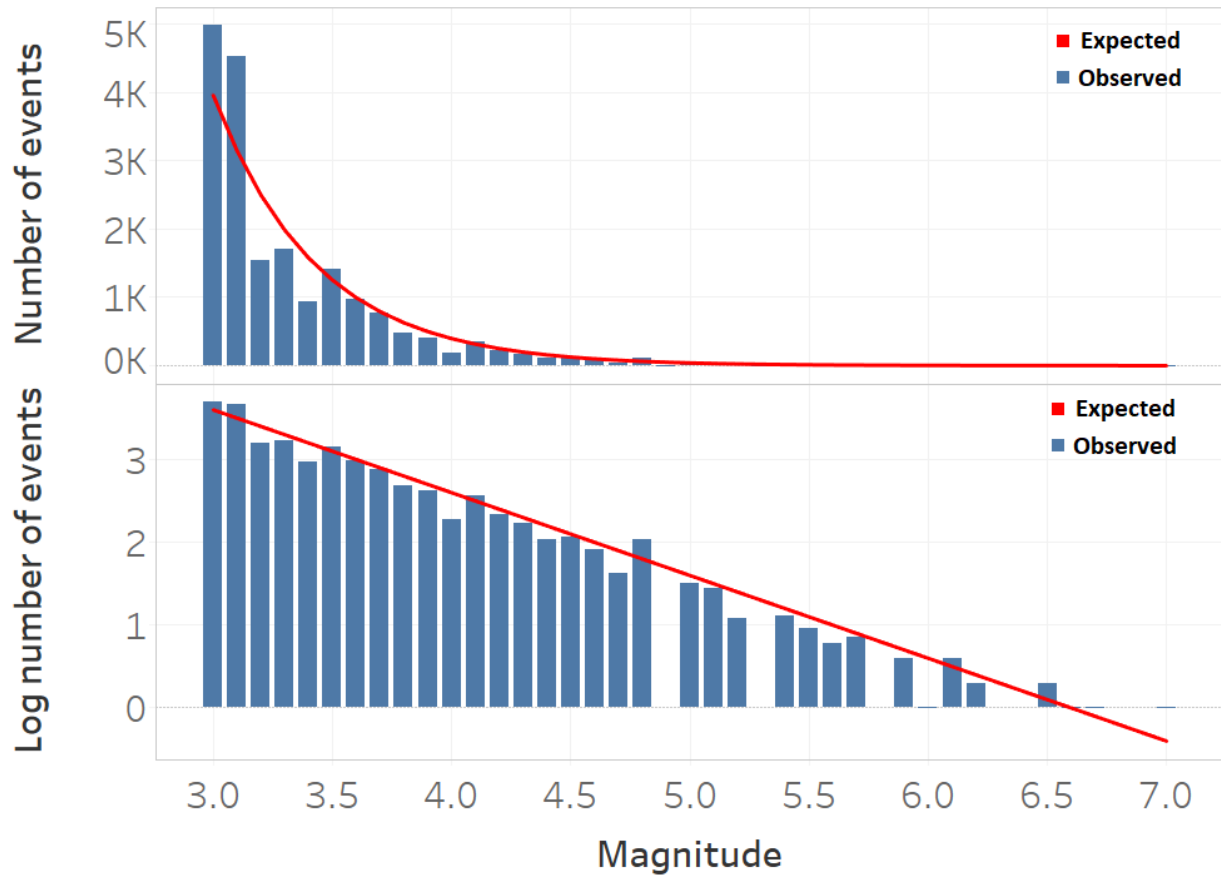


Figure 4: Distribution of the number of events. The blue bars in the top figure show the number of events for each magnitude and the red line is the number of events corresponding to the Gutenberg-Richter law. In the bottom figure, the logarithm base 10 of the same data is shown. If the magnitude distribution follows the Gutenberg-Richter law, the blue bars in the bottom figure should correspond to the red line.

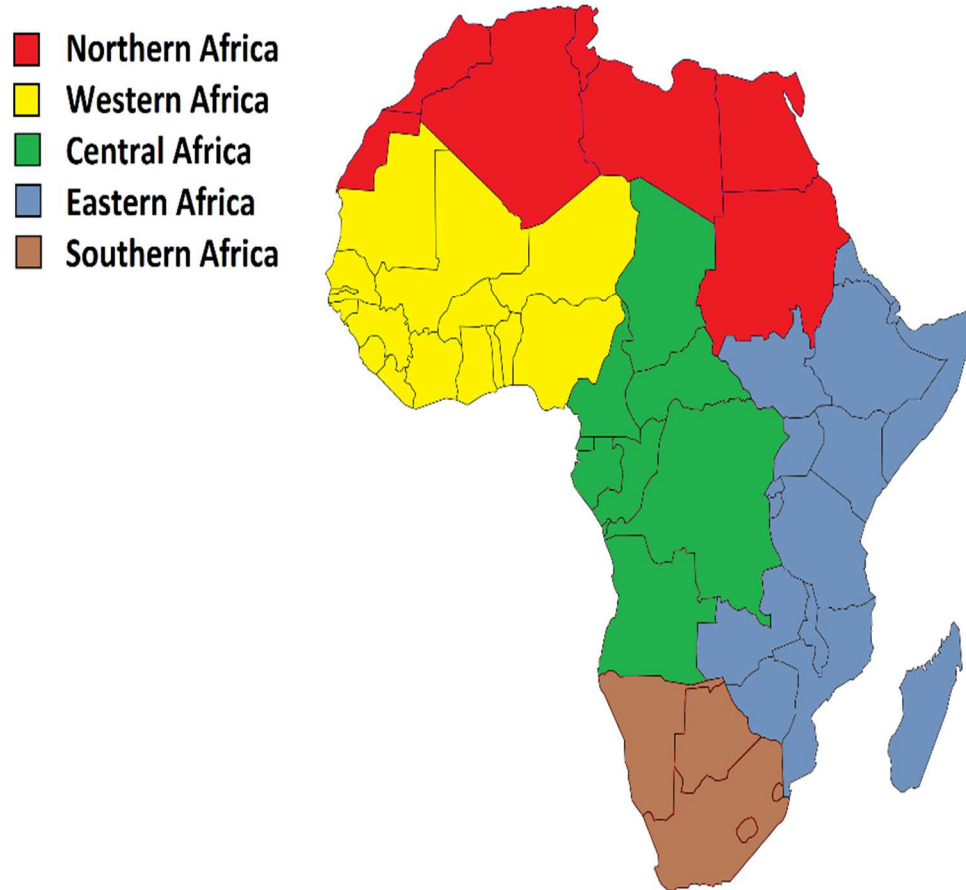


Figure 5: The five major regions of Africa: North, West, Central, East and South. Source: Wikimedia Commons, the free media repository

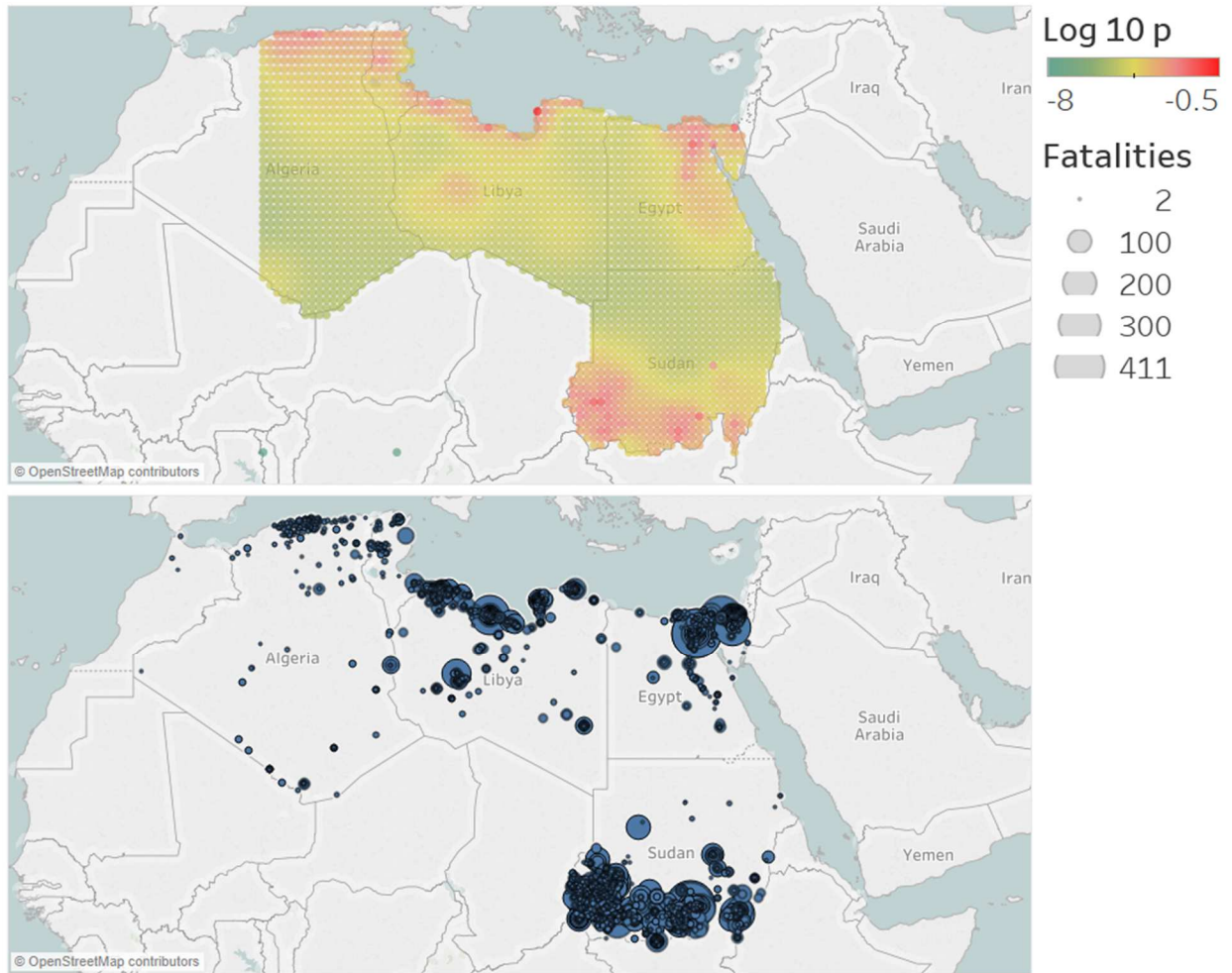


Figure 6: Spatial background distributions of the model with the largest log-likelihood for the region North Africa. The area is divided into $N_c = 4070$ cells. In these cells, the background intensity is assumed to be homogeneous. The cells give the probability of a background event occurring in that cell.

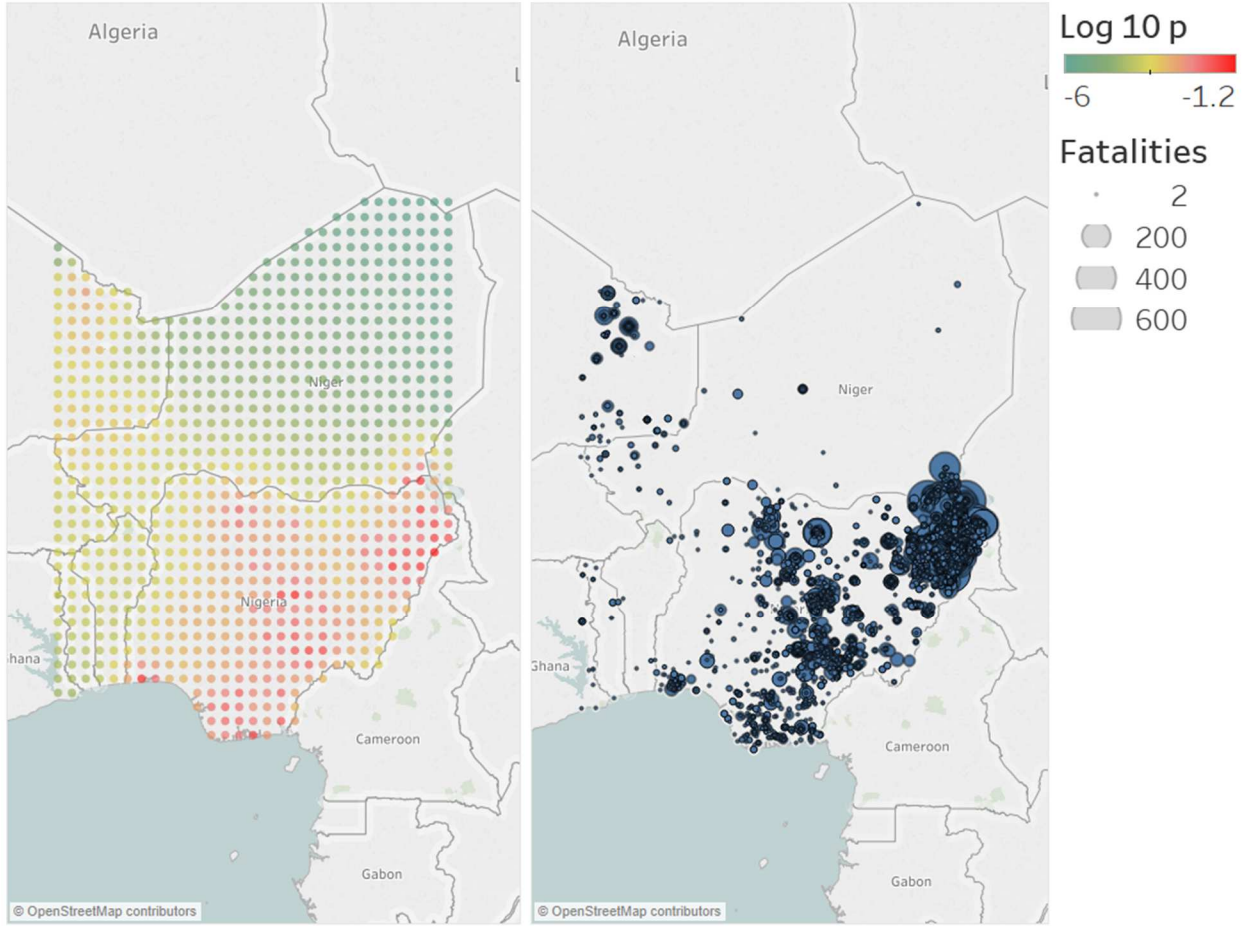


Figure 7: Spatial background distributions of the model with the largest log-likelihood for the region West Africa. The area is divided into $N_c = 1102$ cells. In these cells, the background intensity is assumed to be homogeneous. The cells give the probability of a background event occurring in that cell.

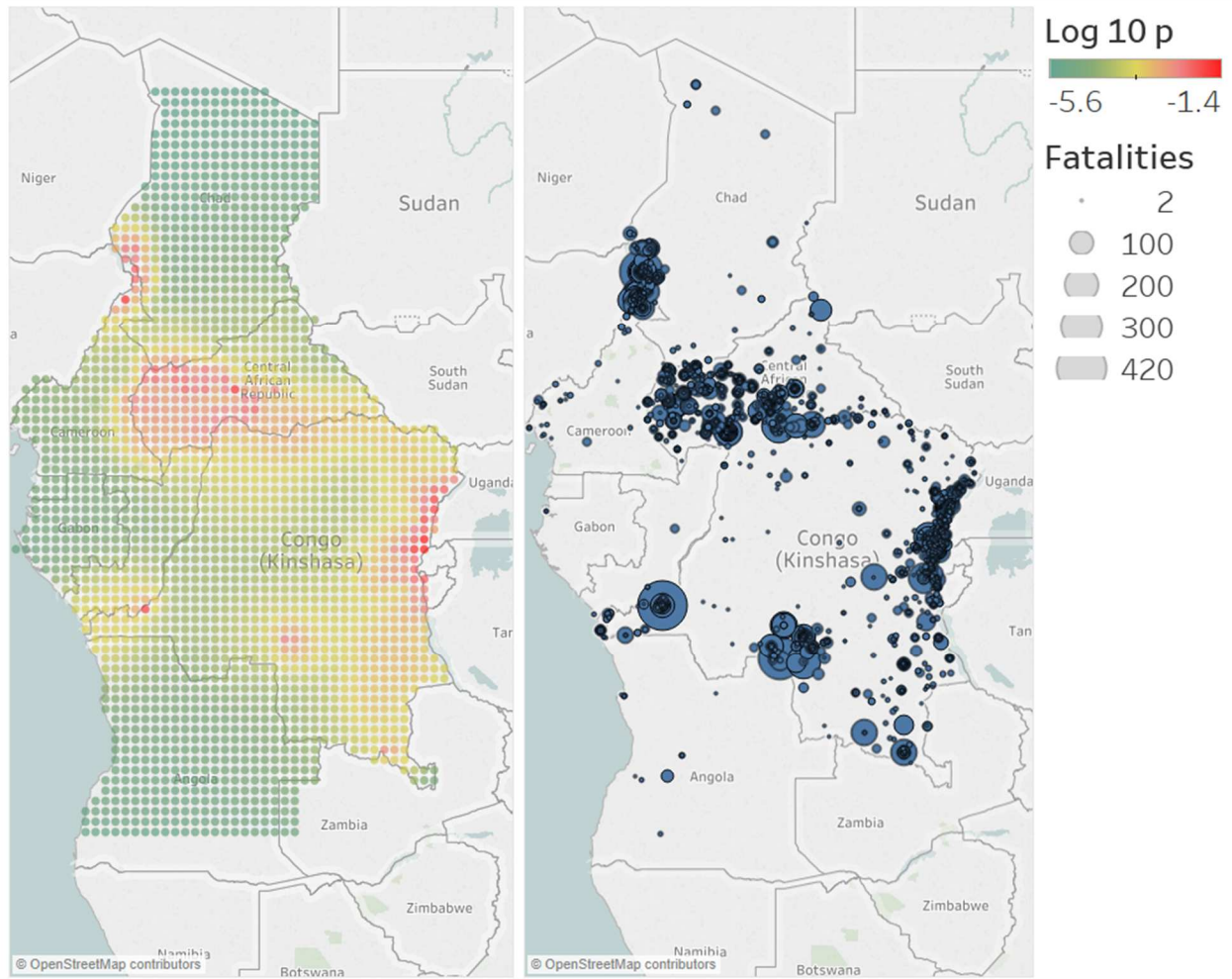


Figure 8: Spatial background distributions of the model with the largest log-likelihood for the region Central Africa. The area is divided into $N_c = 3330$ cells. In these cells, the background intensity is assumed to be homogeneous. The cells give the probability of a background event occurring in that cell.

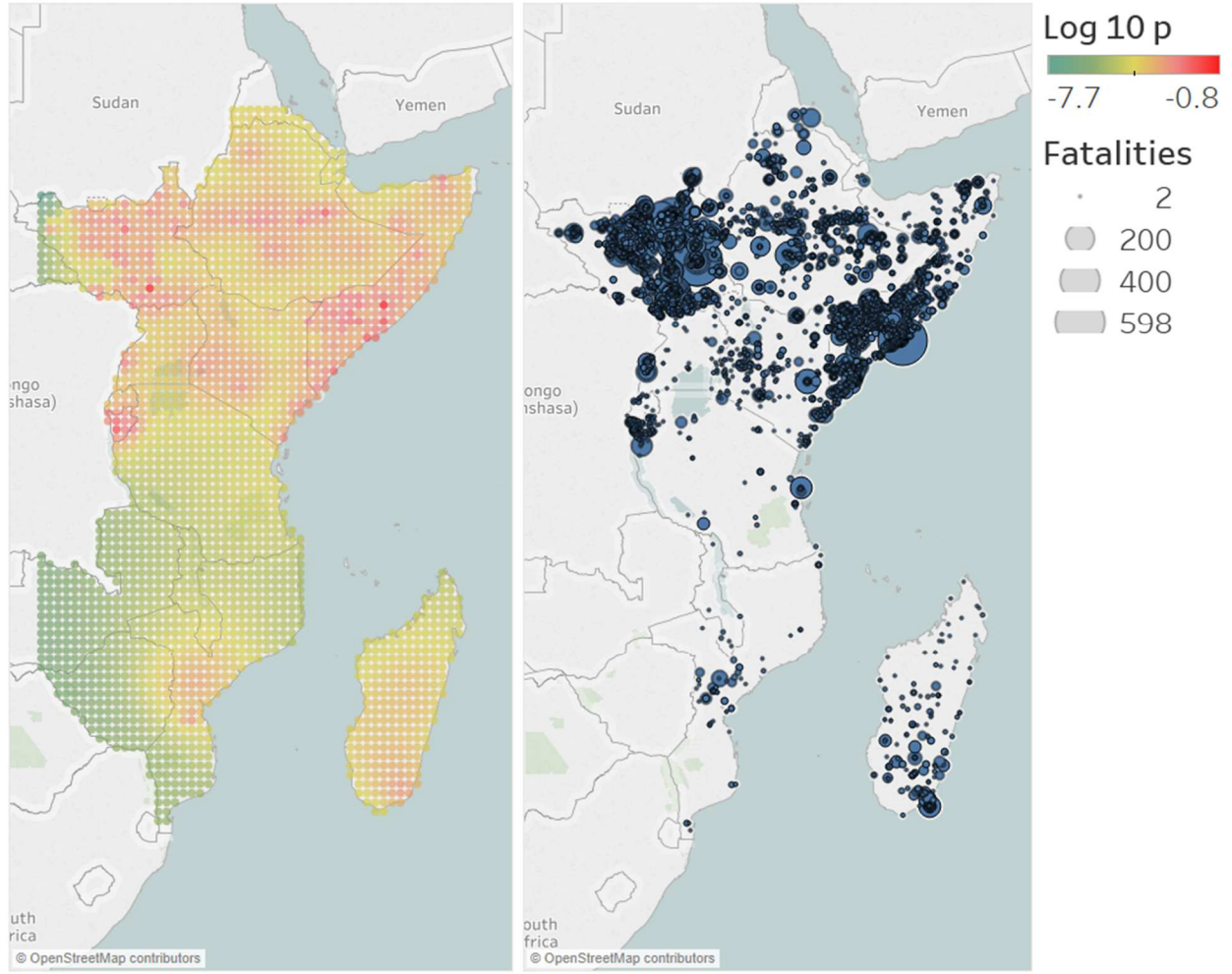


Figure 9: Spatial background distributions of the model with the largest log-likelihood for the region East Africa. The area is divided into $N_c = 4368$ cells. In these cells, the background intensity is assumed to be homogeneous. The cells give the probability of a background event occurring in that cell.

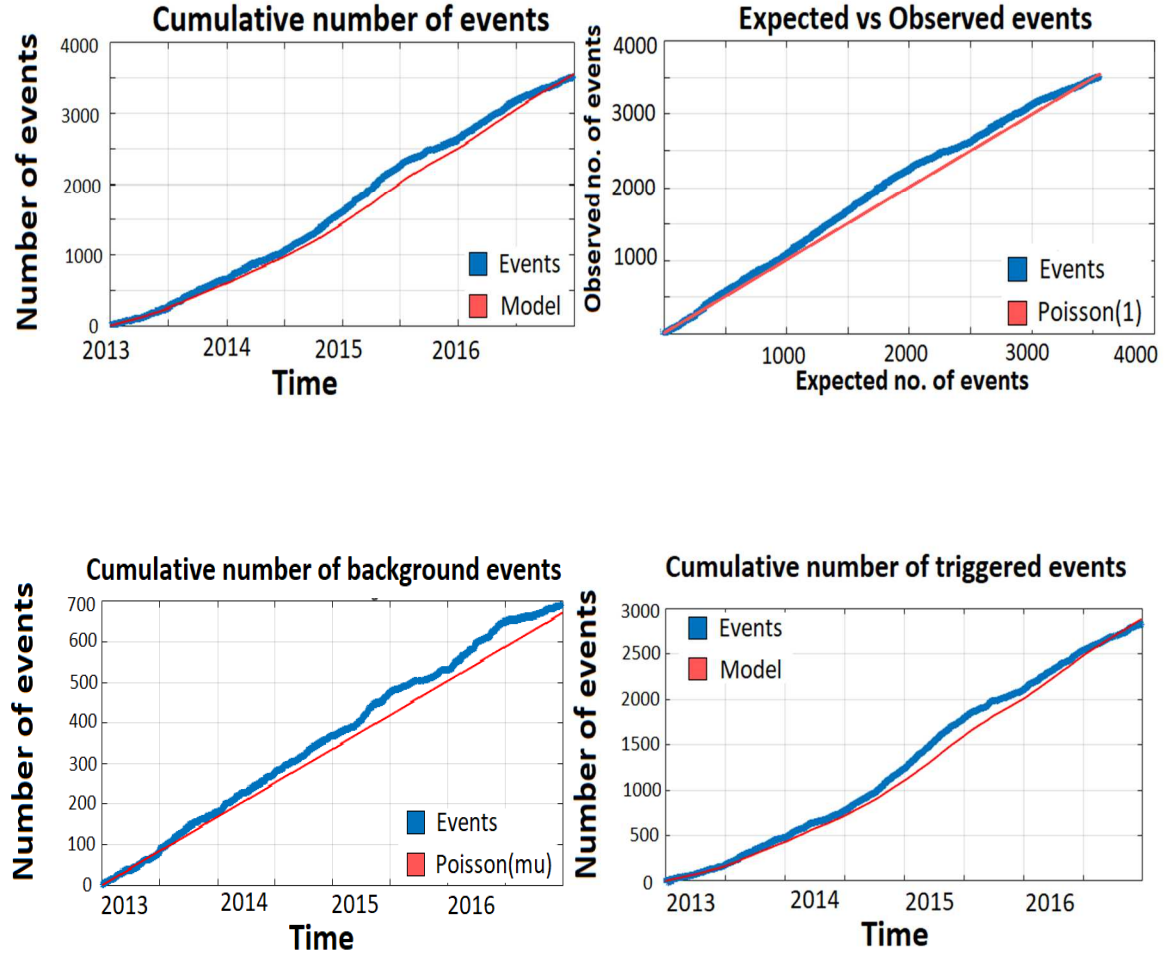


Figure 10: Results of the residual analysis for the North Africa catalog. In the top left figure the observed number of events is plotted against the transformed times from (11). In the top right figure, the observed number of events (blue) is plotted against the expected number of events by the model (red). In the bottom two figures, the observed number of background and triggered events (blue) are compared with the number of events expected by the model (red).

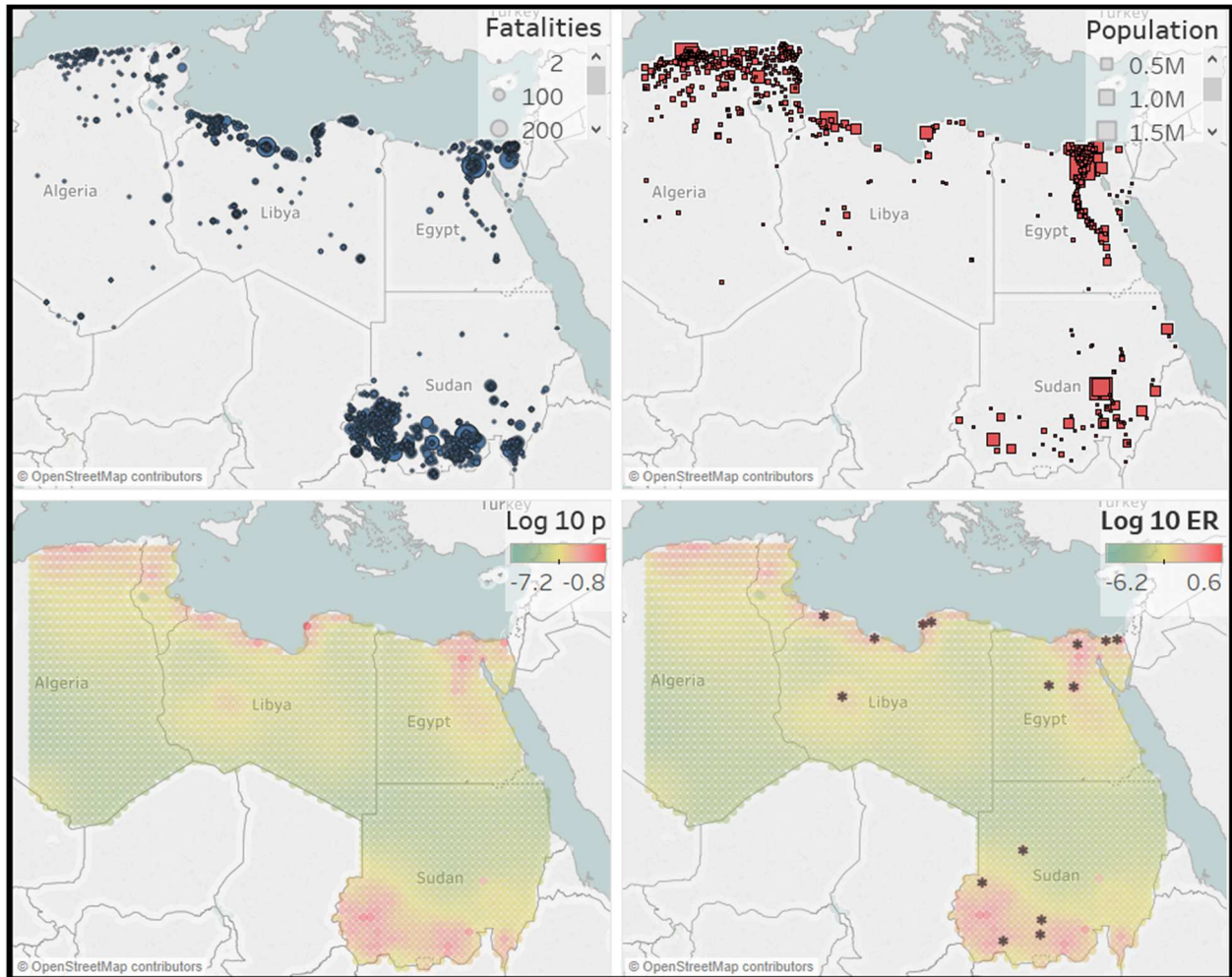


Figure 11: Forecasts for *North Africa*. Top left: event catalog 2012-2016. Top right: City map with cities with 1000+ population. Bottom left: spatial background distribution estimated from the data from 2012-2016. Bottom right: forecast for the expected number of events with $M \geq 4$ in each cell in the year 2017, together with the locations of the actual conflicts that have taken place with $M \geq 4$.

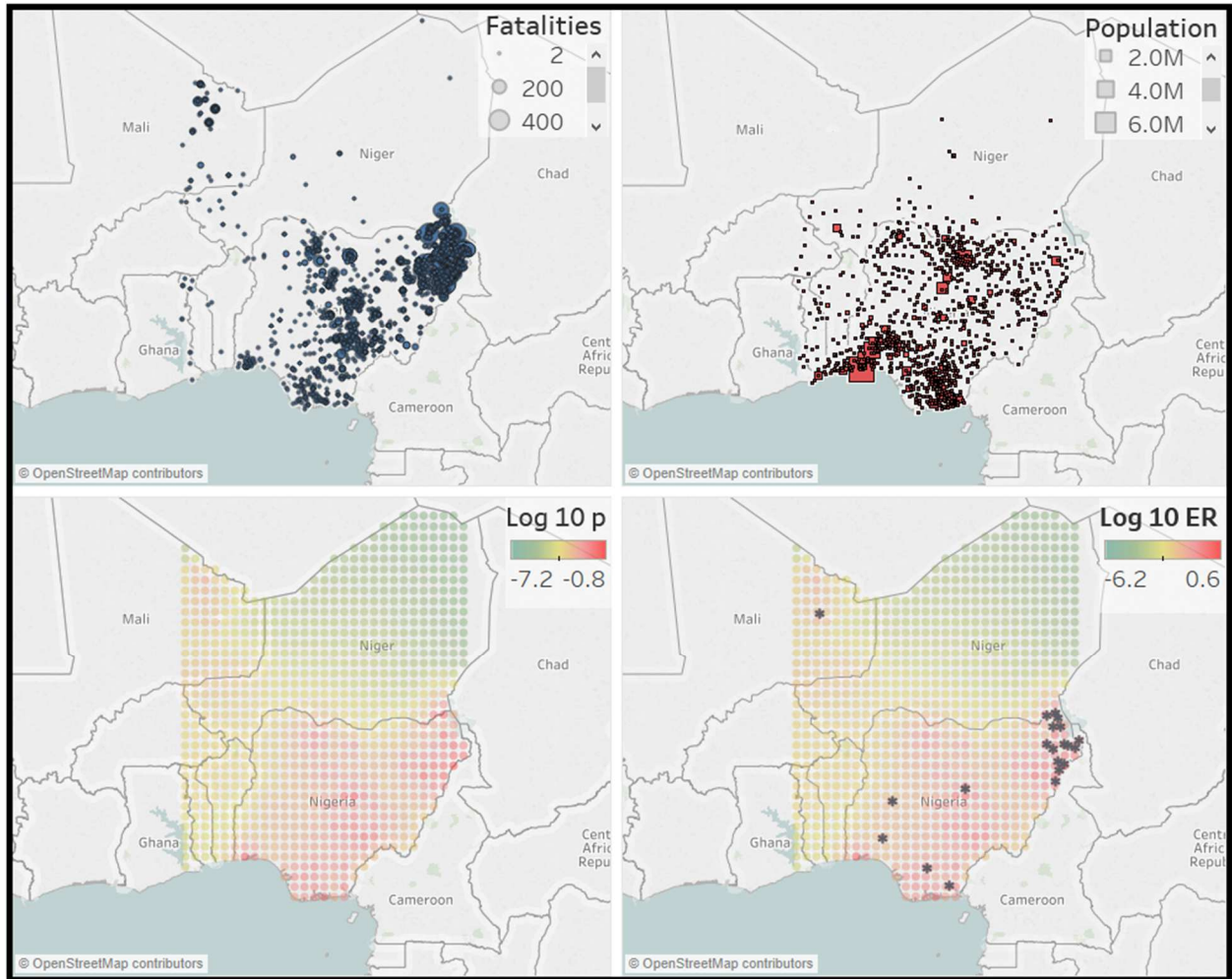


Figure 12: Forecasts for *West Africa*. Top left: event catalog 2012-2016. Top right: City map with cities with 1000+ population. Bottom left: spatial background distribution estimated from the data from 2012-2016. Bottom right: forecast for the expected number of events with $M \geq 4$ in each cell in the year 2017, together with the locations of the actual conflicts that have taken place with $M \geq 4$.

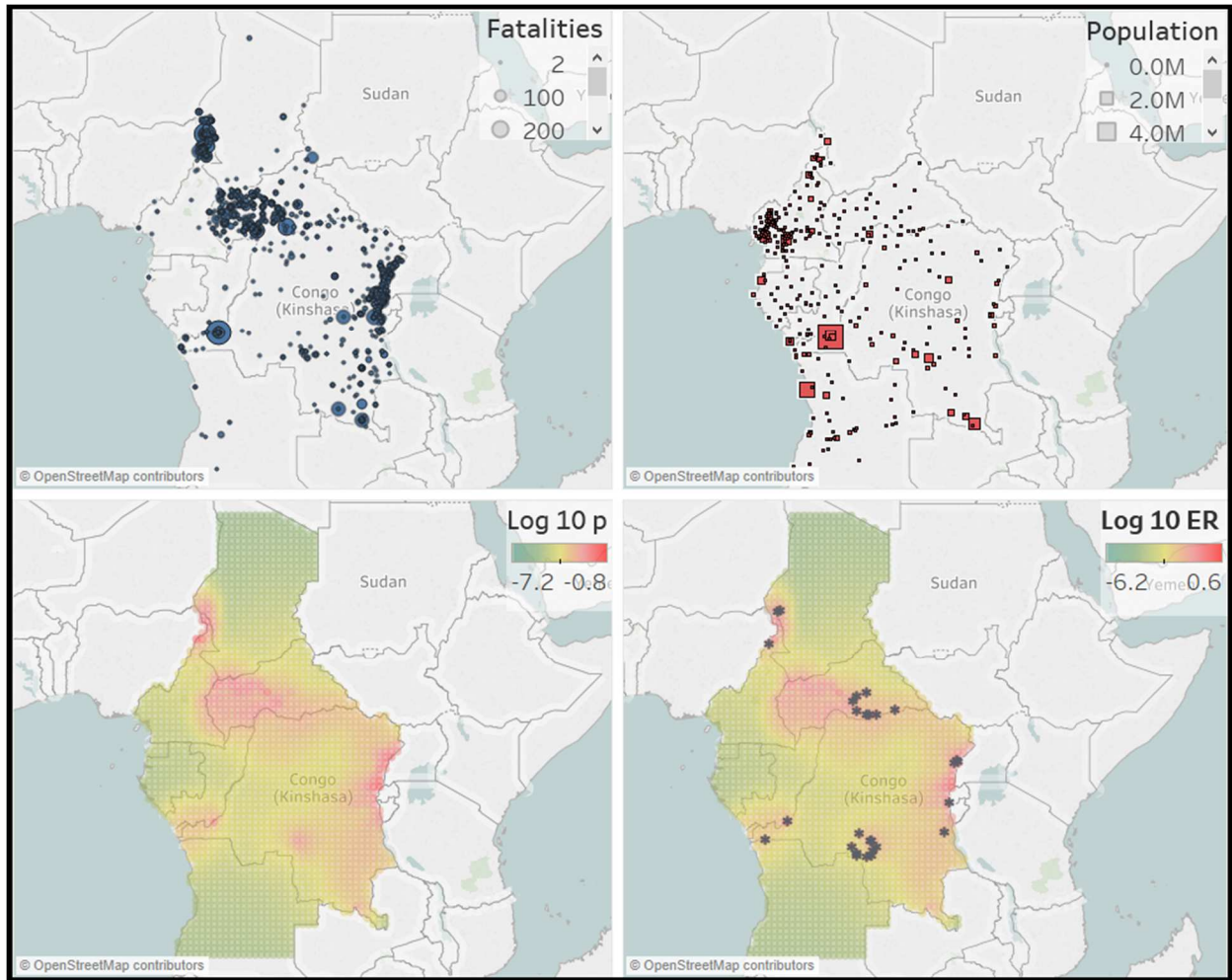


Figure 13: Forecasts for *Central Africa*. Top left: event catalog 2012-2016. Top right: City map with cities with 1000+ population. Bottom left: spatial background distribution estimated from the data from 2012-2016. Bottom right: forecast for the expected number of events with $M \geq 4$ in each cell in the year 2017, together with the locations of the actual conflicts that have taken place with $M \geq 4$.

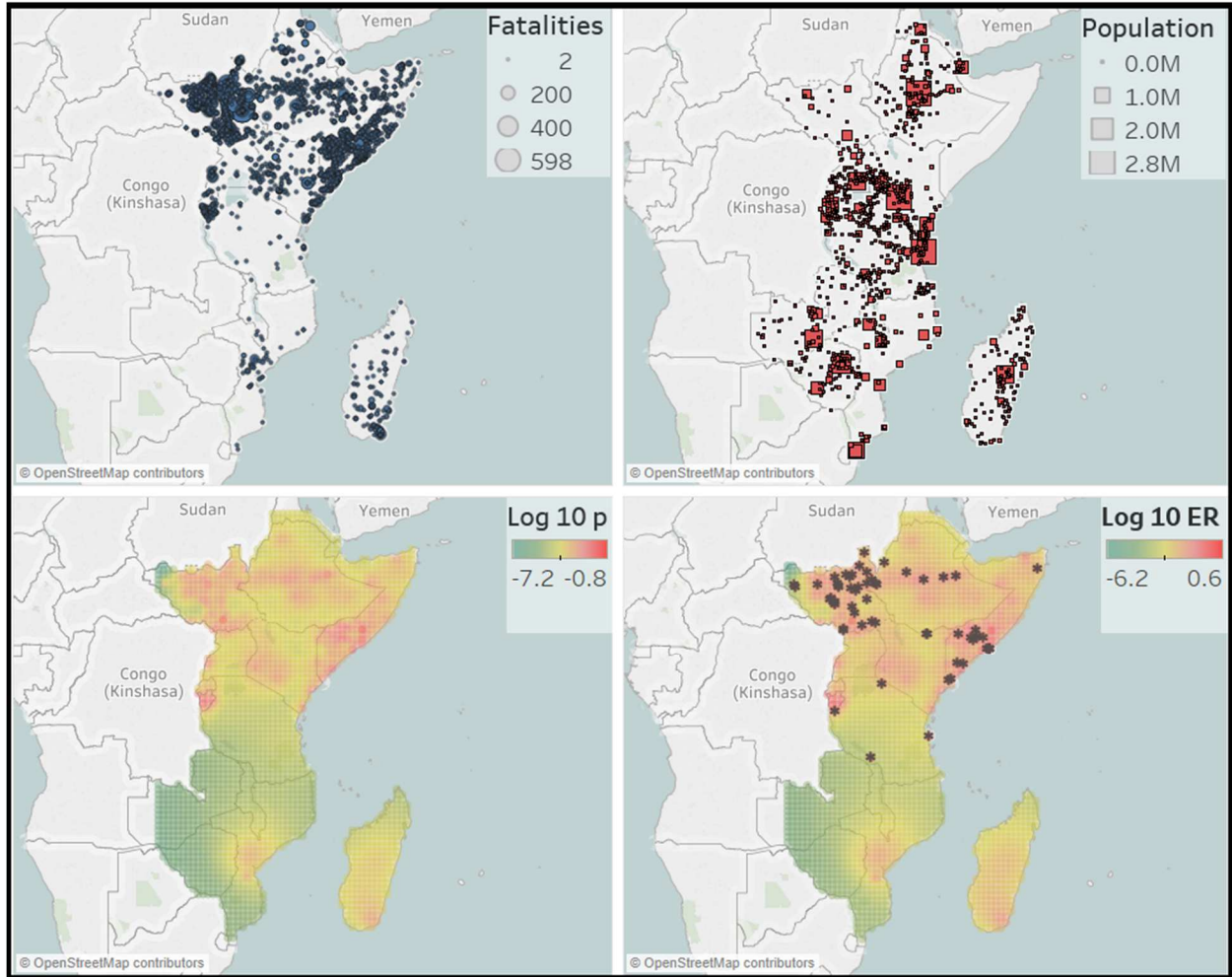


Figure 14: Forecasts for *East Africa*. Top left: event catalog 2012-2016. Top right: City map with cities with 1000+ population. Bottom left: spatial background distribution estimated from the data from 2012-2016. Bottom right: forecast for the expected number of events with $M \geq 4$ in each cell in the year 2017, together with the locations of the actual conflicts that have taken place with $M \geq 4$.

Tables

Table 1: Fatality quantiles of events in the ACLED dataset between 2012 and 2017 with at least 2 fatalities

Quantile	Fatalities	Quantile	Fatalities
0	2	0.91	20
0.1	2	0.92	22
0.2	2	0.93	25
0.3	3	0.94	28
0.4	3	0.95	31
0.5	4	0.96	38
0.6	6	0.97	47
0.7	8	0.98	58
0.8	10	0.99	96
0.9	19	1	600

Table 2: The magnitude distribution of the social conflict data. For each magnitude there is the corresponding number of fatalities, the percentage of events expected by the Gutenberg-Richter law, the observed percentage of events, the expected absolute number of events and the observed absolute number of events. The full table appears in the Appendix.

Magnitude	Fatalities	% expected	% observed	expected	observed
3	2	20.57	26.94	3565	4669
4	22	2.06	1.77	357	306
5	100	0.21	0.33	36	57
6	400	0.02	0.00	4	0
7	600	0.00	0.01	0	1

Table 3: Parameter restrictions in our ETAS model. These restrictions are based on the restrictions used for modeling earthquake occurrences and on values that are physically desirables, see Ogata (1998).

Parameter	Lower bound	Upper bound
μ	0	1
k	0.001	0.1
p	0.5	2
c	1.00E-5	0.1
α	0	2
d	0.01	1
q	1	3
γ	0	2

Table 4: Number of conflict events in each region for the learning period (2012), study period (2013-2016) and forecasting period (2017) and in total (2012-2017).

	North	West	Central	East	South	Total
N learning period	291	378	203	708	40	1620
N study period	3518	2164	1834	4617	117	12250
N forecasting period	612	617	589	1630	16	3464
N total	4421	3159	2626	6955	173	17334

Table 5: Summary statistics of the magnitudes and number of fatalities in each region. The columns are the minimum (Min.), the first quantile (1st Q.), the median (Med.), Mean, third quantile (3rd Q.) and the maximum (Max.).

	Region	Min.	1st Q.	Med.	Mean	3rd Q.	Max.
Magnitude							
	North	3.0	3.0	3.2	3.4	3.5	6.1
	West	3.0	3.1	3.2	3.4	3.7	7
	Central	3.0	3.1	3.2	3.3	3.5	6.1
	East	3.0	3.0	3.2	3.3	3.5	6.7
Fatalities							
	North	2	2	4	10	10	411
	West	2	3	5	13	12	600
	Central	2	3	4	9	10	420
	East	2	2	4	9	10	598

Table 6: Maximum likelihood estimates of the parameter b in the Gutenberg-Richter law in (1).

Region	b	Standard error
North	0.956	0.025
West	0.854	0.024
Central	0.990	0.034
East	1.006	0.022

Table 7: Parameter estimates *North Africa*. Estimation results for the eight parameters of the ETAS model, the log likelihood and the number of events for the North Africa conflict catalog. Estimates marked with a * are at, or very close to, their constraint values. The best estimate, that is, the estimate of the run with the highest log likelihood, is given for each parameter. Furthermore, the median and the 95% confidence intervals are given.

Parameter	Best value	Median	95% Confidence Interval
μ	4.60e-1	5.05e-1	{4.17e-1, 5.45e-1}
k	3.76e-2	3.58e-2	{2.57e-2, 6.13e-1}
p	7.46e-1	7.43e-1	{6.79e-1, 8.53e-1}
c	1.00e-5*	1.04e-5	{1.00e-5, 9.91e-2}
α	0.39e-01	2.48e-1	{5.21e-3, 5.37e-1}
d	1.00e-2*	1.00e-2	{1.00e-2, 1.00e-2}
q	3.00e0*	3.00e0	{2.98e00, 3.00e0}
γ	6.94e-4*	3.05e-3	{1.58e-5, 9.04e-3}
logL	1.33e3	1.31e3	{6.36e2, 1.33e3}
Events	3.56e3	3.63e3	{3.35e3, 3.75e3}

Table 8: Parameter estimates *West Africa*. Estimation results for the eight parameters of the ETAS model, the log likelihood and the number of events for the West Africa conflict catalog. Estimates marked with a * are at, or very close to, their constraint values. The best estimate, that is, the estimate of the run with the highest log likelihood, is given for each parameter. Furthermore, the median and the 95% confidence intervals are given.

Parameter	Best value	Median	95% Confidence Interval
μ	5.02e-1	5.21e-1	{4.13e-1, 5.40e-1}
k	2.26e-2	2.24e-2	{1.75e-2, 3.02e-2}
p	7.01e-1	7.22e-1	{6.61e-1, 7.52e-1}
c	1.00e-5*	1.04e-5	{1.00e-5, 1.15e-2}
α	2.46e-01	4.76e-1	{4.45e-2, 5.62e-1}
d	1.00e-2*	1.00e-2	{1.00e-2, 1.00e-2}
q	3.00e0*	3.00e0	{2.98e00, 3.00e0}
γ	1.97e-4*	2.15e-3	{2.54e-5, 6.74e-3}
LogL	-6.16e3	-6.17e3	{-6.20e3, -6.16e3}
Events	2.16e3	2.26e3	{2.04e3, 2.31e3}

Table 9: Parameter estimates *Central Africa*. Estimation results for the eight parameters of the ETAS model, the log likelihood and the number of events for the Central Africa conflict catalog. Estimates marked with a * are at, or very close to, their constraint values. The best estimate, that is, the estimate of the run with the highest log likelihood, is given for each parameter. Furthermore, the median and the 95% confidence intervals are given.

Parameter	Best value	Median	95% Confidence Interval
μ	3.79e-1	3.59e-1	{3.01e-1, 4.34e-1}
k	3.52e-2	3.80e-2	{2.02e-2, 4.29e-2}
p	7.95e-1	7.96e-1	{7.48e-1, 8.28e-1}
c	1.00e-5*	1.00e-5	{1.00e-5, 1.15e-5}
α	3.03e-1	1.86e-1	{4.49e-2, 8.27e-1}
d	1.00e-2*	1.00e-2	{1.00e-2, 1.01e-2}
q	3.00e0*	2.99e0	{2.96e0, 3.00e0}
γ	3.21e-5*	5.65e-3	{8.02e-5, 2.39e-2}
LogL	-3.13e3	-3.15e3	{-3.30e3, -3.13e3}
Events	1.89e3	1.90e3	{1.68e3, 2.00e3}

Table 10: Parameter estimates *East Africa*. Estimation results for the 8 parameters of the ETAS model, the log likelihood and the number of events for the East Africa conflict catalog. Estimates marked with a * are at, or very close to, their constraint values. The best estimate, that is, the estimate of the run with the highest log likelihood, is given for each parameter. Furthermore, the median and the 95% confidence intervals are given.

Parameter	Best value	Median	95% Confidence Interval
μ	8.38e-1	8.54e-1	{7.69e-1, 9.56e-1}
k	3.07e-2	2.96e-2	{2.25e-2, 3.88e-2}
p	7.14e-1	6.95e-1	{6.56e-1, 7.80e-1}
c	1.00e-5*	1.05e-5	{1.00e-5, 1.19e-5}
α	8.93e-2	6.13e-1	{1.09e-2, 5.29e-1}
d	1.00e-2*	1.00e-2	{1.00e-2, 1.00e-2}
q	3.00e0*	3.00e0	{2.97e0, 3.00e0}
$\gamma\backslash$	8.18e-5*	6.57e-3	{8.18e-5, 1.50e-2}
LogL	-8.39e3	-8.42e3	{-8.56e3, -8.39e3}
Events	4.59e3	4.68e3	{4.38e3, 4.88e3}

Table 11: Parameter estimates for the parameters μ , k , p and α of the ETAS model. The values are the values from the 'Best Value' columns in Tables 7 to 10. The last column gives the number of events in the study period.

Region	μ	k	p	α	Events
North	0.460	0.0376	0.746	0.139	3518
West	0.502	0.0226	0.701	0.246	2164
Central	0.379	0.0352	0.795	0.303	1834
East	0.838	0.0307	0.714	0.0893	4617

Table 12: Total number of expected and observed events, background events and triggered events for each region. The expected numbers of events are calculated using (12) to (14)

Region	All		Background		Triggered	
	Expected	Observed	Expected	Observed	Expected	Observed
North	3558	3516	671	691	2887	2824
West	2160	2163	733	705	1427	1457
Central	1889	1834	553	525	1337	1308
East	4589	4616	1223	1255	3365	3361

Table 13: Results for the Number of Events test for the different regions. The estimated ETAS model for each region is used to simulate $N_{SIM} = 100$ catalogs. The median and 95% confidence interval are calculated for the simulated catalogs, as well as the relative difference (% dif.). Furthermore, the probability (Prob.) that we observe the number of events in the event catalog or more events is in the final column.

Region	Median	Observed	% dif.	95% Confidence Interval}	Prob.
North	3101	3516	-11.8%	{2737, 3592}	0.064
West	1933	2163	-10.6%	{1774, 2107}	0.019
Central	1861	1834	+1.5%	{1652, 2076}	0.62
East	4506	4616	-2.4%	{4090, 4852}	0.34

Table 14: Forecasts for the number of events with magnitude $M \geq 4$ in 2017. The average number of these events in the period 2012-2016 is given, the number in 2016, the number predicted by the model and the observed number of events in 2017.

Region	Average	Last year	Model	Observed
North	54.2	58	16.7	17
West	61.8	39	18.3	22
Central	23.2	13	13.8	39
East	58.4	50	30.5	69

References

- Ingber, L. (2000), Adaptive Simulated Annealing (ASA): Lessons learned, *arXiv preprint cs/0001018*.
- Kumazawa, T. and Y. Ogata (2014), Nonstationary ETAS models for nonstandard earthquakes, *The Annals of Applied Statistics*, 8 (3), 825-1852.
- Lin, F.-T., C.-Y. Cao and C.-C. Hsu (1993), Applying the genetic approach to simulated annealing in solving some NP-hard problems, *IEEE Transactions on Systems, Man and Cybernetics*, 23 (6), 1752-1767.
- Lombardi, A.M. (2015), Estimation of the parameters of ETAS models by simulated annealing, *Nature Publishing Group*, 5 (February), 8417
- Lombardi, A.M. (2017), SEDA: A software package for the Statistical Earthquake Data Analysis, *Nature Publishing Group*, 7 (February), 1-16.
- Ogata, Y. (1988), Statistical models for earthquake occurrences and residual analysis for point processes, *Journal of the American Statistical Association*, 83 (March), 9-27.
- Ogata, T. (1998), Space-time point-process models for earthquake occurrences, *Annals of the Institute of Statistical Mathematics*, 50 (2), 379-402.
- Sen, M.K. and P.L.K. Stoffa (2013), *Global Optimization Methods in Geophysical Inversion*, Cambridge UK: Cambridge University Press.
- Van den Hengel, G. (2018), An epidemic type aftershock sequence model on conflict data in Africa, Unpublished MSc thesis, Econometric Institute, Erasmus School of Economics.

Zhuang, J., Y., Ogata, and D. Vere-Jones (2002), Stochastic declustering of space-time earthquake occurrences, *Journal of the American Statistical Association*, 97 (458, June), 369-380.