Predicting Conflict Zones in Kenya Using a Point Process Model

Article *in* International Journal of Data Science and Analysis · February 2024

DOI: 10.11648/ijdsa.20241001.11

Henry Athiany
Jomo Kenyatta University of Agriculture and Technology

International Journal of Data Science and Analysis
2024; 10(1): 1-10
http://www.sciencepublishinggroup.com/j/ijdsa
doi: 10.11648/ijdsa.20241001.11
ISSN: 2575-1883 (Print); ISSN: 2575-1891 (Online)

Predicting Conflict Zones in Kenya Using a Point Process Model

Carol Ogira*, Roselynn Kamau, Shallom Kamau, Bridgette Kerubo Bwoma, Bonaya Kiinywi Komora, Henry Athiany

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Email address:

carologira8@gmail.com (Carol Ogira), rozeykamau@gmail.com (Roselynn Kamau), shallomkamau@gmail.com (Shallom Kamau), bridgetkerubo@gmail.com (Bridgette Kerubo Bwoma), kinywiibonaya@gmail.com (Bonaya Kiinywi Komora), henry.athiany@jkuat.ac.ke (Henry Athiany)

*Corresponding author

To cite this article:

Carol Ogira, Roselynn Kamau, Shallom Kamau, Bridgette Kerubo Bwoma, Bonaya Kiinywi Komora et al. (2024). Predicting Conflict Zones in Kenya Using a Point Process Model. *International Journal of Data Science and Analysis*, *10*(1), 1-10. https://doi.org/10.11648/ijdsa.20241001.11

Received: January 5, 2024; **Accepted**: January 23, 2024; **Published**: February 5, 2024

Abstract: In the past decade, Kenya has continued to experience high levels of conflict, which has affected the country in various ways. This study presents a method for analyzing and predicting conflict zones in Kenya using a Point Process Model. Data utilized in the analysis was obtained from the Armed Conflict Location & Event Data (ACLED) Project and Open Data for Africa. The focus was to develop a point process model, test its predictive capability, and predict conflict zones in Kenya for a specified period of time. The study highlights the framework of the model, focusing on the intensity, and effects of covariates such as population density and spatial coordinates. Spatial data analysis was carried out using the *spatstat* package of the R Statistical Software, mapping the distribution of the conflict events and further developing the model using the Berman-Turner algorithm. Parameter estimates required for the prediction were obtained from the algorithm. For the 18-year period considered (2004 - 2021), the number of conflict events increased significantly as the election period drew near, during and after the election period. Geographically, the Central and Western parts of Kenya exhibited greater intensity of conflict events, spreading to their surroundings. The spike in the number of conflict events during the electioneering period can be explained by the political differences seen in the country which fuel violence among citizens. Furthermore, population density played a major role in the high cases of conflict as is evident from the many cases of conflict recorded in Nairobi County. Other high cases of conflict during some years in the study period were associated with counties with pastoral communities, such as Mandera. Evaluating the trend of the past conflict events, the model prediction indicates that the capital city of Kenya (Nairobi) and its environs would be more prone to conflict during elections.

Keywords: Spatial Analysis, Armed Conflict, Poisson Point Process, Berman-Turner Algorithm

1. Introduction

The number of conflict events witnessed in Kenya has been on a steady rise in the past two decades [1]. With just over 3500 violent events recorded between 1997 and December 2021, Kenya is ranked 7th according to the Armed Conflict Location and Event Data Project (ACLED) [2]. Moreover, the country has continued to experience high levels of conflict that has "destabilized a swathe of her peripheral counties" even resulting in terrorist attacks in the capital, Nairobi [3].

According to ACLED, armed conflict is the use of force by a group with a political purpose or motivation [4, 5]. Violence being part of this, remains a predominant and persistent issue that needs addressing, varying only in area and time of occurrence. While statistical models that can be used for predicting conflict in a region or country already exist, the goal of this study was to develop one that may be more accurate in prediction by incorporating both the spatial and temporal aspects. These predictions would be helpful to the government, individuals or any other interested institution in planning and allocation of resources to areas affected by conflict. For instance, further research on childhood exposure to conflict and nutritional health outcomes in Africa can benefit from the results of this and similar studies [6, 7].

For several years, statistical models of spatial point patterns have been developed and successfully applied to various academic fields [8, 9]. However, the ability to fit these models to data is essential in assessing the predictive gain associated with individual variables [10]. Despite the uncertainty involved in forecasting, we believe that making predictions of internal armed conflict has several potential advantages [11]. So far, no model can accurately predict the future, but the insights given can be useful in planning and making important decisions [12].

Studies have shown how conflict can impact the day-to-day activities of the affected areas [13]. Different regions in Kenya have experienced conflicts of varying intensities and forms, such as political conflict during the 2007-2008 post-election violence witnessed in Kenya [5, 13] or communal conflict, especially in arid and semi-arid lands (ASAL), which occurs without notice with incidents such as cattle rustling. These events have ripple effects, and the end results include loss of property and lives, displacement of families, and a decline in the country's economy [13, 14]. The loss of revenue leads to slow and inefficient service delivery and the destruction of infrastructure, which forces the government to incur the repair cost at the expense of taxpayer money. These call for measures to mitigate the conflict cases and their effects.

Avoiding conflict is the ideal outcome but even in a case where that is not achievable, an accurate prediction could go a long way in helping the country and her people be prepared, from the valuable and actionable insights on where the government can prioritize its resources to maximize on reduction of conflict. Furthermore, this would help with timely response in situations of conflict and improve intelligence among security agencies.

In this study, we developed a model to predict zones likely to be in a state of conflict over a defined period of time. We explored how to identify and predict conflict zones in Kenya using the Point Process Model. Of interest was also testing its predictive capability using an ACLED dataset and predicting conflict events and where they might occur in future.

2. Methods

2.1. Data Source

The data used for this study were from two main sources. First, data was obtained from ACLED [15]. ACLED generates its data from four major sources, that is, traditional media, reports from international institutions and NGOs, local partner data, which is the major contributor since it has less biased data, and new media, which includes social media platforms such as twitter, though the data from this source is thoroughly reviewed for quality. ACLED collects information on specific conflict event types such as battles, riots, remote violence, protests, and violence against civilians and fatalities therein. A summary of the classification of these events have been documented by Kimani, Mugo and Athiany [5]. Moreover, it also considers the unique variables of the countries, that is, the geography of the regions, population sizes, and the types of violence in the specified country/ region, providing written reports describing the events after the collected data has been cross validated by multiple coders. Data from this source was therefore of great value to the study because it is aggregated into locations, specifying the longitude and latitude coordinates, and months of the occurrence of these conflict events, thus incorporating the spatial-temporal aspect of the events. ACLED data is updated regularly on a weekly basis and has currently been updated to give access on request.

Data from ACLED can be extracted by data visualization and descriptive statistics, which include graphically displaying the number of events per month in the different areas of the country to give a pictorial overview of the occurrences. However, this alone does not exhibit the dynamic behavior of the events in space and time, therefore not stipulating statistically based predictions [16]. This influenced the necessity of employing a model that would be used to efficiently analyze the available data and use the generated output to draw predictive inferences about conflict events, for the present dataset, and even for future analyses.

For the covariate - population density, data was obtained from the Open Data for Africa website [17] that hosts data from the Kenya Population and Housing Census (KPHC), in this case, the 2019 report was utilized. This data presents the numerical distribution of enumerated population density at county and sub-county levels.

2.2. Data Preprocessing

2.2.1. Study Period and Variables

For analyses, ACLED data [4] on conflict events in Kenya was arbitrarily chosen for a period of 18 years, 1/01/2004 - 31/12/2021 (the most current data at the time of study). The dataset had 31 variables. However, only 14 variables that were of interest for the study were selected.

Table 1. *Selected Variables from the ACLED dataset.*

Variable	Description
Year	Year event occurred
Event type	The fundamental unit of observation recorded as battles, riots, or protests.
Actor1	Name of first actor, identified as rebel group, militia etc.
Inter1	A numeric code indicating the type of ACTOR1
Actor2	Name of second actor
Inter2	A numeric code indicating the type of ACTOR2
Interaction	A numeric code indicating the interaction between types of ACTOR1 and ACTOR2
Admin1	County
Location	The location where the event occurred
Latitude	The latitude of the event location
Longitude	The longitude of the event location
Notes	Additional notes
Fatalities	The integer value of fatalities that occurred
Event Date	Date of conflict, recorded in the DD-MM-YY format

Missing values were identified from the data and imputed with appropriate values. Moreover, the latitude and longitude variables were converted to numeric data type to obtain the coordinates that would be used in the analysis. A new column named "*month*" was introduced, extracting only the month from the Event date column.

2.2.2. Population Density

For covariates to be compatible with the Poisson point process model, they should be an image, a function, a window, a tessellation, or a single number. Therefore, the numerical population densities for each county were converted to a raster file. The *raster* package of R software was used to convert the numerical population densities of various regions in the country to raster file.

2.2.3. The Kenya Handshake 2018

Kenyan politics have for a long period been a major influencer of conflict events especially among people from different regions with different political views. The election process has constantly been declared unfair by losing parties and this has led to increased political violence. On 9th March 2018, the Kenyan president and opposition leader agreed to reconciliatory efforts between their various political parties which were symbolized by a handshake between them. This was to enhance peace between the different conflicting parties and to ensure economic prosperity of the country. It was therefore necessary to consider this in our analysis, and possibly see the effect of incidences associated with political events in the year 2018 and to ascertain whether the handshake served its purpose.

2.3. The Poisson Point Process

The Poisson point process is usually characterized by the Poisson distribution. In this case, the probability of finding x points in a given subset N of the observational window, with measure $\mu(N)$ follows a Poisson distribution with mean $\mu(N)$, that is,

$$P(Y = x) = \frac{(\mu(N))^{x} * e^{-(\mu(N))}}{x!} \tag{1}$$

The process is therefore termed as a *Poisson Point Process*. There are two categories of the Poisson point process, namely the Homogeneous Poisson Process and the Inhomogeneous Poisson Process. Homogeneous Poisson Point Process, also known as Complete Spatial Randomness (CSR), is described by a single parameter, the intensity, which is uniform. The points are assumed to be independent and uniformly distributed in the region or space. On the other hand, in Inhomogeneous Poisson Point Process, the points are not uniformly distributed, and the intensity of the regions is not uniform.

Assumptions of the Inhomogeneous Poisson Process

a) The point events are independent of each other

b) The intensity $\lambda(s)$ varies spatially, and so is indexed by location, s.

To identify the type of Poisson Point Process the dataset exhibits, the *Kolmogorov Smirnov Test* of CSR is used, and the appropriate conclusions are made from the outputs. The most important aspects of this model are therefore: (a) Intensity and (b) Covariate effects.

2.3.1. Intensity

Intensity is the average density of points; expected number of points per unit area. It is a measure of the frequency of the events recorded by the points. Conflict data from ACLED are aggregated in a discrete-time format, that is, the day the event occurred. We therefore let t denote the time index set. Thus, for each t, the intensity function is:

$$\lambda_t(s) = e^{W_t(s)} \tag{2}$$

where $W_t(s)$ incorporates the explanatory variables (covariates) in the model.

We tested the hypothesis whether the point process is homogeneous at 0.05 level of significance. With the results, D = 0.2619 and p < 0.0001. We then rejected the null hypothesis and concluded that the Poisson Point Process was inhomogeneous, implying that the intensity is not uniform.

2.3.2. Covariate Effects

From the point pattern dataset, the covariate data was critical for investigating whether the resultant intensity depends on the covariate. The covariates of interest for this model were the population density of Kenya and the spatial coordinates.

Population density

Including the population density as a covariate, the resulting intensity function becomes:

$$\lambda_t(s) = e^{\beta_0 + \beta_1 x_1(s)} \tag{3}$$

Where β_0 is the constant and β_1 is the coefficient of the population density of a given space, *s*.

Spatial coordinates

The spatial coordinates were obtained from the ACLED dataset, particularly the columns with the latitude and longitude information. Incorporating these spatial coordinates, the intensity function becomes:

$$\lambda_t(s) = e^{\beta_0 + \beta_1 x_1(s) + \beta_2 x_2(s)} \tag{4}$$

where:

$x_1(s)$- Latitude of the space/area, s

$x_2(s)$- Longitude of the space, s

β_0- A constant

β_1- Coefficient of the latitude

β_2- Coefficient of the longitude

2.3.3. The Null Model

In the absence of any of the covariates, the function reduces to the null model expressed as:

$$\lambda_t(s) = e^{\beta_0} \tag{5}$$

However, the null model assumes a common intensity value for all the locations, which clearly does not adequately represent the structure of the data being analyzed. The Homogeneous Poisson Process is usually taken as the appropriate null model for a point pattern [18]. Obtaining the intensity function for each point of conflict event, in a given time, would therefore help in determining the areas with potentially high risks of experiencing conflict events in the future, and those with lower risks.

2.4. Model Fitting

2.4.1. Fitting Model to the Data

The Berman-Turner Algorithm [19], developed by Mark Berman and Rolf Turner is used to find the Maximum Likelihood Estimates of the parameters, say θ. The algorithm works by finding a similarity between the Poisson log-likelihood and the log-linear Poisson regression. This algorithm is implemented in the spatstat library of the R software [20], and consists of a series of steps for computing the fitted intensity function. The function used to fit the model is called a point process model (PPM), with the syntax:

$$ppm(P, trend) \tag{6}$$

where P is the point pattern dataset, and $trend$ specifies the form of the logarithm of the intensity function. The fitted model would therefore be represented by the value returned by the ppm function. Parameter estimates, including the predictions, are also obtained from the computations and used to evaluate the spatial trend of conflict events.

2.4.2. Model Evaluation and Prediction

Once the point process model was fitted to the data, it was necessary to check whether the model fits the data well, and whether the variables used were appropriate. The fit of a parametric point process model is often assessed using a likelihood score such as the Akaike Information Criterion (AIC), which is defined as:

$$2p - 2L(\theta) \tag{7}$$

where p is the number of fitted parameters in the model.

The AIC rewards a model for higher likelihood and penalizes a model for overfitting, thus, lower AIC indicates better fit [21]. Among the two models with covariates, the model with the smaller AIC was considered and therefore used in the next step, which is prediction. Next, the function rmh of the spatstat package was used for simulation, with the syntax:

$$rmh(H) \tag{8}$$

here H is the selected model with the best fit.

This function prevised the hotspot areas during the next election year, from the trend displayed by the analyzed dataset. The output therefore indicated the regions that would be more prone to conflict and those that are less likely to experience conflict. For statistical analyses, we used the R Version 4.0.2 [20].

3. Results

The focus of this section was to examine the data for trends, presenting the data analysis results and the possible simulation of conflict events using the point process model. The analysis was carried out for data collected over a period of 18 years (1/01/2004 - 31/12/2021).

3.1. Exploratory Data Analysis

Frequency of Conflict Events across Years

Figure 1 displays the total number of conflicts recorded per year. From the figure, a noticeable spike in the number of conflicts during the election year (2007) and the year that succeeded it (2008), compared to the year preceding it (2006). For 2013 and 2017, there is just an increase in the number of conflicts in the year, but not necessarily translated to the next year. In 2007, elections were done in December, while in 2013 it was in March and 2017 it was done in August. However, we observe a steady increase in the number of conflict events beyond the year 2010 as compared to before that time.

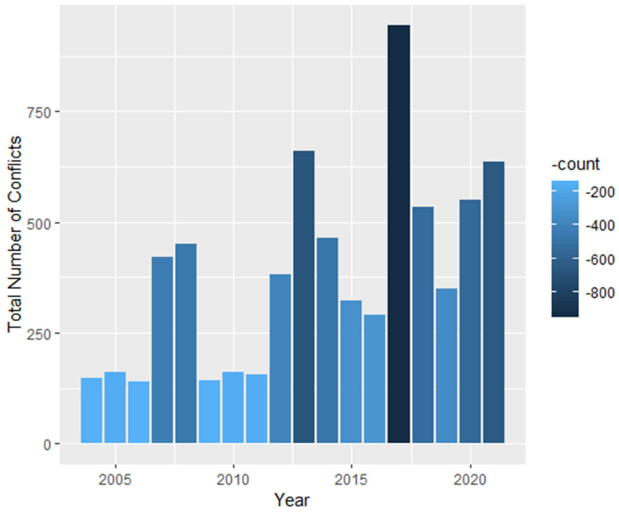

Figure 1. *Frequency Plot for Number of Conflicts Events from 2004 - 2021.*

Figure 2 illustrates the counts of events over the period of the study by top ten counties. Nairobi County had the highest number of events followed by Nakuru and Mandera.

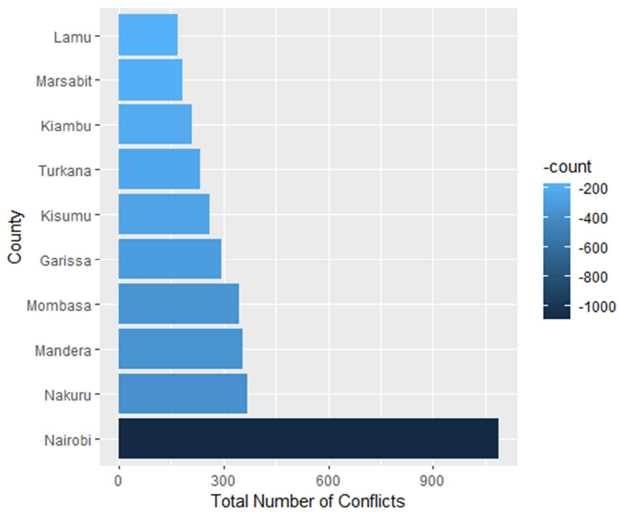

Figure 1. *Total number of conflicts for the top 10 violent counties.*

Frequency of Fatalities over Years

Results in Figure 3 summarizes the total fatalities per year for the period under study. A striking number of fatalities are noted in the periods around the years Kenya held the national elections, that is, 2007-2008, 2012-2014, and 2017.

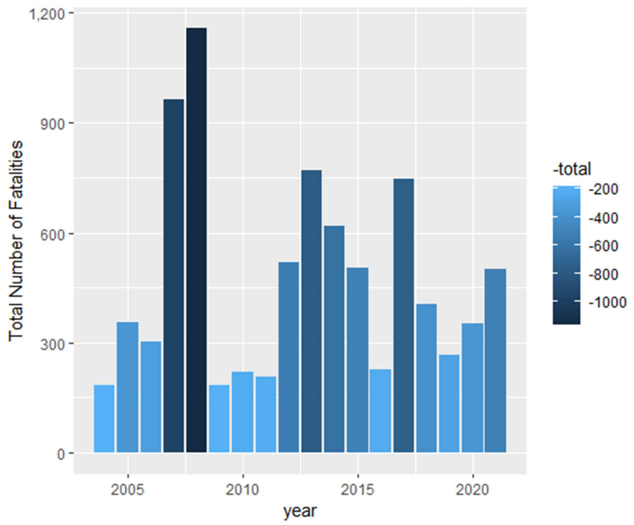

Figure 2. *Number of Fatalities due to Conflict Events from 2004-2021.*

Figure 4 presents a side-by-side comparison between the fatalities in the year 2007 (election year) and 2008 the following year. Clearly, the fatalities were more in the months of December 2007 and January 2008, just after elections. The total number of fatalities recorded in February went down drastically, then a slight increase in March 2018 before a steady decrease up to June 2018.

In august 2017, an election was carried out, presidential results nullified, and a repeat election scheduled for October the same year. A handshake took place in March 2018. The number of fatalities recorded over this period clearly indicates a decrease in the number of fatalities in 2018, as shown in Figure 5.

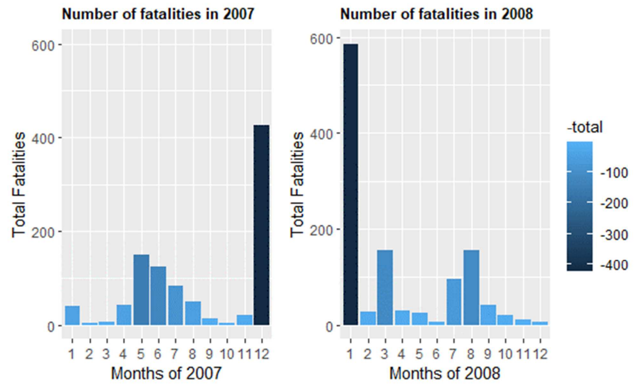

Figure 4. *Side by side comparison of total fatalities in 2007 and 2008.*

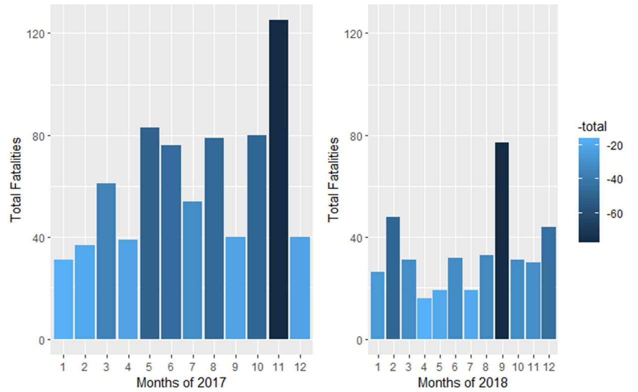

Figure 5. *Side by side comparison of total fatalities in 2017 and 2018.*

3.2. Mapping the Events

To give clearer insights, representations of the geographical coordinates of areas that experienced the conflict incidents as points, were plotted on the map of Kenya (Figure 6).

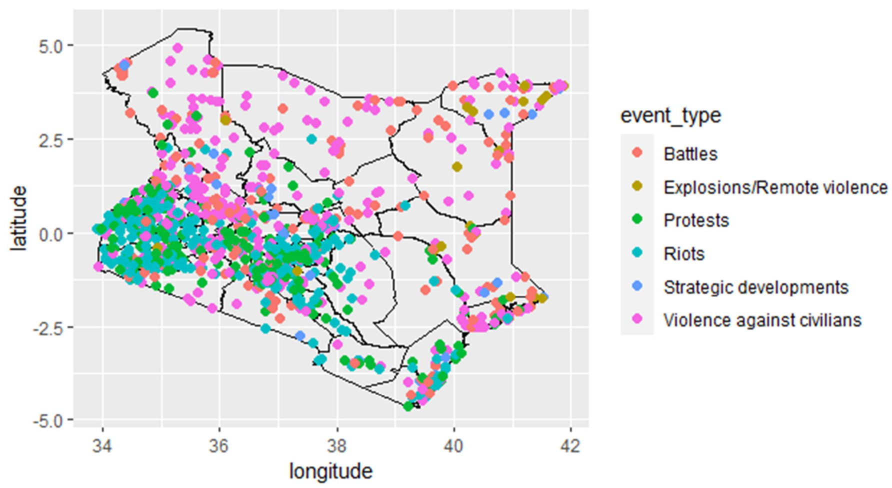

Figure 6. *Scatter plot of conflict incidents in Kenya.*

Additionally, side by side plots of conflict events before, during and after the electioneering years were generated to compare the extent of violence in those years (Figure 7, Figure 8 and Figure 9).

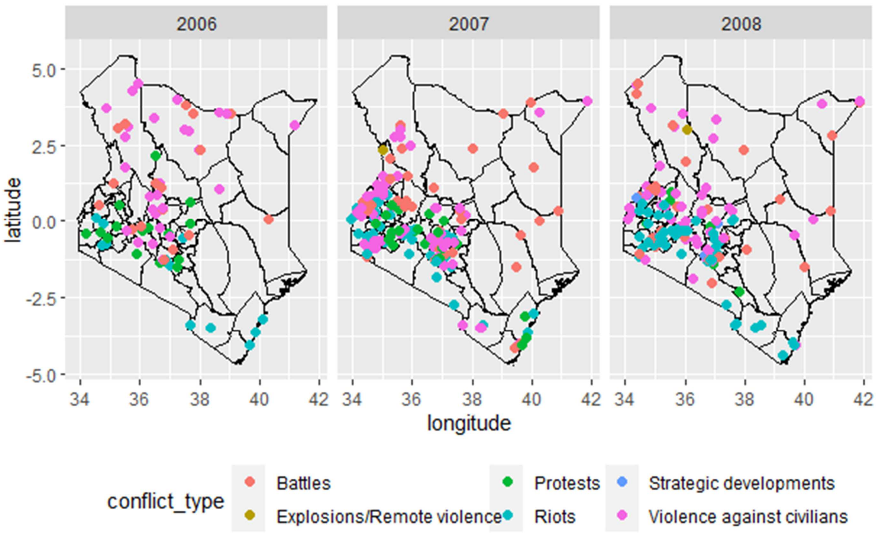

Figure 7. *Side by side comparison of Conflict Events from 2006 – 2008.*

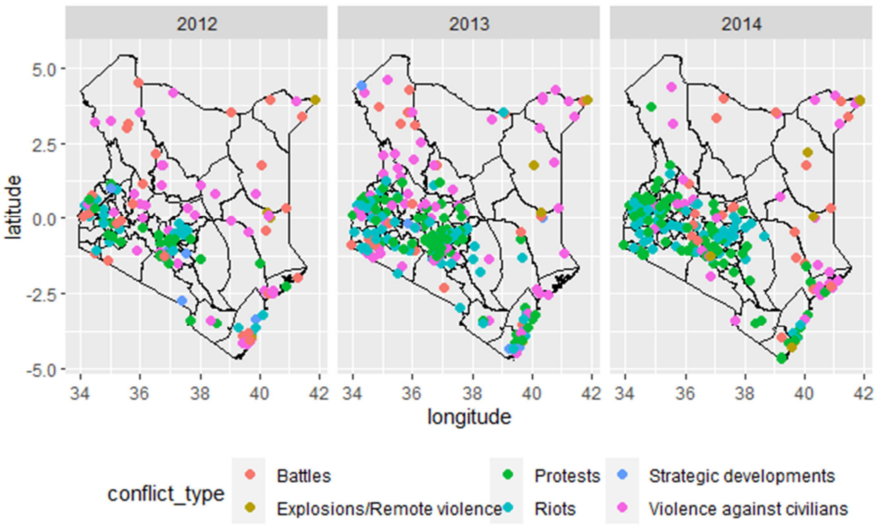

Figure 8. *Side by side comparison of Conflict Events from 2012 – 2014.*

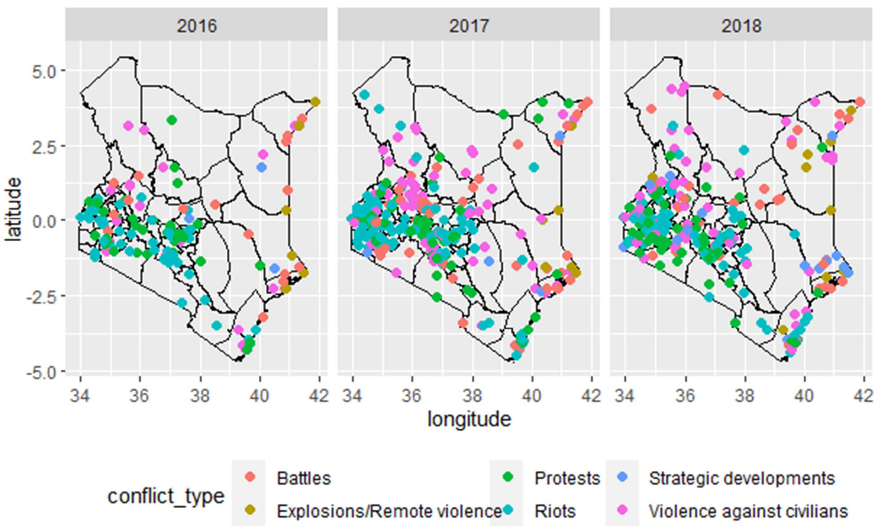

Figure 9. *Side by side comparison of Conflict Events from 2016- 2018.*

From these figures, it is evident that the Western region, part of the Central region, and the Coastal region have the most points. This implies that from the dataset, these regions experienced more conflict events compared to the rest of Kenya.

3.3. The Point Process Model

The map of Kenya was converted into an observational window, and the conflict incidents converted into planar points. The different shapes are the marks, representing the different event types and their distribution throughout the country. This is illustrated in Figure 9.

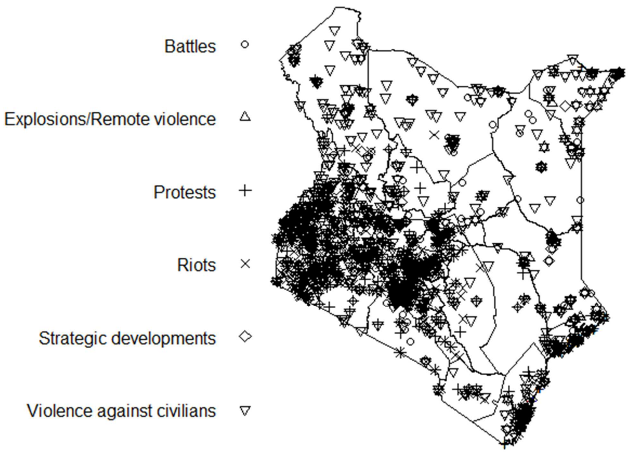

Figure 10. *The point pattern of conflict events on the map of Kenya.*

Finally, a plot the Kernel density of the conflict events showing the number of points per unit area was generated as shown in Figure 11.

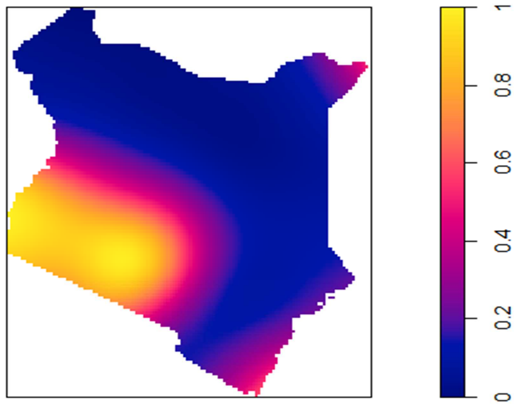

Figure 11. *Empirical density of the conflict incidents.*

Model 1: The null model

Using the data as it is, without including any covariates, with the syntax of the model being $ppm(P, trend)$, the trend for the null model is 1. This gives an intensity value of 142.4537 (95% CI 139.1185 – 145.8690), obtained from the equation as

$$\lambda(s) = exp(4.9590) \tag{9}$$

Model 2: Population density as a covariate

Using population density as a covariate to build the model, the following results were obtained for intensity as summarized in Table 2.

Table 2. *Parameter estimates for the Population Density Covariate Model.*

	Estimate	S.E	CI95.lo	CI95.hi	Z value
Intercept	4.6333	0.01397	4.6060	4.6607	331.5844
Pop-density	0.00086	0.000005	0.00085	0.00087	182.2476

The resulting intensity function was:

$$\lambda(s) = e^{4.6333+0.00086X_1(s)} \tag{10}$$

where $X_1(s)$ represents the population density of a given location.

Model 3: Spatial coordinates as predictors (spatial trend along the coordinates)

Using the spatial coordinates from the longitudinal and latitudinal data as predictor variables, intensity value was obtained as shown in Table 3.

Table 3. *Parameter estimates for the Spatial Coordinates as Predictors.*

	Estimate	S.E	CI95.lo	CI95.hi	Z value
Intercept	15.5439	0.2606	15.0332	16.0546	59.6551
X	-0.2817	0.00703	-0.2955	-0.2679	-40.0652
Y	-0.2527	0.00610	-0.2647	-0.2407	-41.4259

The resulting intensity function was:

$$\lambda(s) = e^{15.5439 \pm 0.2817x \pm 0.2527y} \tag{11}$$

where x and y represent the longitude and latitude respectively.

Model Selection

To select the best model from the three formulated, the Akaike Information Criterion (AIC) was used and the summary of the model parameters together with the AIC value of each model is given in the table below:

Table 4. *Summary of the three models with their AIC values.*

Model Intercept	Parameter Estimates (95% CI)	AIC
Model 1	4.9590 (4.9353; 4.9827)	-54189.03
Model 2 (Pop density)	4.6333 (4.6060; 4.6607) 0.00086 (0.00085; 0.00087)	-67216.58
Model 3 (Longitude(x) Latitude(y))	15.5439 (15.0332; 16.0546) -0.2817 (-0.2955; -0.2679) -0.2527 (-0.2647; -0.2407)	-57053.40

From Table 4, the second model is the best of the three models, where population density is used as the predictor variable.

3.4. Prediction

Here, the second model, which was the best among the three models was used to simulate density of conflict incidents. The selected model is used to simulate the intensity of the incidents, using the rmh function of the spatstat package. The function extracts the model information and evaluates the trend, giving the output as shown in Figure 12. The simulated density gives a future projection of the regions in conflict during the next election year, in relation to our present data. The output suggests that the South-Central and Western parts of Kenya have greater intensity spreading to their surroundings. This implies that these regions are more prone to conflict as compared to the regions in blue which have lesser intensity.

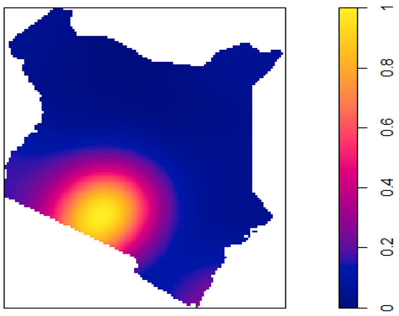

Figure 12. *Simulated density of conflict incidents.*

Finally, a side-by-side comparison of the simulated density and the actual density of conflict incidents during the election year 2022 was plotted as shown in Figure 13.

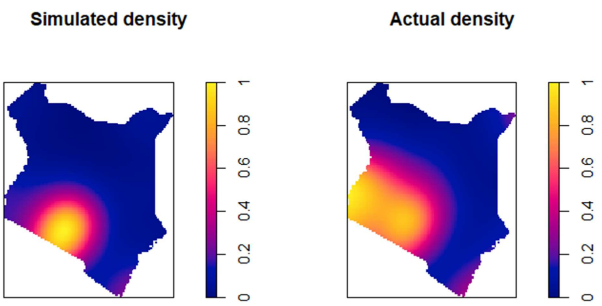

Figure 13. *Side-by-side comparison of the simulated and actual density of conflict incidents in 2022.*

4. Discussion

This study sought to analyze and predict the incidence of armed conflict in Kenya over the period of 2004 to 2021 using data collected from two sources. From the exploratory results, armed conflict recorded in the years or months just before the elections have been shown to increase and in particular regions such as Nairobi, Nakuru among others. These results are quite similar to the results presented by Kimani, Mugo and Athiany in an exploratory analysis carried out in 2020 [5]. As per ACLED data [15], the number of armed conflict incidents in Kenya has a non-linear trend, with distinct peaks observed in the years that the general elections were held [5].

Previously, studies on conflict have reflected on several geographical and environmental conditions contributing to conflict [10]. Some of the variables that have been attributed to conflict include economic growth, distance to capital and border [10, 22], land and resources and ethnicity [1, 13]. This study mainly focused on identifying the regions that are more prone to conflict in Kenya. Assessing the influence of population density on conflict incidence, it was noted that the number of people living in an area largely contributes to the number of conflict events in the area.

Guided by the insights from the general trends observed in the frequency of the incidents, major focus was put on the electioneering years. From the various exploratory visualizations generated, fatalities due to conflict largely increased during election periods. This was largely attributed to the disputes and claims of the unfairness of the election results and therefore regions supporting different parties were almost in constant conflict [1]. The recorded fatalities in the 2007-2008 post-election period were at an average of 224 people in the subsequent three months. This proved to be hugely reduced after the handshake, where the fatalities due to politically related conflict incidents reduced to 23 on average within a timeline of the next three months. This emphasizes the positive impact that the political handshake had on reducing the number of conflict incidents, particularly its effect in terms of fatalities.

Additionally, regions with high population density were also seen to be in continuous conflict which could have been attributed to the fight for the few available resources [1].

From the covariates added to the point process model in this study, population density was the most important predictor for intensity of conflict incidents. Using this model to simulate potential hotspot areas during election periods, the South-Central region, which includes the capital city, had the highest intensity, which gradually spread towards the Rift and Western regions.

The actual conditional intensity in the 2022 election period did not show much variation from the results of the simulation. A direct comparison between the simulated and actual intensity immediately confirms the high intensity around the capital city as it spread towards the Western areas of the country. However, comparing the actual intensity to both the empirical and simulated intensity, the intensity of reported conflict incidents in the Western regions greatly increased. This can be attributed to the spike of reported rioting and protest activity across the region, which was due to the increased cost of living in the just elected regime [23].

The findings of this study demonstrate an optimistic feasibility of using the Poisson point process modelling approach to examine empirical trends from spatial conflict data and generate predictions. Moreover, a point pattern records the number of events in a given area, with the location depending on the underlying spatial process, modelled using an intensity function represented by equation (2). Other approaches such as the log-Cox point process may also be considered for modelling similar events, or the stochastic partial differential equation (SPDE) models for predictive analysis of the ACLED data. Such models are able to analyze the point level data much better with the continuous Gaussian field represented as discrete indexed Gaussian Random Markov Field (GRMF) [24].

This study, however, had some limitations in terms of the quality of data. The frequency of national population census in Kenya is usually 10 years, which limited the availability of credible annual population density for more informative insights. Additionally, ACLED's reliance on media sources for data, which portrays the risk of media bias and verifiability of some of the reports collected. Secondly, estimation of point process models and validity of their inferences rely on geolocation accuracy, and covariate scale [25]. The availability of other potential important covariates of conflict incidents in Kenya was limited in the context of this study. This limited the covariates added to the model in this study, and consequently limited the precision of the simulated conditional intensity.

5. Conclusions

From the empirical results, it can be concluded that as the election period drew near, the country experienced an increase in the number of armed conflict incidents. Population density played a big role in the number of conflict cases. For instance, Nairobi recorded many cases of conflict over the years, and it could be attributed to the high population density in the county. Additionally, in some of the years, the high cases of non-political conflict were associated with counties that have pastoral communities such as Mandera. Based on the past trend of events as shown in the study, the model prediction indicates that the capital city of Kenya (Nairobi) and its environs would be more prone to conflict during elections.

Contributions

CO, RK, SK, BKB, BKK, HA, conceived and designed the study; CO, RK, SK, BKB, BKK conducted the analysis, interpreted the results and drafted the manuscript under the supervision of HA; CO, RK, SK, BKB, BKK, HA critically revised and discussed the manuscript for its content and approved its final version for publication.

ORCID

0000-0003-2749-5157 (Henry Athiany)

Acknowledgments

Data used for this research was provided by the Armed Conflict Location & Event Data (ACLED) Project (https://acleddata.com) and Open Data for Africa (https://africaopendata.org). We are also grateful for the valuable comments received from the SUSAN 2021 (Kenya) and IBC 2022 (Latvia) conferences.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] B. Rohwerder, "Conflict analysis of Kenya," *Birmingham, UK: GSDRC, University of Birmingham*, pp. 1–33, 2015.

[2] C. Dowd and C. Raleigh, "Regional Report : Kenya December 2013," *ETH Zürich*, vol. 2013, no. December, pp. 1–7, 2013.

[3] J. Lind, P. Mutahi, and M. Oosterom, "Tangled ties: Al-Shabaab and political volatility in Kenya," 2015. [Online]. Available: https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/6018/ER130_TangledTiesAlShabaabandPoliticalVolatilityinKenya.pdf

[4] ACLED, "Armed Conflict Location & Event Data Project (ACLED) Codebook," *Armed Conflict Location & Event Data Project*, pp. 1–36, 2019, [Online]. Available: https://acleddata.com/resources/general-guides/

[5] P. Kimani, C. Mugo, and H. Athiany, "Trends of Armed Conflict in Kenya from 1997 to 2021: An Exploratory Data Analysis," *International Journal of Data Science and Analysis*, vol. 7, no. 6, p. 161, 2021, doi: 10.11648/j.ijdsa.20210706.14.

[6] O. A. Makinde, E. Olamijuwon, I. Mgbachi, and R. Sato, "Childhood exposure to armed conflict and nutritional health outcomes in Nigeria," *Confl Health*, vol. 17, no. 1, pp. 1–12, 2023, doi: 10.1186/s13031-023-00513-0.

[7] D. K. Kinyoki *et al.*, "Conflict in Somalia: impact on child undernutrition," *BMJ Glob Health*, vol. 2, no. 2, 2017, doi: 10.1136/bmjgh-2016-000262.

[8] E. Velázquez, I. Martínez, S. Getzin, K. A. Moloney, and T. Wiegand, "An evaluation of the state of spatial point pattern analysis in ecology," *Ecography*, vol. 39, no. 11, pp. 1042–1055, Nov. 2016, doi: 10.1111/ECOG.01579.

[9] J. Lancaster and B. J. Downes, "Spatial point pattern analysis of available and exploited resources," *Ecography*, vol. 27, no. 1, pp. 94–102, 2004, doi: 10.1111/j.0906-7590.2004.03694.x.

[10] S. Schutte, "Regions at risk: predicting conflict zones in African insurgencies," *Political Sci Res Methods*, vol. 5, no. 3, pp. 447–465, 2017.

[11] H. Hegre, J. Karlsen, H. M. Nygård, H. Strand, and H. Urdal, "Predicting armed conflict, 2010–2050," *International Studies Quarterly*, vol. 57, no. 2, pp. 250–270, 2013, [Online]. Available: https://www.jstor.org/stable/24016137

[12] B. C. Boekestein, "A predictive logistic regression model of world conflict using open source data," *AFIT Scholar*, pp. 2–97, 2015, [Online]. Available: https://scholar.afit.edu/etd/101

[13] J. K. Nyiri, "The impact of ethnic conflict on economic development: the case of post-election violence in Kenya, 2007-2008.," University of Nairobi, 2014. [Online]. Available: http://erepository.uonbi.ac.ke/handle/11295/75407

[14] V. Owuor and S. Wisor, "The Role of Kenya's Private sector in peacebuilding: the case of the 2013 election cycle," *Broomfield, CO: One Earth Future Foundation*, pp. 1–36, 2014, doi: 10.18289/OEF.2014.003.

[15] C. Raleigh, A. Linke, H. Hegre, and J. Karlsen, "Introducing ACLED: An armed conflict location and event dataset," *J Peace Res*, vol. 47, no. 5, pp. 651–660, 2010, doi: 10.1177/0022343310378914.

[16] A. Zammit-Mangion, M. Dewar, V. Kadirkamanathan, A. Flesken, and G. Sanguinetti, *Modeling Conflict Dynamics with Spatio-temporal Data*. Springer, 2013.

[17] Open Data for Africa, "openAFRICA," Kenya Data Portal. [Online]. Available: https://africaopendata.org/

[18] A. Baddeley and others, "Analysing spatial point patterns in R," 2008.

[19] A. Baddeley, E. Rubak, and R. Turner, *Spatial point patterns: methodology and applications with R*. CRC press, 2015.

[20] R Core Team, "R: A Language and Environment for Statistical Computing." Vienna, Austria, 2022. [Online]. Available: https://www.r-project.org/

[21] F. P. Schoenberg, "Introduction to point processes," *Wiley Encyclopedia of Operations Research and Management Science*, p. 1, 2010, doi: 10.1002/9780470400531.

[22] N. Weidmann and M. D. Ward, "Spatial–temporal modeling of civil war: The example of Bosnia," in *GROW-Net Conference, Zurich*, 2008, pp. 1–10.

[23] C. Raleigh and C. Wafula, "Kenya's Political Violence Landscape in the Lead-Up to the 2022 Elections," 2022. [Online]. Available: https://acleddata.com/2022/08/09/kenyas-political-violence-landscape-in-the-lead-up-to-the-2022-elections/

[24] P. Kimani, H. Athiany, and C. Mugo, "Spatial dynamics of armed conflict in Kenya during the electioneering periods of 1997, 2002, 2007, 2013 and 2017 general elections," *Strathmore University*, pp. 1–7, 2019, [Online]. Available: https://su- plus.strathmore.edu/handle/11071/10501.

[25] L. Zhu, S. J. Cook, and M. Jun, "The promise and perils of point process models of political events," *arXiv preprint arXiv:2108.12566*, pp. 1–25, 2021.