Measuring Violence Risk in Space and Time using Kernel Density Estimation*

Augustin Tapsoba†

September 23, 2019

Abstract

Being able to assess conflict risk at local level is crucial for preventing political violence or mitigating its consequences. This paper develops a new methodological approach for measuring violence risk across space and time that improves the prediction of future conflict events. Violence is modeled as a stochastic process with an unknown underlying distribution. Each conflict event observed on the ground is interpreted as a random realization of this process and its underlying distribution is estimated using kernel density estimation methods in a three-dimensional space. The optimal smoothing parameters are estimated to maximize the likelihood of future conflict events. An illustration of the practical gains (in terms of out-of-sample forecasting performance) of this new methodology compared to standard space-time autoregressive models is shown using data from Ivory Coast.

JEL Codes: C1, O12, J13

Keywords: Conflict - Insecurity - Kernel Density Estimation

*I am grateful to Joan Llull and Hannes Mueller for their invaluable guidance and support. I also thank Michael Creel, Ada Ferrer-i-Carbonell, Pau Milan, and seminar/conference participants at Toulouse School of Economics, University of Amsterdam, NHH Bergen, ENS Lyon, Universite Libre de Bruxelles, Universitat Autònoma de Barcelona, EUDN Ph.D. workshop at CERDI, and Barcelona GSE Ph.D. jamboree for useful comments and suggestions. Financial support from the Spanish Ministry of Economy and Competitiveness through an FPI Grant at Barcelona GSE is gratefully acknowledged.

†Toulouse School of Economics. Email: augustin.tapsoba@tse-fr.eu

1 Introduction

Violent conflicts are serious humanitarian and economic threats in many developing countries.[1] Preventing political violence or mitigating its consequences for local populations is a daunting challenge that requires an efficient allocation of scarce resources in peace keeping missions, humanitarian aid, and other development projects. Anticipation of when and where violence is likely to occur is key for these types of operations and there is a growing demand for reliable tools for conflict forecasting from countries, non-governmental organizations (NGOs) and international organizations such as the United Nations (UN) and the World Bank (O'brien, 2010).

At local level, violence history is one of the most relevant and easily available sources of information for assessing violence risk (Bazzi et al., 2019; Hegre et al., 2019). This is mostly due to the persistence of violence over time and its diffusion in space through contagion. Data on other risk factors is either time invariant (terrain, geographic location, etc.) or rarely available in conflict-prone areas because they lack the stability and resources to produce them.

Using information on the timing and location of conflict events, this paper develops a new approach for measuring violence risk that can improve conflict prediction at local level. It models violence as a space-time point process with an unknown underlying distribution that has two main characteristics.[2] First, the occurrence of an event increases the likelihood that another event occurs in the same area. Second, violence can spread through the local environment via contagion.[3] The density of the underlying violence process is backed out of the observed pattern of conflict events. To do so, each event observed on the ground is interpreted as a random realization of the underlying process. Its density is estimated using kernel density estimation methods (Li and Racine, 2007; Silverman, 1986) in which the optimal smoothing parameters can be estimated to maximize the likelihood of occurrence of future events.

The basic principle behind this approach is that each event has its own contribution to the overall density at a given location. This contribution is given by a trivariate Gaussian kernel function. It is highest at the exact location of the event and fades out as we move away from

[1] More than 3/4 of countries in sub-Saharan Africa have experienced civil war since 1960 (Gleditsch et al., 2002) and the Internal Displacement Monitoring Centre estimated that more than 2 million people were newly displaced in Africa during the first six months of 2017 alone due to conflict.

[2] A space-time point process is a collection of random variables where each point represents the time and location of an event (Diggle, 2013). They have been used in the literature to model events such as crime, disease incidence, or the occurrences of natural disasters (fires, earthquakes, lightning strikes, tsunamis, etc.).

[3] An illustration of the first case is when battles between organized actors trigger waves of retaliatory or follow-up violence in a given area. For the second case, troops repeatedly attack clusters of nearby targets. This may happen because local vulnerabilities are well-known to them. Moreover, armed combatants can easily migrate from one area to another and violence in a given city can disrupt regional economic stability.

it in space or time. The dispersion of the kernel mass is controlled by a matrix of smoothing parameters. The density at a given point in space and time is obtained by summing up all the contributions. The estimated density is proportional to the statistical risk of violence in this framework. The performance of this new approach can therefore be evaluated based on its ability to predict future violence. The choice of the optimal smoothing parameters is crucial in this kernel density estimation approach. This paper also proposes a new bandwidth selection approach in which the optimal smoothing parameters are chosen to maximize the likelihood of future events.[4]

This simple density estimation approach allows for the use of all the information on the exact timing and location of conflict events when measuring violence risk. It also takes into account higher moments of the underlying violence process, such as its variance/dispersion. This is not feasible with the standard space-time autoregressive approach used so far in the literature to exploit conflict history data (Bazzi et al., 2019; Hegre et al., 2019; Weidmann and Ward, 2010). This approach consists indeed in slicing a given window of interest (country or region in a given time period) into space-time units.[5] Events that occur in each unit are counted and used to run space-time autoregressive regressions. The information on the exact timing and location of conflict events is therefore aggregated at a certain level leading to a potential loss of valuable data.[6] Moreover, this approach only uses distance to neighboring events (first order moment of violence process) to measure risk at a given location. The kernel density estimation approach imposes instead some desirable structure on the extent to which the risk generated by each event spreads in space and time. It assumes that this risk fades out at an exponential rate following a Gaussian kernel function. The dispersion of the kernel mass is estimated to maximize the likelihood of future events which gives flatter kernel functions for processes with higher dispersion. This framework allows one to run conflict prediction equations equivalent to the space-time autoregressive models but with just one regressor at the right hand side in which we can feed the exact timing and location of past events. The entire pattern of violence observed until time t is therefore used to measure the statistical risk of future violence in a single index.

An illustration of the practical gains, in terms of out-of-sample prediction performance, of the new methodology proposed here compared to space-time autoregressive models is shown us-

[4]The goal in modeling space-time processes is often to predict the risk of an incident occurring in a given area. Being able to target the prediction of future events when choosing the degree of smoothing to apply to the data is therefore a substantial advancement for space-time kernel density estimation methods in general.

[5]The spatial dimension of the units can be administrative subdivisions or geographical cell grids and the temporal dimension can be monthly, quarterly, yearly, etc.

[6]In the space-time autoregressive approach, one has to use coarse space-time units and limited lags to reduce the number of regressors in order to improve out-of-sample forecasting performance.

ing data from Ivory Coast. This country has experienced a relatively low-intensity but highly disruptive conflict between 2002 and 2011 that opposed the sitting government at the time to rebel groups that were powerful enough to control half of the country for a long period of time. Most of the conflict events that happened during the two civil wars in Ivory Coast were clashes between well organized armed groups.[7] This paper shows that kernel density approach systematically outperforms the standard space-time autoregressive models in out-of-sample prediction of violence. This gain is still present even in the prediction of violence onset at local level. This suggests that the gain in forecasting power does not only come from the persistence of violence over time in conflict affected areas. It also comes from its diffusion across space. The results are robust to several alternative data and model specifications.

Beyond its ability to make better conflict predictions based on violence history data, the methodology proposed here to measure violence risk is general and can be useful for several other purposes. In particular, it provides a new metric of violence risk that goes beyond the incidence of conflict events in a given space-time window. This is crucial because economic agents often change their consumption/production behavior in reaction to violence risk even before or without any manifestation of violence around them. The kernel density estimation method can therefore provide a better measure of exposure to the adverse effects of conflict. Based on this intuition, Tapsoba (2018) uses this method to show that insecurity in conflict-prone areas can lead to major health setbacks for young children even in absence of immediate violence around them.

The kernel density approach also has the advantage of modeling violence in an agnostic way in the sense that it tries to capture variations in the density of the equilibrium violence process without making any claim on the mechanisms behind it. Whether violence occurs in a given space-time window because of terrain, presence of natural resources, negative income shocks, propagation of heinous rhetoric, etc., it always mobilizes actors and resources (at local level) that can be used to perpetrate future attacks in the same area or in neighboring locations. The clustering of events in space and time is an indication of high violence risk irrespective of what is the initial source of this violence.

This paper is related to 3 main strands of literature. First, it is related to the conflict prediction literature. This literature has made significant progress in assessing violence risk at

[7] Civil wars that oppose organized armed groups are more likely to diffuse in space and time by escalation or relocation (Schutte and Weidmann, 2011) and the method proposed here is able to capture such diffusion. Even in cases of one-sided violence and mass repression such as the Rwandan genocide, Yanagizawa-Drott (2014) shows that "heinous" radio broadcasts increased militia violence directly by influencing behavior in villages with radio reception but also indirectly by increasing participation in violence in neighboring villages. He also shows that only 10 percent of the total violence could be attributed to radio.

country level, with recent contributions that use text analysis of newspaper content to predict civil wars (Mueller and Rauh, 2016; Chadefaux, 2014). At local level, this literature has relied heavily on the diffusion of violence across space and over time to forecast future violence. Weidmann and Ward (2010) use, for instance, spatial and temporal autoregressive models to predict conflict in Bosnia. Hegre et al. (2019) use an ensemble model to predict conflict at disaggregated level in Africa. A substantial part of their prediction power comes from violence history data that they also exploit using space-time autoregressive models. Bazzi et al. (2019) use a wide range of machine learning techniques and hundreds of annual risk factors to predict violence in Colombia and Indonesia. Their models are able to identify persistent, high-violence hot spots but not variations of violence risk over time.[8] The new approach used in this paper models violence as a stochastic process in space and time in order to aggregate conflict history data into a single index that captures variations in violence risk. This new approach outperforms (in terms of out-of-sample prediction) the standard space-time autoregressive approach used in all the papers mentionned above. It is therefore a better way of incorporating violence history data into ensemble models such as those developed in Bazzi et al. (2019) and Hegre et al. (2019).

This paper also belongs to a small but growing literature in both Economics and Political Science that focuses on understanding how (and why) violence spreads across space and time. There is evidence that violent events are clustered in space and time (Townsley et al., 2008; Schutte and Weidmann, 2011). Using the case of Northern Ireland, Mueller et al. (2017) show that distance between attackers and targets of attacks is a cost that can explain the distribution of violent events. Another diffusion mechanism explored in this literature is the feasibility (or sustainability) of rebellions and insurgencies. Berman et al. (2017) show that a rise in world price of specific minerals leads to more violence around the mining areas of these minerals and more attacks perpetrated by the armed groups that control them. Yanagizawa-Drott (2014) also shows that radio broadcast played an important role in the diffusion of violence during the genocide in Rwanda. The new approach proposed in this paper to measure violence risk is complementary to these studies in the sense that it uses data on the realizations of conflict events to estimate the density of the underlying violence generating process. This underlying process is an equilibrium state that could have been generated by any combination of the mechanisms studied in this literature. Irrespective of the mechanism behind, violence begets violence that often persists over time or spreads across space so conflict history is able to inform us on future risk.

[8] Other attempts of violence forecasting at sub-national level include Blair et al. (2017), Chiba and Gleditsch (2017) and Witmer et al. (2017).

Finally, this paper is related to the literature on the cost of conflict, specially at micro level. Standard approach in this literature consists in counting the number of events in a given space-time window to define exposure to the adverse effects of conflict (Dagnelie et al., 2018; Akresh et al., 2016; Leon, 2012). The few papers that look beyond the incidence of violence include Besley and Mueller (2012) that uses a Markov switching process over time to model violence in Northern Ireland; Rockmore (2017) that uses self-reported data on perceived risk in Uganda; and Arias et al. (2014) that uses data on presence of armed groups in Colombia. The new methodology proposed here can be used to capture violence risk beyond the incidence of conflict events in order to estimate correctly the cost of insecurity on households and firms. Its main advantage compared to the approaches used in the other papers is that it provides a measure of the statistical risk of violence and relies solely on the observed pattern of events in space and time.

The remainder of the paper is organized as follows. The next section points out the limits of the standard approach used so far to measure violence risk in the literature. Section 3 presents the new methodological approach proposed here to measure violence risk across space and time. Section 4 discusses an application to conflict prediction in Ivory Coast. Section 5 concludes.

2 Limits of Standard Space-Time Autoregressive Approach

Violence history is one of the most relevant sources of information for predicting future violence (Bazzi et al., 2019; Hegre et al., 2019). Standard conflict event datasets such as ACLED (Armed Conflict Location and Event Dataset (Raleigh et al., 2010)) and UCDP-GED (Sundberg and Melander, 2013) provide information on exact timing and location of different conflict events. The standard space-time autoregressive approach used so far in the literature is not able to fully account for all the information contained in such detailed violence history data. This approach consists indeed of splitting the area considered into space-time windows of a given size in order to run a regression with spatial and temporal lags as regressors following Equation 1.[9]

$$Y_{i,t+1} = \sum_{k=0}^{K} \sum_{q=0}^{Q} \alpha_k^{(t-q)} Y_{\mathcal{N}_k(i),t-q} + \eta + \epsilon_{i,t+1}, \tag{1}$$

[9]The space windows can be administrative units or geographic cells of a certain size and time window can be monthly, quarterly or yearly. The size of the space-time window is usually hand-picked.

where $\mathcal{N}_k(i), t-q$ is the neighborhood order k of cell i at period $t-q$, $Y_{i,t+1}$ is dummy equal 1 if cell i experiences a conflict event at period $t+1$.[10]

Ideally, one could split the area of interest into very fine space-time windows to account for all the information on the exact timing and location of conflict events. However, this will increase substantially the number of regressors in the spatio-temporal autoregressive model. It leads at best to an overfitting of the regression model which translates into very poor out-of-sample forecasting performance.[11] An other alternative is to run a LASSO (Least Absolute Shrinkage and Selection Operator) regression to reduce the number of lags to include in the prediction regression, but this leads to a mechanical exclusion of some lags without any improvement in out-of-sample performance specially with fined-grained space-time windows.[12]

The second limitation of the standard approach is that it relies only on distance in space and time to different conflict events to capture the risk of occurrence of future events. It is not able to account for higher moments of the violence process such as the dispersion of conflict events or the trajectory of the process in space. The new approach proposed in this paper allows us to deal with these limitations of the existing methods by modeling violence as a stochastic process across space and time.[13]

3 Estimation of Violence Risk in Space and Time: Kernel Density Estimation Approach

In order to predict conflict events in space and time, we need a way to quantify violence risk at the local level. To do this, I model the observed conflict events as random realizations of an underlying process that can be backed out of the violence pattern, and used to predict the likelihood of future events. In this section, I show how non-parametric density estimation methods can be used for this purpose.

Non-parametric density estimation methods have been initially used in the literature to as-

[10]Spatio-temporal autoregressive models like the ones in Harari and Ferrara (2018) are special cases of the one shown here. They are used in spatial econometrics to estimate the impact of a given regressor (climate shocks in Harari and Ferrara (2018)) on a dependent variable (violence) accounting for the spatial and temporal correlation in both the dependent variable and the variable of interest. This method uses spatial and temporal filters to remove the spatial and temporal autoregressive terms in the corresponding regression equation.

[11]In extreme cases we can end up with more regressors than observations which is not feasible in the standard regression framework.

[12]LASSO model is a logistic regression model that penalizes large coefficients and forces all but the most important ones to zero.

[13]This paper tries to predict violence at $t+1$ given information at hand at the end of period t. This is the reason why there is no spatial lag variables at $t+1$ in Equation 1. The Kernel density estimation approach also uses the same information set.

sess basic characteristics of an unknown distribution such as skewness, tail behavior, number, location and shape of modes (Silverman, 1986). Nowadays, they play a major role in machine learning, classification and clustering.[14] They are popular methods in crime literature for hotspot mapping and in seismology literature for seismic hazard estimation. They can also be used (like in this paper) as input for more sophisticated analysis. DiNardo et al. (1996) used kernel density estimation method to build counter-factual densities in order to study the effects of institutional and labor market factors on changes in the U.S. distribution of wages in the 80s. Kernel density estimation methods have also been extensively used in poverty analysis to measure poverty from grouped data (mean incomes of a small number of population quantiles) through the estimation of the underlying global income distribution (Sala-i-Martin, 2006; Minoiu and Reddy, 2014; Sala-i-Martin, 2002).

3.1 Principle of Kernel Density Estimation (KDE)

Let's assume violence process X has an unknown probability density function (pdf) f. Density estimation consists of constructing an estimate of f based on a representative sample of random realizations $\{x_1, ..., x_n\}$ of X.

Let's begin with the simple case of a continuous, univariate random variable X.[15] Each observed event has its own contribution to the density at a given location x. This contribution is given by a kernel Gaussian function $k \sim \mathcal{N}(0, h)$.[16] The value of the kernel function is highest at the exact location of the data point, and fades out at an exponential rate as we move away from it. The density estimate at x is obtained by summing up all the contributions according to Equation 2. Closer events have higher contributions to the density and clustered events generate peaks corresponding to the modes of the underlying density function as shown in Figure 1.

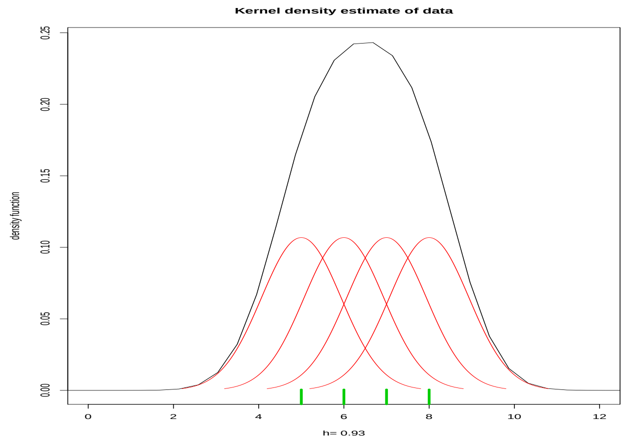
$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} k\Big(\frac{x - x_i}{h}\Big) \tag{2}$$

[14] For instance, some clustering methods are based on bump hunting, i.e., locating the modes in the density and Bayes classifiers are based on density ratios that can be implemented via density estimation. Applications of density estimation in machine learning and classification are discussed in more depth in the books of Izenman (2008) and Hastie et al. (2009).

[15] To estimate violence risk across space and time, an extension to 3 dimensions (latitude, longitude and time) is shown below.

[16] The bandwidth is the most crucial choice to make in KDE because it controls the degree of smoothing applied to the data. The kernel form is only responsible for the regularity of the resulting estimate (continuity, differentiability) and Gaussian kernels give estimated density function that has derivatives of all orders (Silverman, 1986).

Figure 1: Aggregation of Individual Kernels in KDE



Notes: Kernel density estimation with 4 data points (in green). The Red Gaussian functions are individual kernels aggregated to get the density estimate in black.

3.2 Generalization to More Than One Dimension

The univariate kernel density estimation can easily be extended to the multivariate case. The estimated density is given by

$$\hat{f}(\bar{x}) = \sum_{i=1}^{n} \frac{1}{n \mid H \mid} K\Big(H^{-1}(\bar{x_i} - \bar{x})\Big),$$

where $\bar{x_i}$ and $\bar{x}$ are d-dimensional vectors, H is a dxd symmetric matrix of parameters to be estimated and K is a multivariate kernel function.

To measure conflict risk in space and time, we need to consider 3 dimensions: latitude, longitude and time. I assume independence between each of the 3 dimensions because of the relatively small number of conflict events given the estimation space. I also follow the space-time kernel density estimation literature and assume that latitude and longitude dimensions have the same smoothing parameter h_s. The matrix of smoothing parameters H is therefore parametrized in the most simplistic way possible as follows:[17]

$$H = \begin{pmatrix} h_s & 0 & 0 \\ 0 & h_s & 0 \\ 0 & 0 & h_t \end{pmatrix}.$$

The density at a given point $\bar{x} = (x, y, t)$ in space and time is therefore given by:

[17] See Appendix C for a discussion on the role of off-diagonal elements of the matrix H.

$$\hat{f}(x,y,t) = \frac{1}{nh_s^2 h_t}\sum_{i=1}^{n} k\Big(\frac{x-x_i}{h_s}\Big)k\Big(\frac{y-y_i}{h_s}\Big)k\Big(\frac{t-t_i}{h_t}\Big),$$

where n is the number of events that occurred before time t. The value of this function reflects the likelihood of an incident occurring around location (x, y) at time t. Space-time kernel density estimation of this type have been used to study patterns of forest fire (Tonini et al., 2017), crime (Brunsdon et al., 2007; Nakaya and Yano, 2010), occurence of disease (Eaglin et al., 2017), etc.

Choice of Smoothing Parameters and Violence Prediction

Kernel Density Estimation method imposes some structure in the contribution of each conflict event to violence risk. The space-time smoothing parameters control the dispersion of the kernel mass. The choice of the smoothing parameters h (for space or time) is crucial in this method. Small values of h reduce the bias by putting all the mass just around each data point and the density estimate displays spurious variations in the data. When h is too big, each individual kernel becomes flatter and important details in the distribution can be obscured (see Appendix Section A for an illustration).

For the purpose of predicting future conflict events, the optimal smoothing parameters can be chosen to minimize Mean Integrated Square Error in a 2-stage approach, following the existing literature, or to directly maximize the likelihood of future events across space following a new approach developed in this paper.

The 2-Stage Approach

In the 2-stage approach, the optimal smoothing parameters are estimated by minimizing the Mean Integrated Squared Error (MISE)

$$\begin{aligned} MISE(H) &= \int [\hat{f}(\bar{x},H) - f(\bar{x})]^2 d\bar{x} \\ &= \int \hat{f}(\bar{x},H)^2 d\bar{x} - 2\int \hat{f}(\bar{x},H) f(\bar{x}) d\bar{x} + \int f(\bar{x})^2 d\bar{x}, \end{aligned}$$

where $\hat{f}(\bar{x}, H)$ is the kernel density estimate and $f(\bar{x})$ is the unknown density.

$MISE(H)$ can be evaluated and minimized without knowing explicitly $f(\bar{x})$ by using a bootstrap method (Taylor, 1989) as described in Section B of the appendix.

The density of the violence process (risk measure) is then estimated, using the optimal smoothing parameters, for each location (i, t) in space and time as follows:

$$Risk_{i,t} \propto \sum_{j \in \mathcal{H}(t)} K\Big((H)^{-1}(\bar{x_j} - \bar{x_i})\Big), \tag{3}$$

where $\mathcal{H}(t)$ is the set of events that happened up till time t. The estimated risk can then be used to predict future events:

$$Y_{i,t+1} = \alpha Risk_{i,t} + \eta + \epsilon_{i,t+1}. \tag{4}$$

The Direct Approach

A more intuitive approach is to estimate the optimal smoothing parameters directly by maximizing the likelihood of future events across space. To do so, I use a latent variable representation of violence risk in a logit framework. The latent violence state Y^* is given by:

$$Y^* = \gamma \sum_{j \in \mathcal{H}(t)} K\Big((H)^{-1}(\bar{x_j} - \bar{x})\Big) + \eta + \epsilon, \qquad \epsilon \sim F(.),$$

where $F(.)$ follows standard logistic distribution. We observe violence only if latent risk is high enough: $Y = 1 \Leftrightarrow Y^* > 0$

The probability of occurrence of an event in a given space-time window is:

$$P[Y_{i,t+1} = 1|\mathcal{H}(t)] = P\Big[\epsilon > -\gamma \sum_{j \in \mathcal{H}(t)} K\Big((H)^{-1}(\bar{x_i} - \bar{x})\Big) - \eta\Big].$$

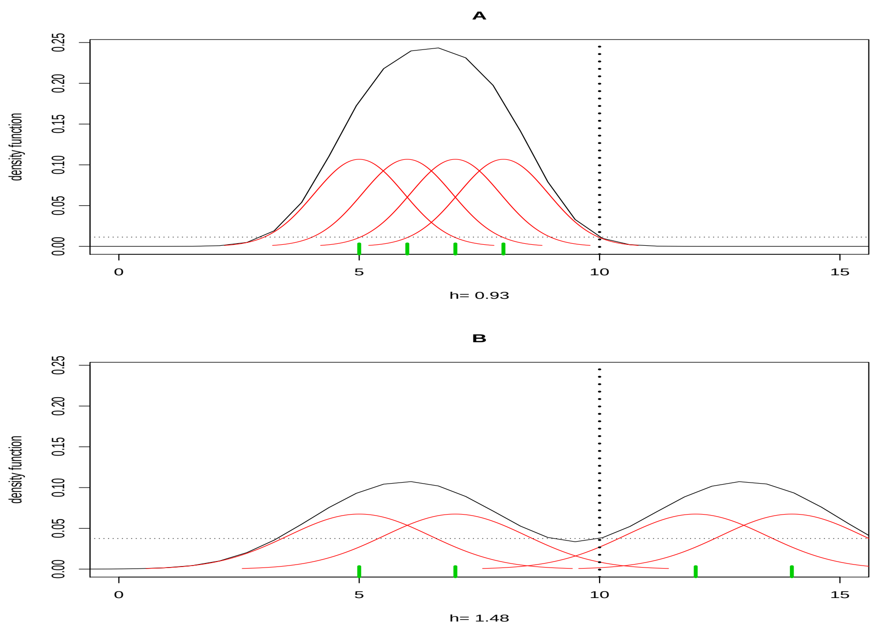
The log-likelihood of observed sample of events is:

$$\mathcal{L}(H, \eta, \gamma) = \sum_{i,t} Y_{it} ln\big(P[Y_{i,t} = 1|\mathcal{H}(t-1)]\big) + (1 - Y_{it}) ln\big(P[Y_{i,t} = 0|\mathcal{H}(t-1)]\big).$$

The optimal smoothing parameters are therefore estimated to maximize the likelihood of future events in the data. This new method turns the kernel density estimation approach into a "logit-like" estimation in which we can feed all the events from the violence history with their exact timing and location without any issue of dimensionality. It also maximizes the prediction power of the kernel density estimation approach compared to choosing the smoothing parameters based on the minimization of $MISE(H)$.

The dispersion of the observed sample of events is a feature of the underlying density, so minimizing the $MISE(H)$ (or maximizing the likelihood of future events) leads to larger optimal smoothing parameters for samples with more dispersion. The smoothing parameter is indeed an increasing function of the sample variance.[18] This is illustrated (for 1 dimension) in Figure 2 where both panels show individual kernels and density estimates from 4 data points at equal distance from a given location $x = 10$. There is less dispersion in the event data in panel A compared to panel B and location x should be more at risk in B than A. The kernel density estimation is able to distinguish these two cases when estimating them as two separate processes. The standard approach is not able to distinguish these two situations because it relies only on distance to neighboring events.

Figure 2: Fixed KDE and Sample Variance in Separate Processes



3.3 Discussion and Comparison With Standard Approach

The KDE approach proposed in this section imposes some structure in the way in which each event contributes to the likelihood of future events. This very simple approach allows us to use all the information on exact timing and location of conflict events by shrinking them into a single indicator of risk for conflict prediction.

The choice of the smoothing parameters controls the speed at which the contribution of each individual event to the risk fades out. The data-driven methods used here to choose these

[18] In the particular case of the underlying pdf being normally distributed, one can show that the optimal smoothing parameter h_{opt} is actually proportional to sample variance Silverman (1986).

parameters ensure that higher moments of violence process such as the dispersion of conflict events are used in the estimation of the risk. In particular, the idea of choosing KDE bandwidth by maximizing the likelihood of future events can be useful for other applications that use space-time kernel density estimation approach.

Finally, the KDE approach is also less sensitive to measurement errors in the conflict event data. It only requires, for instance, a representative sample of random realizations of the underlying process unlike the standard approach that requires an exhaustive list of events. Moreover, isolated events or coding errors in the timing or location of few events in the conflict data will still translate into low violence risk with the KDE approach. In this approach, it is only the aggregation of contributions from several nearby events that can lead to substantial increase in violence risk.

The KDE approach presented here assumes that all the conflict events are equivalent, irrespective of their nature (violence against civilians, battle between armed groups, etc.) or their human toll (number of casualties).[19] With a large enough sample, one can split conflict events into the relevant sub-categories and estimate different smoothing parameters for each of them in the direct approach. Events that generate high number of fatalities, for instance, could therefore have smaller or larger smoothing parameters compared to low intensity events.[20] Similarly, events that happen across country or regional borders could also have different smoothing parameters if there are enough conflict events to estimate them.

4 Application: Conflict Prediction in Ivory Coast

4.1 Background and Data

Ivory Coast is a previous French colony that enjoyed a prolonged period of economic growth since its independence in 1960 until 1990. Political instability has been sparked by the power struggle following the death of the country's first president, Felix Houphouet-Boigny, in 1993. The Interim President, Henri Konan Bedie, in order to secure a win in the 1995 elections, changed the electoral code to exclude his rival and former Prime Minister, Alassane Dramane Ouattara, from running based on his status of son of a migrant. The 1995 election was then boycotted by the other

[19] A possible interpretation of the assumption that conflict events are equivalent, irrespective of their human toll, is that the timing and location of events are determined by the underlying violence process but the number of casualties that they could each generate is random and does not affect how risk spreads in space and time.

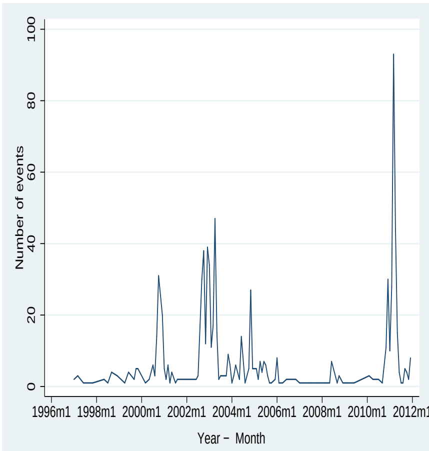
[20] The smoothing parameters control the dispersion of kernel mass. Smaller values mean that most of the mass is concentrated around the data point and larger values mean the mass is more spread. Whether events with high death toll should have smaller or larger smoothing parameters is an empirical issue.

opposition leaders in protest against this discrimination and Bedie was elected with 96% of the votes.

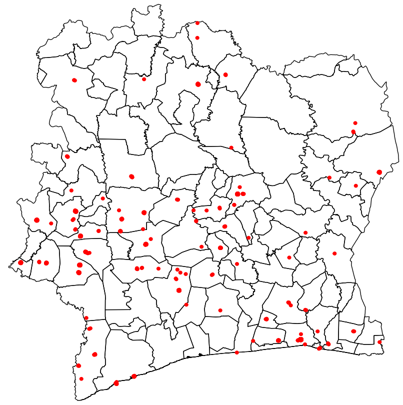
In July 1999, Alassane Ouattara, left his job at the International Monetary Fund and returned to run for the 2000 presidential elections. His plan to challenge Bedie again split the country along ethnic and religious lines based on their own origins which determined their electoral basis. This political turmoil gave a pretext to a group of army soldiers to intervene and overthrow Bedie by a coup in December 1999. Robert Guei, an army officer was chosen to lead a transition to new elections and restore the order. New elections took place in October 2000, but Alassane Ouattara was still excluded from the electoral process because of not being "Ivorian enough". General Guei proclaimed himself president after announcing he had won the presidential elections, but was forced to flee in the wake of a popular uprising and was replaced by his challenger Laurent Gbagbo. Fighting erupted between President Gbagbo's mainly southern Christian supporters and followers of his main opponent Alassane Ouattara, who were mostly Muslims from the north. Tensions lasted for almost a year before both challengers agreed to work towards reconciliation in March 2001.

This fragile reconciliation process was disrupted in September 2002 when a mutiny in Abidjan grew into a full-scale rebellion with Ivory Coast Patriotic Movement rebels seizing control of the north. French interposition troops were sent to limit the clashes between the two armies as shown in Figure A1 but intense battles between the two sides took place until March 2003 when the first peace agreement was signed and rebels made their entry into the government. Many other peace talks were held as actors were resuming clashes at one point or another during the implementations of the different peace agreements. This went on until March 2007 when a power sharing deal was signed. Under this deal, Guillaume Soro, leader of the rebel group, was made Prime Minister of Ivory Coast and the new government was put in charge of preparing the elections that would end the crisis and restore a stable constitutional order. The elections after being postponed twice were finally held in December 2010 but led to another crisis. The electoral commission declared Mr Ouattara the winner of presidential election run-off. Mr Gbagbo refused to accept these results and the dispute between the two camps soon escalated into extremely violent clashes until the capture of Gbagbo in April 2011 after the loyalist army has been defeated by the rebels backed by French troops under UN mandate.

Figure 3: Distribution of Violent Events in Ivory Coast Between 1997 and 2012



(a) Monthly Number of Events



(b) Spatial Distribution of Conflict Events

Conflict Event Data

To locate violence in space and time, I use information from ACLED on the exact timing (day) and location (latitude and longitude) of conflict events that happened in the country. These events are obtained from various sources, including press accounts from regional and local news, humanitarian agencies or research publications during the conflict. ACLED records all political violence, including battle between armed groups, violence against civilians, rioting and protesting. I focus on war-related events: battles, explosions/remote violence and violence against civilians. I also only keep all the events recorded with geographical precision at municipality level or lower (more than 90 percent of all the events). Duplicates (same location and time) are eliminated to have a list of different days and location with at least one conflict event.

Figure 3a shows the number of recorded events per month across the whole country. We can see that the recorded events are consistent with the timing described in the preceding paragraphs with some violence peaks after the 2000 elections, during the first months following the start of the first civil war in 2002 and the second civil war following the 2010 elections. Fewer events were recorded during the negotiations period until the comprehensive power sharing treaty was signed in 2007. The spatial distribution of the events is also shown in Figure 3b. One can see that conflict incidents are mostly clustered in the central and western parts of the country.

4.2 Out-of-Sample Prediction and Performance Metrics

The sample of conflict events is split between training and evaluation sets. The training set uses data from 1997 to 2005 to estimate parameters of the different models discussed in this paper. To run space-time autoregressive models, I use PRIO-GRID cells (Tollefsen et al., 2012) for Ivory Coast. PRIO-GRID is a vector grid network with a resolution of 0.5 x 0.5 decimal degrees (around 50 x 50 kilometers), covering all terrestrial areas of the world. For a country such as Ivory Coast, it provides a full grid of the entire country divided in sub-national units. The temporal frequency considered is monthly. From the centroid of each PRIO-GRID cell, I define spatial lag variables by drawing several rings of increasing radius around. I then count the number of events that occur between ring $k-1$ and rink k for a given month.[21] This approach allows for more flexibility in defining the size of the spatial lags to use in the space-time autoregessive method.[22]

The outcome variable in each model is a dummy variable $Y_{i,t+1}$ equal 1 if there is at least one conflict event that happenned in cell grid i during month t. Each model also uses violence history data up till month t to predict violence next period. In the space-time autoregressive models, I use violence history data of events that happened within 200 kilometers from the centroid of cell i up to 2 years before time t. I also use events that happen in the same space-time window to compute the risk estimate for each cell $Y_{i,t}$ in the KDE approach in order to properly compare the two methods.

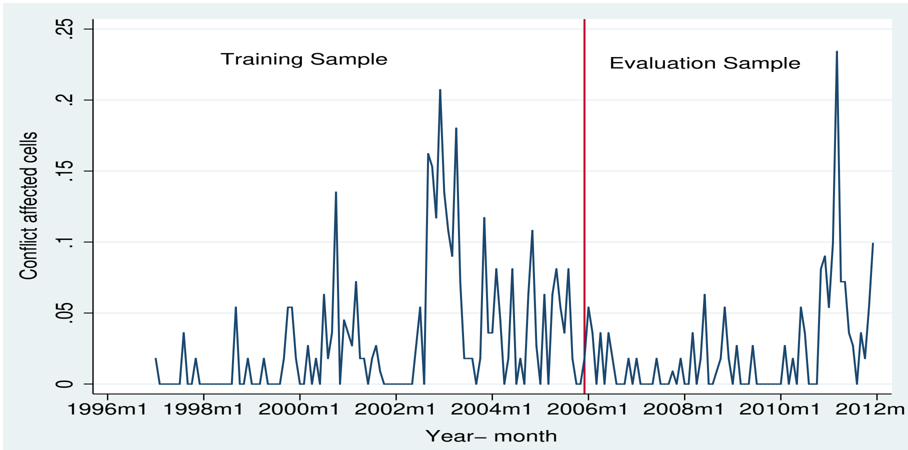
Figure 4 shows the share of conflict affected cells per year-month. More than 20 percent of the 111 PRIO cells in Ivory coast were affected by a conflict event during violence peaks in 2002 and 2011. There are also few months with no violent event in the whole country and substantial variation in the share of conflict-affected cells between these two extremes.

Prediction power of each model is evaluated out-of-sample using receiver operating characteristic (ROC) curves with data from 2006 till 2011. These curves plot the tradeoff between true and false positive rates in conflict prediction for a given model. The area under the curve, or AUC, captures the probability that a randomly chosen pair of observations is correctly ordered

[21] First spatial lag corresponds to events that occur between 50km and 60km (70 kilometers) if we consider a lag frequency of 10 kilometers (20 kilometers) for instance.

[22] The standard approach in the literature is to consider the 50 x 50 kilometers cell grids and their first and second degree neighbors for instance. This is equivalent to drawing rings of 50, 100 and 150 kilometers. First degree neighborhood corresponds to area between 50 and 100 kilometers, second degree neighborhood corresponds to area between 100 and 150 kilometers.

Figure 4: Share of PRIO Cells With at Least One Conflict Event



in terms of predicted risk of violence. A model that performs no better than chance would have an AUC of 0.5 (45 degree line in ROC curve) while a perfect model would have an AUC of 1: a true positive rate of 1 at a false positive rate of 0.[23]

4.3 Violence Forecasting at Local Level

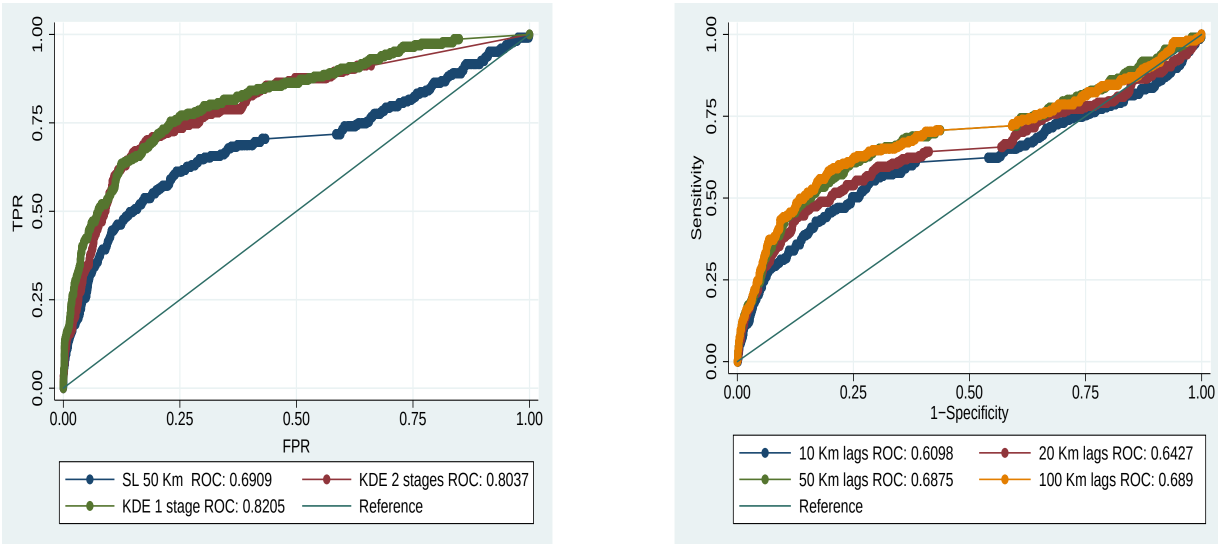
Figure 5a shows that the "direct" (or 1-stage) approach in the KDE method has an AUC of 0.82 which is slightly higher than the AUC of the "2-stage" approach (0.80). The KDE methods strictly outperform the spatio-temporal autoregressive approach which gives an AUC of 0.69. At a false positive rate (FPR) of 25 percent both KDE methods reach a true positive rate (TPR) of about 75 percent versus 60 percent for the spatio-temporal autoregressive method. At a false positive rate of 50 percent, the gap widens even more. Figure 5b shows that using finer lags in space does not improve out-of-sample performance of the space-time autoregressive model.[24]

To check whether the gain in prediction power comes mostly from persistence of violence in conflict affected cells over time, I also compare the performance of the different methods in the prediction of violence onset. This is to rule out the concern that all the forecasting gain shown previously could be coming from situations in which conflict cell x month windows are following each other. Figure 6a and 6b show performance in prediction of violence onset at local level. In Figure 6a, onset is defined as the occurrence of a conflict event after at least 3 months without violence in a given grid cell and in Figure 6b after 6 months. They both show that predicting onset as opposed to incidence is harder in general but KDE methods still outperform substantially

[23]An AUC of 1 means that all conflict windows (space and time) are predicted without raising any false alarms.

[24]I also show in appendix Figure A3 that using the number of events in each space-time lag rather than a dummy variable does not improve prediction power of the standard model.

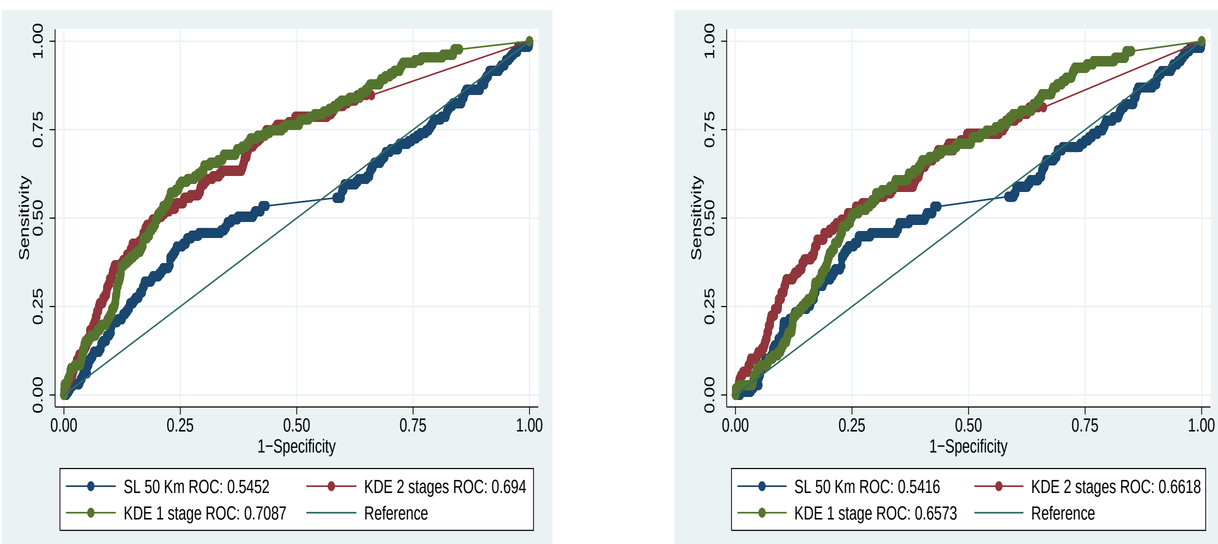
Figure 5: ROC Curves for Out-Of-Sample Violence Prediction from 2006 to 2012



(a) KDE Approach

(b) Spatio-temporal Autoregressive Approach

Figure 6: ROC Curves for Out-Of-Sample Violence Onset Prediction from 2006 to 2012



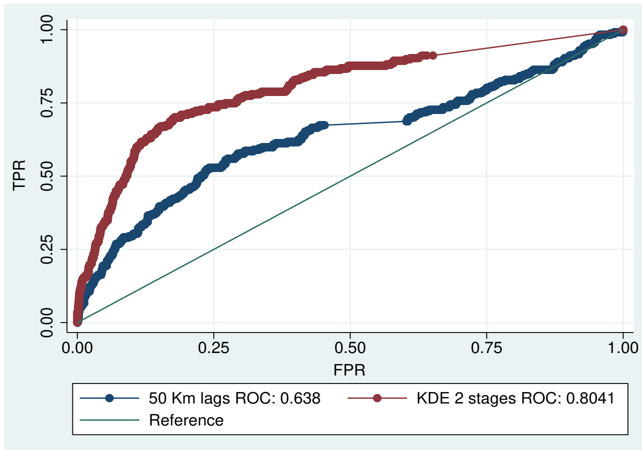
(a) 3 Months Onset

(b) 6 Months Onset

the standard approach with AUC improving by up to 0.15. This suggests that KDE approach is not just tapping into the persistent effect of violence in time but also its diffusion in space.

Figure 7 shows out-of-sample forecasting power after controlling for cell fixed effect. The space-time autoregressive model now performs less while the KDE approach performs almost equally as before. At a false positive rate of 25 percent, the gap between the two approaches is now of 20 percent (as opposed to 12 percent before). This suggests that a larger part of variation in the predicted conflict risk using the autoregressive model is explained by the fixed effect.

Figure 7: ROC Curves for Out-Of-Sample Violence Prediction Using Cell Fixed Effects



5 Conclusion

This paper introduced a new way of exploiting violence history information to measure conflict risk at local level. Violence is modeled as a stochastic process with an unknown underlying distribution that is backed out of the observed pattern of conflict events using kernel density estimation methods. A new method for estimating the optimal smoothing parameters is also proposed.

This new approach allows us to use all the information on the exact timing and location of conflict events without any issue of dimensionality in prediction equations. It also uses higher moments of the violence process to estimate risk compared to standard space-time autoregressive models. An application of this approach to conflict prediction in Ivory Coast shows that it outperforms (in terms of out-of-sample prediction) the standard space-time autoregressive models. This gain in performance is substantial even when predicting violence onset at local level.

The kernel density estimation approach can be used in combination with other sources of information to build reliable early warning systems for governments, NGOs and international organizations that intervene in conflict-prone areas. Ensemble models like those proposed in Bazzi et al. (2019) and Hegre et al. (2019) could be improved by using the proposed methodology to model the violence history component of such models.

The new methodology developed here also opens the doors for revisiting some of the findings in the conflict cost literature. Insecurity in conflict-prone areas often triggers changes in economic behavior even in absence of immediate violence. With the new method proposed here, one can easily build a new metric of exposure to the adverse effects of conflict. The integral of the estimated density of the underlying process over a given space-time window gives indeed a measure of the likelihood of occurrence of a conflict event in this window. This metric captures violence risk beyond the incidence of conflict events. The statistical risk of violence risk can be high in space-time windows with no conflict event and conversely, it can be low in windows with isolated events. If we assume that beliefs of economic agents on the ground are driven by the underlying violence process, this methods provides better way of defining treatment when estimating the cost of insecurity in conflict-prone areas.

References

Akresh, Richard, German Caruso, and Harsha Thirumurthy, "Detailed Geographic Information, Conflict Exposure, and Health Impacts," 2016.

Arias, María Alejandra, Ana Ibáñez, and Andres Zambrano, "Agricultural Production Amid Conflict: The Effects of Shocks, Uncertainty, and Governance of Non-State Armed Actors," Technical Report, Universidad de Los Aandes-Cede 2014.

Bazzi, Samuel, Robert A Blair, Christopher Blattman, Oeindrila Dube, Matthew Gudgeon, and Richard Merton Peck, "The Promise and Pitfalls of Conflict Prediction: Evidence from Colombia and Indonesia," Technical Report, National Bureau of Economic Research 2019.

Berman, Nicolas, Mathieu Couttenier, and Dominic Rohner, "This Mine is Mine! How Minerals Fuel Conflicts in Africa," *American Economic Review*, 2017, *107* (6), 1564–1610.

Besley, Timothy and Hannes Mueller, "Estimating the Peace Dividend: The impact of violence on house prices in Northern Ireland," *The American Economic Review*, 2012, *102* (2), 810–833.

Blair, Robert A, Christopher Blattman, and Alexandra Hartman, "Predicting local violence: Evidence from a panel survey in Liberia," *Journal of Peace Research*, 2017, *54* (2), 298–312.

Brunsdon, Chris, Jonathan Corcoran, and Gary Higgs, "Visualising space and time in crime patterns: A comparison of methods," *Computers, environment and urban systems*, 2007, *31* (1), 52–75.

Chadefaux, Thomas, "Early warning signals for war in the news," *Journal of Peace Research*, 2014, *51* (1), 5–18.

Chiba, Daina and Kristian Skrede Gleditsch, "The shape of things to come? Expanding the inequality and grievance model for civil war forecasts with event data," *Journal of Peace Research*, 2017, *54* (2), 275–297.

Dagnelie, Olivier, Giacomo De Luca, and Jean-François Maystadt, "Violence, Selection and Infant Mortality in Congo," *Journal of Health Economics*, 2018.

Diggle, Peter J, *Statistical analysis of spatial and spatio-temporal point patterns*, Chapman and Hall/CRC, 2013.

DiNardo, John, Nicole M Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, September 1996, *64* (5), 1001–1044.

Eaglin, Todd, Isaac Cho, and William Ribarsky, "Space-time kernel density estimation for real-time interactive visual analytics," in "Proceedings of the 50th Hawaii International Conference on System Sciences" 2017.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand, "Armed conflict 1946-2001: A new dataset," *Journal of peace research*, 2002, *39* (5), 615–637.

Harari, Mariaflavia and Eliana La Ferrara, "Conflict, climate, and cells: a disaggregated analysis," *Review of Economics and Statistics*, 2018, *100* (4), 594–608.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, "Unsupervised Learning," in "The Elements of Statistical Learning," Springer, 2009, pp. 485–585.

Hegre, Håvard, Marie Allansson, Matthias Basedau, Michael Colaresi, Mihai Croicu, Hanne Fjelde, Frederick Hoyles, Lisa Hultman, Stina Högbladh, Remco Jansen et al., "ViEWS: A political violence early-warning system," *Journal of Peace Research*, 2019, *56* (2), 155–174.

Izenman, Alan Julian, *Modern Multivariate Statistical Techniques: Regression, Classification and Manifold learning*, Springer, 2008.

Leon, Gianmarco, "Civil conflict and human capital accumulation the long-term effects of political violence in perú," *Journal of Human Resources*, 2012, *47* (4), 991–1022.

Li, Qi and Jeffrey Scott Racine, *Nonparametric econometrics: theory and practice*, Princeton University Press, 2007.

Minoiu, Camelia and Sanjay G Reddy, "Kernel Density Estimation on Grouped data: The Case of Poverty Assessment," *The Journal of Economic Inequality*, 2014, *12* (2), 163–189.

Mueller, Hannes Felix and Christopher Rauh, "Reading Between the Lines: Prediction of Political Violence Using Newspaper Text," *CEPR Discussion Paper Series*, 2016.

_ , **Dominic Rohner, and David Schönholzer**, "The Peace Dividend of Distance: Violence as Interaction Across Space," *CEPR Discussion Paper Series*, 2017.

Nakaya, Tomoki and Keiji Yano, "Visualising crime clusters in a space-time cube: An exploratory data-analysis approach using space-time kernel density estimation and scan statistics," *Transactions in GIS*, 2010, *14* (3), 223–239.

O'brien, Sean P, "Crisis early warning and decision support: Contemporary approaches and thoughts on future research," *International studies review*, 2010, *12* (1), 87–104.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre, and Joakim Karlsen, "Introducing acled: An armed conflict location and event dataset: Special data feature," *Journal of peace research*, 2010, *47* (5), 651–660.

Rockmore, Marc, "The Cost of Fear: The Welfare Effect of the Risk of Violence in Northern Uganda," *The World Bank Economic Review*, 2017, *31* (3), 650–669.

Sala-i-Martin, Xavier, "The World Distribution of Income (Estimated from Individual Country Distributions)," Technical Report, National Bureau of Economic Research 2002.

_ , "The World Distribution of Income: Falling Poverty and ...Convergence, Period," *The Quarterly Journal of Economics*, 2006, *121* (2), 351–397.

Schutte, Sebastian and Nils B Weidmann, "Diffusion patterns of violence in civil wars," *Political Geography*, 2011, *30* (3), 143–152.

Silverman, Bernard W, *Density estimation for statistics and data analysis*, Vol. 26, CRC press, 1986.

Sundberg, Ralph and Erik Melander, "Introducing the UCDP georeferenced event dataset," *Journal of Peace Research*, 2013, *50* (4), 523–532.

Tapsoba, Augustin, "The Cost of Fear: Impact of Violence Risk on Child Health During Conflict," Technical Report, Households in Conflict Network 2018.

Taylor, Charles C, "Bootstrap Choice of the Smoothing Parameter in Kernel Density Estimation," *Biometrika*, 1989, *76* (4), 705–712.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug, "PRIO-GRID: A unified spatial data structure," *Journal of Peace Research*, 2012, *49* (2), 363–374.

Tonini, Marj, Mário Gonzalez Pereira, Joana Parente, and Carmen Vega Orozco, "Evolution of forest fires in Portugal: from spatio-temporal point events to smoothed density maps," *Natural Hazards*, 2017, *85* (3), 1489–1510.

Townsley, Michael, Shane D Johnson, and Jerry H Ratcliffe, "Space time dynamics of insurgent activity in Iraq," *Security Journal*, 2008, *21* (3), 139–146.

Weidmann, Nils B and Michael D Ward, "Predicting conflict in space and time," *Journal of Conflict Resolution*, 2010, *54* (6), 883–901.

Witmer, Frank DW, Andrew M Linke, John O'Loughlin, Andrew Gettelman, and Arlene Laing, "Subnational violent conflict forecasts for sub-Saharan Africa, 2015–65, using climate-sensitive models," *Journal of Peace Research*, 2017, *54* (2), 175–192.

Yanagizawa-Drott, David, "Propaganda and conflict: Evidence from the Rwandan genocide," *The Quarterly Journal of Economics*, 2014, *129* (4), 1947–1994.

Appendix

Figure A1: Partition of Ivory Coast during Conflict

A Importance of the Smoothing Parameter in KDE

To illustrate the importance of the smoothing parameter h in KDE, I draw 500 data points from a bi-modal distribution given by a mixture of 2 normal distributions with means 5 and 15 and standard deviation 3. The underlying distribution is shown in panel (a) of Figure A4. Panel (b) shows the estimated density with a small and large smoothing parameter h. The first graph still shows some spurious variations from the data while the second one over-smooths the distribution to the extent of almost not reflecting its bi-modal nature.

Figure A2: Spatial Distribution of Events in Ivory Coast by Year

1997

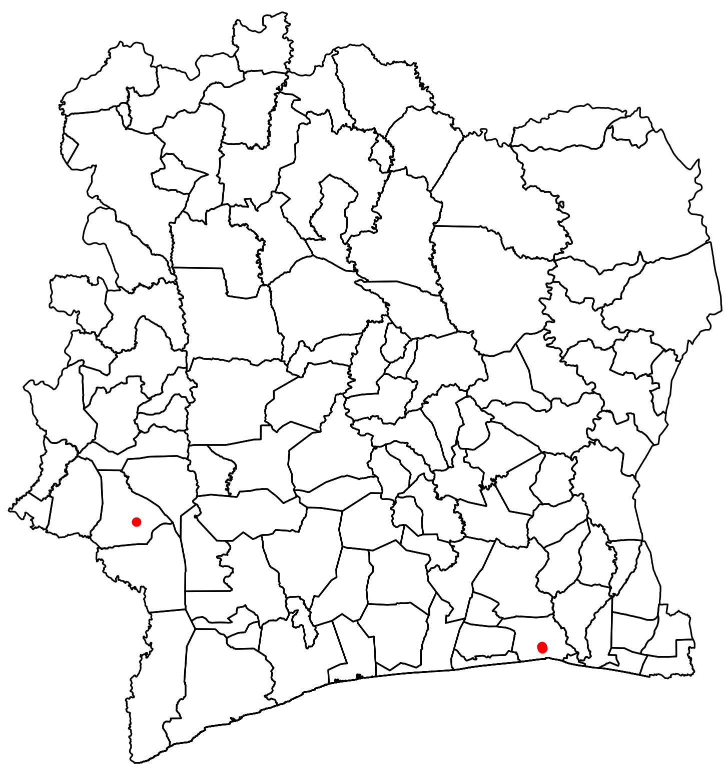


Figure A3: ROC Curves With Space-Time Lags: 50km by Month

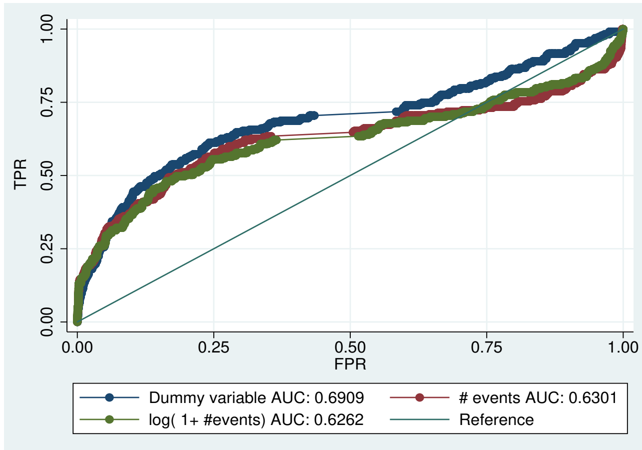
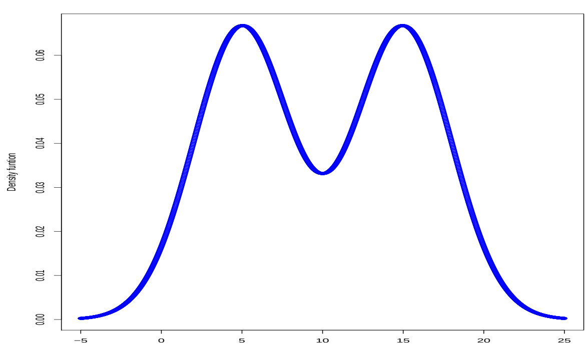
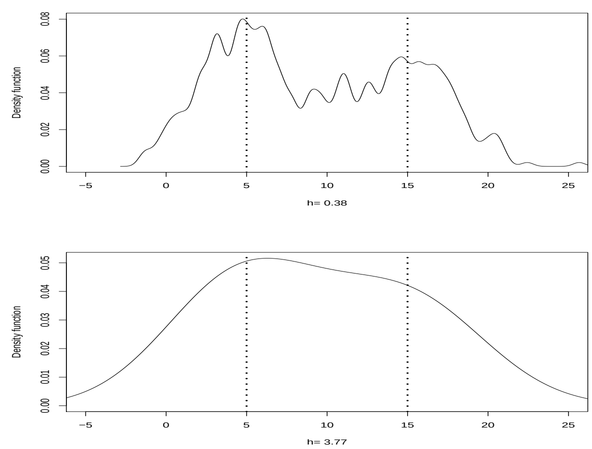


Figure A4: The Role of the Smoothing Parameter



(a) Underlying density



(b) KDE with small and large bandwidth

B Bootstrap

The idea of using bootstrap resampling to choose a smoothing bandwidth has been introduced by (Taylor, 1989). The bootstrap approach is used in conjunction with the $MISE(h)$ as a target criterion. The basic idea is to construct a "reference" density estimate of the data at hand, repeatedly simulate data from that reference density, and calculate the empirical integrated squared error at each iteration; doing so at different bandwidths. The bandwidth that minimises the bootstrap-estimated MISE is taken as the optimal value.

- Select a pilot bandwidth g and compute estimator $\hat{f}_g$ of f
- Draw bootstrap samples X_1^*, ..., X_J^* from $\hat{f}_g$

- Compute the bootstrap version of the $MISE$ and minimise it over h:

$$J^{-1}\sum_{j=1}^{J}\int[\hat{f}_h(y|X_j^*) - f_g(y|X)]^2dy$$

- Set new pilot bandwidth to the value h_0 that minimizes the $MISE$ and iterate untill it converges

C Parametrization of H

Given the relatively small number of realizations of the violence process in Ivory Coast compared to the estimation space (entire country for over 15 years), I parametrize H in the most simplistic way possible. I only consider one smoothing parameter h_s for both latitude and longitude and another smoothing parameter h_t for time.

In multivariate kernel density estimation, the parametrization of the matrix of smoothing parameters H is crucial. Diagonal elements correspond to the smoothing parameter with respect to each dimension and off diagonal elements capture the trajectory of the process. The violence process could spread more on one dimension (latitude) rather than the other (longitude). Setting off-diagonal parameters to zero implies that the process spreads in symetric way along both axis.

In theory the trajectory/symmetry of the violence process in space can also play an important role. To illustrate that, let's ignore time dimension and consider a 2 dimensional violence process that follows a bivariate normal distribution.

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 7 \\ 7 & 16 \end{pmatrix}\right]$$

Figure A5 shows the contour plots of the underlying density. I draw sample N=200 from this distribution and compare performance of kernel density estimation when allowing for full parametrization of matrix H or not.

Figure A6 shows that allowing for a full parametrization of H fits better the true density compared to the restricted parametrization. It also implies that everything else equal, locations that are on the trajectory of the violence process will have higher density estimate.

Figure A5: Density countour plot of bivariate normal distribution

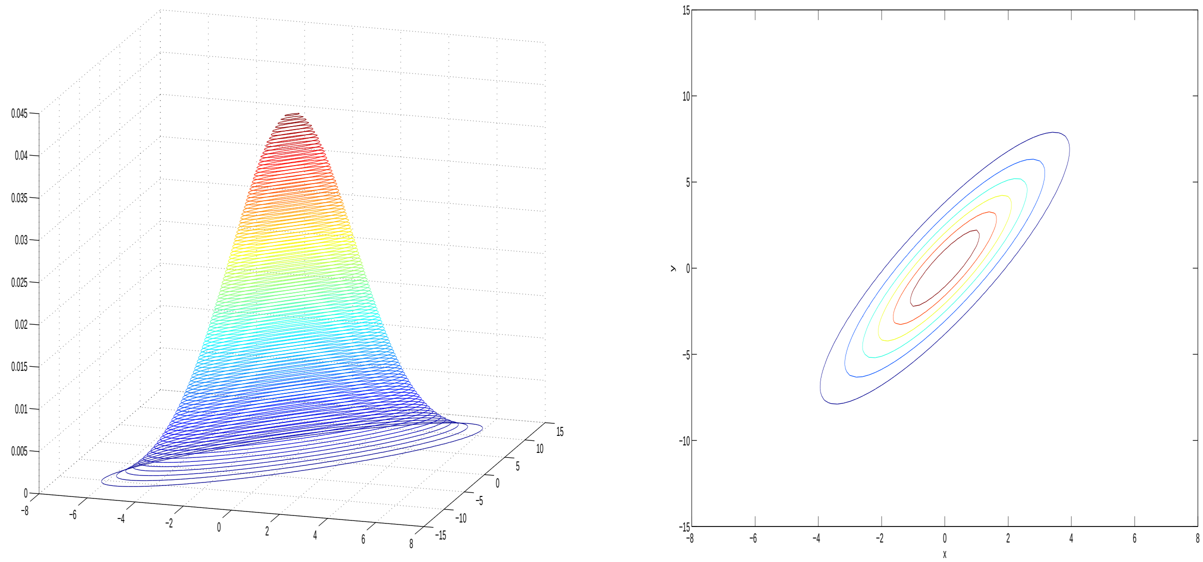


Figure A6: Contour plot KDE with full and diagonal matrix H