Masters Program in Geospatial Technologies

SPATIO-TEMPORAL ANALYSES OF THE RELATIONSHIP BETWEEN ARMED CONFLICT AND CLIMATE CHANGE IN THE EASTERN AFRICA

Riazuddin Kawsar

A dissertation submitted in partial fulfillment of the requirements
for the Degree of *Master of Science in Geospatial Technologies*

UNIVERSIDADE NOVA
UNIVERSITAT JAUME I
Supported by: Education and Culture ERASMUS MUNDUS

SPATIO-TEMPORAL ANALYSES OF THE RELATIONSHIP BETWEEN ARMED CONFLICT AND CLIMATE CHANGE IN THE EASTERN AFRICA

Supervisor:

Prof. Dr. Edzer J. Pebesma.

Institute for Geoinformatics (IFGI),

Westfälische Wilhelms-Universität (WWU),

Münster, Germany.

Co-supervirors:

Prof. Dr. Jorge Mateu M

Department of Mathematics,

Universitat Jaume I (UJI)

Castellon, Spain.

&

Prof. Mário Caetano

Instituto Superior de Estatística e Gestão de Informação (ISEGI)

Universidade Nova de Lisboa (UNL),

Lisbon, Portugal.

February 2011

&

Prof. Dr. Pedro Cabral

Instituto Superior de Estatística e Gestão de Informação (ISEGI)

Universidade Nova de Lisboa (UNL),

Lisbon, Portugal.

February 2013

Dedication

To my family

Acknowledgments

I am very grateful to my supervisor Edzer Pebesma, co-supervisors Professor Jorge Mateu for taking time off a busy schedule to supervise me at every stage of this project. Thanks to my co-supervisors Professor Mário Caetano and Professor Pedro Cabral of ISEGI-UNL, Lisbon for the face-lifting comments on the manuscript.

I am most grateful to the European Commission for, through the Erasmus Mundus scholarship scheme, making my studies and stay as comfortable as they were during these eighteen months. Thanks to the staff of the three partner universities - Institute for Geoinformatics Muenster, Universitat Jaume I, Castellon Spain and New University of Lisbon Portugal for ensuring a smooth run of the entire program.

I am indebted to the Research Groups Photogrammetry & Remote Sensing, Department of Geodesy and Geoinformation, Vienna University of Technology, especially Richard Kidd for allowing me to use data (Surface Soil Moisture) collected by the project. I also thank to Uppsala Conflict Data Program's (UCDP) for the Armed conflict dataset.

I am also grateful to Professor Pablo Juan Verdoy of UJI, Spain for helping me with R software.

Many thanks to my classmates who made the period of my studies an interesting ride. I sincerely thank my family and friends for supporting me and for giving me the vote when I decided to take up this Master program. Their prayers and support were invaluable.

Above all, I am thankful to God Almighty for His amazing grace. At every point of the work, I really felt my prayers answered.

And to the many others who gave valuable input to this work in whatsoever form but whom I have not mentioned, I mean no ingratitude. I still have you at heart and I sincerely thank you.

Declaration of Originality

I declare that the submitted work is entirely my own and not belongs to any other person. All references, including citation of published and unpublished sources have been appropriately acknowledged in the work. I further declare that the work has not been submitted to any institution for assessment for any other purpose.

Münster, 28th February 2013

Riazuddin Kawsar

SPATIO-TEMPORAL ANALYSIS OF THE RELATIONSHIP BETWEEN ARMED CONFLICT AND CLIMATE CHANGE IN THE EASTERN AFRICA

Abstract

Despite recent methodological improvements and higher data availability, the Climate Change (CC) and Armed Conflict (AC) studies are suffering from poor data and inappropriate research designs (e.g., Incompatibilities of scale). This study fills the gaps by taking the climate conflict analyses into a different scale (e.g., 55 km x 55 km sub-national cell/year) and uses high resolution Geo-referenced data sets. This study presents the results from 10 years (1991-2000) of observations and a rigorous modelling methodology to understand the effects of climate change on the conflict occurrence in the Eastern Africa. The main objective of the study is to identify and understand the conflict dynamics, verify the pattern of conflict distribution, possible interaction between the conflict sites and the influence of climatic covariates of conflict outbreak. We have found that if the climate related anomaly increases, the probability of armed conflict outbreak also increases significantly. To identify the effect of climate change on armed conflict we have modeled the relationship between them, using different kinds of point process models and Spatial Autoregressive (SAR) Lag models for both spatial and spatio-temporal cases. In modelling, we have introduced one new climate indicator, termed as Weighted Anomaly Soil Water Index (WASWI), which is a dimensionless measure of the relative severity of soil water containment indicating in the form of surplus or deficit. In all the models the coefficients of WASWI were found negative and to be significant, predicting armed conflict at 0.05 level of significance for the whole period. The conflicts were found to be clustered up to 200 kilometers and the local level negative relationship between conflict and climate suggests that change in WASWI impacts changes in AC by -0.1981 or -0.1657. We have also found that the conflict in the own cell associated to a (app. 0.7) increase in the probability of conflict occurances in the neighbouring cell and also to a (app. 0.6) increase of the following years (spatio-temporal). So, climate change indicators are a vital predictor of armed conflict and provides a proper predictive framework for conflict expectation. This study also provides a sound methodological framework for climate conflict research which encompasses two big approaches, point process modelling and lattice approach with careful modelling of spatial dependence, spatial and sptio-temporal autocorrelation, etc.

Keywords

- Armed Conflict
- Climate Change
- Spatial Point Pattern analysis
- Spatial distribution pattern
- Spatio-temporal modelling
- Spatial Autoregressive model
- Climate conflict relationship
- R

Acronyms

- AC - Armed Conflict
- AI - Area Interaction
- AIC - Akaike's Information Criterion
- CC - Climate Change
- CIESN - Center for International Earth Science Information Network
- CSR - Complete Spatial Randomness
- ERS - European Remote Sensing
- GDP - gross domestic product
- GED - Geo-referenced Event Dataset
- GPW - Gridded Population of the World
- ICP - Inhomogeneous Cluster Process
- ICTP - Inhomogeneous Cluster Thomas Process
- IMCP - Inhomogeneous Matern Cluster Process
- IPCC - Intergovernmental Panel on Climate Change
- IPP - Inhomogeneous Poission Process
- OLS - Ordinary Least Square
- PPP - Point Pattern Process
- SAG - Space Advisory Group
- SAR - Spatial Autoregressive Model
- SCM - Spatial Cluster Process model
- SDM - Spatial Durbin Model
- SEM - The Spatial Error Model
- SPI - Standerized Precipitation Index
- SSA - Sub Saharan Africa
- SSM - Surface Soil Moisture
- STIKhat - Space-time inhomogeneous K-function
- SWI - Soil Water Index
- UCDP - Uppsala Conflict Data Program
- WASWI - Weighted Anomaly Soil Water Index

Table of content

List of figures

List of tables

1 Chapter 1: Background

1.1 Armed conflict

Armed conflict can be termed as a conflict between or among several parties which involves armed forces. In 2001 Wallenstein and Sollenberg redefined the definition of armed conflict. According to them, "an armed conflict is a contested incompatibility which concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths." Armed conflict can be categorized into several categories based on their magnitude, involving parties, duration of conflict etc. but Uppsala Conflict Data Program (UCDP) categories the organized violence into three major categories in their dataset, Geo-referenced Event Dataset (GED). (Strandow et al. 2011). The categories are (1) state-based armed conflict, (2) non-state conflict and (3) one-sided violence. UCDP had compiled and coded information of organized violence's in event form for all three conflict types, covering the entire time period 1989-2010 for the African continent. The dataset defines armed conflict or events as, "…The incidence of the use of armed force by an organized actor against other organized actor, or against civilians, resulting in at least 1 direct death in either the best, low or high estimate categories at a specific location and for a specific temporal duration" (Strandow et al. 2011). Each event was appended with additional information like the date, scale, perpetrator etc. Different types of events differ in the aspects such as duration, temporal precision and continuity in armed violence. From 1989 to 2010, they have recorded around 22,000 events. For this study we have only considered all kinds of continuous violence for the period 1991 to 2000, which includes 3289 events in our study area in the Eastern African continent (see Figure 1.1 in page no 10).

1.2 Climate change

The temperature increase is not only warming the world but also deciding the fate of human being associated with the implications of warming the earth's surface. So global climate change became a very popular topic in the international research community. Impact of climate change is now evident and water is at the heart of it.

Climate anomaly driving the world towards the days with enormous water stress that determining the agricultural productivity. A vast majority of the people of developing countries depends on rain fed agriculture . So the relationship between the climate change, resource scarcity and the impact on human life, in the part of the developing world is very clear. This relationship can explain some other impacts too. Such as the impact of climate change on armed conflict. In several environmental

security literature, we can find that the access to the natural resources is a major predictor of armed conflict (Homer- Dixon 1991, 1999; Kahl 2006). Study on long term trend in temperature and precipitation change in light of human security can revile the notion. Understanding the impact of climate change on human security can lead us towards better conflict prediction by reconciling climate change and environmental security in the same ground.

Global warming is likely to affect the water availability pattern by affecting the precipitation pattern and that is pushing us to the ground of unpredictability of extreme events and this kind of situation may have implications for peace and security. So, understanding the climate change and its effect on life, cold be a very important study topic.

1.3 Armed conflict and climate change

Scarcity of resources such as minerals and water and conflict over that scarce resource is an old source of armed conflict. Resource scarcity will be intensified by environmental degradation and therefore will contribute to an increase in armed conflict (Gleditsch et al. 2011). Different authors argued otherwise on this topic. Fearon and Laitin (2003) argue that the probability of conflict can be increased by poverty since poor states have a much weaker financial and bureaucratic basis, providing an opportunity for riot. Besides poverty, low economic growth and high dependence on primary commodity exports are also important predictors of civil war. Then again, ethnic and religious diversity as well as democracy may not affect the risk of war (Collier and Hoeffler 2004). On the other hand, Hegre et al. 2001 found that regime type and ethnic heterogeneity matter have a greater impact on development. So, most of the studies on armed conflict have identified several economic, social, demographic or political factors as the main indicators of armed conflict until recently. In the first quantitative study of environmental conflict, Hauge et al. 1998 have found that economic and political factors were the strongest predictors of conflict but that environmental and demographic factors did have some impacts too. The world is generally becoming more peaceful but the debate on climate change presents the climate change as a potential threat to a new source of instability and conflict (Gleditsch et al. 2011).

The effects of climate change are frequently assumed to lead to loss of livelihood, economic decline, and increased insecurity either directly or indirectly (e.g., through forced migration). Interacting with poor governance, societal inequalities, and a bad neighborhood, these factors in turn may promote political and economic instability, social fragmentation, migration, and inappropriate responses from governments and the interplay of these factors can trigger conflict as well.

Several studies have found the potential link between climate change and armed conflict. Different authors argued that climate change is not the only factor of armed conflict but it may make the situation more tense coupled with some other economic, social or political factors. For example, Gleditsch et al., 2011 argued that reduced rainfall and higher temperature that jointly causes droughts and that reduces the access to the natural capital what sustains livelihoods. As a result, existing poverty will be more widespread and this kind of property situation and crisis are potential sources for a greater conflict.

Some more recent studies of temperature variation and conflict (Burke et al., 2009, 2010) claimed to find a link between temperature and civil war in Sub-Saharan Africa for the period 1981–2002 and argued that over a 35- year period climate change has contributed to a major increase in the incidence and severity of civil war in the region. Then again, Bernauer et al., 2012 have applied the temperature and precipitation deviations as a function of economic growth and thereafter have seen that, these variations in climate variables can predict onset of civil conflict in a non-democratic settings. Miguel et al. (2004) also has argued that anomalies of rainfall can be considered as another factor for armed conflict because he has found a relationship between negative rainfall deviation and increased risk of civil war in Africa. Again Miguel (2005) has found that both positive and negative extremes in rainfall increased the frequency of conflicts and has killed a lot of people in a rural Tanzanian district. On the other hand, D'Exelle and Campenhout (2010) have found water scarcity to drive conflicting behavior, particularly so for poor and marginalized households. Several statistical studies of conflict in Africa have found social violence and communal conflict to be most likely in or following wet periods (Raleigh and Kniveton 2012; Theisen 2012). The extreme events of natural disaster are another implication of climate change and one study using survey material on Indonesia finds villages that had suffered a natural disaster during the preceding three years to be more likely to experience violent conflict (Barron et al. 2009). Then again, sea-level rise, which is another implication of climate change, will threaten the livelihood of the populations of small island states in the Indian Ocean, the Caribbean, and the Pacific. Studies that look at non-random sets of cases with out-migration in areas with severe environmental degradation provide suggestive evidence that climate change could trigger more human mobility and that is also a source of conflict (Reuveny 2007; Reuveny and Moore 2009).

So, linking climate change and armed conflict can lead us towards a better understanding to find a way to a peaceful world by predicting the future armed conflict scenarios and taking action right now. Maybe, environmental or climate indicators and demographic stress are not likely to be an equally important risk factor but climate change can act as a multiplier, which needs special attention from now on. This study is an attempt to find out the potential link between climate change and

armed conflict for better understanding which may lead the modeler to a more accurate conflict scenario simulation.

One of the major criticisms of the early environmental security literatures are, those studies tended to neglect important political and economic context factors (Gleditsch 1998), but they are connected to each other. This study is also an attempt at the reconciliation of environmental with limited socioeconomic factors to draw a more justifiable conclusion.

1.4 Why this study?

Economical and political factors are very important for the armed conflict study (Fearon and Laitin 2003). In several studies it's been observed that Poverty increases the likelihood of the war and climate change is putting further pressure on the significance of the poverty level. So, several researchers are considering climate change as a factor or economic growth reduction and negative economic growth as a factor of armed conflict. On the other hand, in some other studies it's been evidenced that when we control the income, ethnic and religious diversity it does not increase the risk of conflict (Collier and Hoeffler 2004). Last few years, the researchers found a new dimension to observe the armed conflict linking climate change and its effect perspective but most of the researches are suffering from different kind of deficient. For an instance, the Case-based Environmental Security literature contains several narratives of violent conflict within the context of resource competition and environmental degradation but suffers from proper methods. Some of these quantitative researches found a potential scarcity-conflict connection but suffer from poor data and inappropriate research designs. There are some potential researches identified above give some indication that climate change increase risk of armed conflict but there are disagreements too. Another limitation of the previous studies can be identified as Incompatibilities of scale as several studies are conducted which focuses on countries not on subnational level, whether it's been pointed that most of the armed conflict cases are a local phenomenon. Besides, spatial statistics wings provide us extraordinary tools and methods to model the relationship between armed conflict armed conflict climate change from spatial time series analysis but almost no spatial time series analysis has been done in such study or suffers from poor data and inappropriate research designs. So it can be concluded that there are still no concrete solutions, which give modellers an idea how to incorporate climate change in armed conflict modeling or the question is still unanswered, how good climate change indicators would be to predict armed conflict?

Then again according to the climate scientist the effect of climate change is going to be worse in the near future and becasue of that it's important to know the dynamics of armed conflict and their relation with climate change what can help us to decide

tomorrow's word peace today. Modelling the relationship between armed conflict and climate change for future scenario prediction can be the start point of the nexus to clime walk towards peace by taking a right decision for the particular space and particular time. This study has tried to fill some of the identified research gaps in the study arena of climate change and armed conflict.

1.5 Data in such study

This study is mainly concentrated on finding the potential link between climate change and armed conflict for better understanding of the complex human environment interaction in the light of armed conflict. By modelling the relationship between armed conflict and climate change we can derive the mathematical link, which might help the researcher to predict the dynamic of armed conflict which are induced by the change in climate.

The instability of a country is the indicator of potential armed conflict which depends on some explanatory independent variable and their link with the risk on the armed conflict. To measure the instability of a country the explanatory variables used in several researches in this arena are basically some structural facts which can be identified as Economic indicators (gross domestic product (GDP) and its growth rate), Socio-demographic (total population, population density, youth bulge, school enrolment, infant mortality rate, ethnic polarization, regional polarization), Resources and their distribution (inequality: GINI Index, natural resource: ratio between primary commodity exports and GDP), Geographical context (Proximity and nature of the border, Terrain characteristics), Regime (level of democracy), Development Indicators (export, import, investment, foreign debt etc.), Climate change factors (precipitation, temperature, sea level rise and extreme disaster events), History of armed conflict (armed conflict event and location) etc. (Burnley et al. 2008).

Due to time constraint, we could not include all the possible range of variables in our modeling procedure as the data processing takes enormous time. So we had to consider very small number of variables of those some are already being used in several climate conflict literatures. For an instance, the number of population which is the most used demographic variable and explains the armed conflict mostly as most of the socioeconomic variable are highly correlated to the population. Besides population, in our study we tried to find some more potential explanatory variables which incorporate maximum possible dimensions of the climate change. In this study we have introduced a new climate variable termed s Weighted Anomaly Standardized soil water index (WASWI) which is a dimensionless measure of the relative severity of soil water containment indicating in the form of surplus or deficit (for detail see in the methodology data section). From the soil moisture literature we

have found that soil moisture depends on the climate change indicator precipitation, temperature, evaporation etc. and can cause extreme events such as draught (for detail about WASWI preparation see the Data section). For an instance Lakshmi et al. (2003) has molded the relationship between the surface temperature and soil moisture and showed that increase of temperature corresponds to a decrease in the soil moisture. So soil moisture level can depict the temperature scenario for a particular region. Then again Xu et al. (2012) have shown the relationship between soil water and rainfall. So, WASWI can be a good indicator of rainfall and temperature though there are other factors related to soil water containment such as soil type. Again, some study provides a clear picture about the relationship between climate change, soil moisture and agricultural production such as Tao et al. (2003) and which corresponds to our study area characteristics and so WASWI is a better explanatory climate variable for the study area like Sub Saharan Africa (SSA). So this WASWI can be considered as a good indicator of climate change as studies have identified SWI as an indicator of climate change already (Seneviratne et al. 2010). WASWI can offer a better picture of climate change as it a factor which not only consider precipitation or temperature alone but also the interplay of these factors. Besides WASWI we also have incorporated other climate change indicators which are good for draught study and already used in climate conflict study (Theisen et al. 2010) is Standardized Precipitation Index (SPI), which is considered as a well-known probabilistic measure of the severity of a dry event (Guttman 1999) and that explains the physical and climatic situation of our study area well (Collier et al. 2008).

1.6 Study area

Almost every country has experienced armed conflict in several locations and in different time periods. To get a more accurate picture it is recommended to do such study on the whole world with a large range of explanatory variables. But due to time, resource and data limitations we needed to consider a small portion of the SSA for this study. Choosing a study region was tricky and again easy enough. It was tricky because a small portion is not enough to draw a relationship between conflict and climate for the whole world as the space is continuous and heterogeneous in characteristics both in the physical and socioeconomic sense. And it was easy to decide because of available data sets. In several literatures it's been termed that Sub Saharan Africa (SSA) is the place on earth which experienced most of armed conflict events and therefore potential for such study. Then again being in the lower latitude the climate change effect will be more moderate than most other locations. So we decided to choose the Eastern part of Sub Saharan Africa as our study area.

One third of African population live in the drought-prone regions and only 4% of arable land in SSA is irrigated so the people have to depend on the rain for their agricultural production. Most of the population can be classified under subsistence

economies as this continent is one of the largest agricultural sectors of the world. Global warming is likely to lead to a drying of most of the Africa. It is generally accepted that Africa will be affected by future global climate change first and most severely. (Low 2005; Collier et al. 2008).

From the figure 01 we can see that most of the armed conflict event over the time period 1991 to 2000 are distributed in the eastern part of the Africa. Then again the study window in Blue boundary can be characterized as a diversified area because of its geographical feature distribution (e.g., water body, dry region, central African forest etc.) as well as heterogeneous distribution of armed conflict (e.g., zone experienced Armed Conflict and zone without experiencing armed conflict).

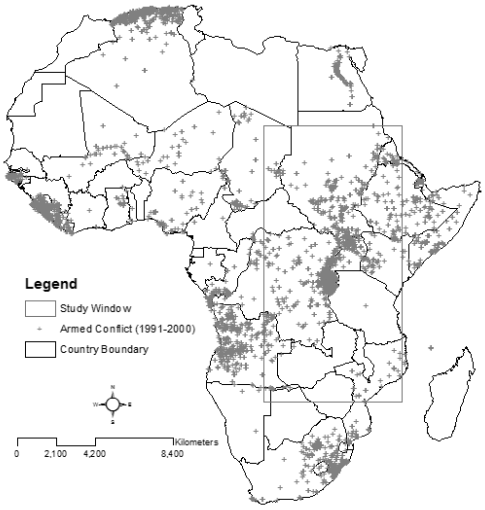


Figure 1-1 the distribution of the Armed Conflict in Africa (Strandow et al. 2011)

The total land area of Africa continent is approximately 29063244 sq.km and our study window is about 8719926 sq.km which means more than 30 percent area are covered by our study window. Then again based of the UCDP armed conflict dataset named geo-referenced Event Dataset (GED), over the period 1991 to 2000 the African Continent experienced around 10708 numbers of armed conflict of what 3289 numbers of armed conflict are inside the study window. Because of the characteristics, the number of conflict event distribution and efficiency of analysis, we have considered the blue banded box as our study window.

1.7 Methods used to support the climate conflict relationship

As identified above there were several attempts to establish the relationships between climate and conflict and those studies have used a range of statistical methods but out of them a significant number of the studies have involved different forms of logistic regression. (Carter and Signorino 2010, Levy et al. 2005; Nel and Righarts 2008; Raleigh and Urdal 2007), whether some others have used multivariate regression

models (Theisen et al. 2010). Some other have adopted linear regression with country fixed effects and time trends (Buhaug 2010). There some other studies too which involved some other method besides logistic regression. For an example Raleigh and Kniveton (2012) has used Negative binomial count regression and composite analysis (or 'epoch superposition') methodology and Miguel et al. (2004) has used a nonparametric version of the local regression method with an Epanechnikov kernel etc.

Continuous development in probability wing Point Process modeling has provided us with more sound techniques to model and predict point process (e.g., event data can be considered as a point process according to the definition of point process. For detail see chapter 3). For an instance Zammi -Mangion et al. (2012) while studying the dynamic of war in Afghanistan has involved dynamic spatio-temporal modelling from point process theory. Until now that was the first involvement of point process theory in the Armed Conflict study but only to model the war dynamics no to study the climate conflict relationship. In our study we tried to model armed conflict (events) and climate change related information from spatial patterns of events; insights from point process theory. We have used five models in several phases. Poisson Process model, Thomas Cluster Model, Matern Cluster Model, Area Interaction Model for both spatial and spatio-temporal process. Space-time inhomogeneous K-function (STIKhat) has been used to assess the dynamic spatio-temporal point process modelling. Besides point process modeling we have also considered lattice approach and have completed some spatial regressive models like Ordinary Least Square (OLS) model, Spatial Autoregressive (SAR) model, Spatial Error Model (SEM) and Spatial Durbin Model (SDM). Then again we have computed spatio temporal models devising SAR model for spatial, temporal and spatio-temporal neighbors.

1.8 Research questions

1. Is there any link between climate change and conflict?
2. Is local resource scarcity (e.g., soil moisture) in terms of climate change offers a better prediction of conflict behavior?
3. Is the climate problem may arise and persist locally?
4. May sub national disaggregated studies provide more support for the resource conflict nexus?
5. If there is any link between climate change and armed conflict, then how good climate change indicators are predicting armed conflict?

So this study has addressed the relationship between local conflict and local climate behavior by modelling the relationship between climate and conflict using a combination of the two approaches and those are point process analyses and lattice

analyses. Here is our basic assumptions were the distribution of conflict location are inhomogeneous due to the varying distribution of climatic factors. For an instance, areas with lower levels of water contained in soil and higher number of population are more prone to conflict.

1.9 Research design

- **Ambition:** modeling the relationship between armed conflict and climate change and analysis of the central environmental security proposition that Climate Change increases the local risk of civil armed conflict
- **Sample:** Eastern region of Sub Saharan Africa (SSA) 1991 – 2000 (10 years)
- **Unit of analysis**
    - Spatial: 0.5 degree latitude x 0.5 degree longitude grid cell observations. In this part of the world each side of the cells corresponds to approximately 55 kilometers
    - Temporal: year
- **Dependent variable:** Geo-referenced armed conflict occurrences which have >25 battle-deaths threshold.
- **Independent Variable**
    - Weighted Anomaly Soil Water Index (WASWI)
    - Soil Water Index (SWI)
    - Standardized Precipitation Index (SPI)
    - Number of Population etc.
- **Method**
    - Different models of point process theory
    - Different models from lattice approach

1.10 Research writing organization

This report is organized in 6 major sections with several subsections, which follow a chronological flow of ideas as follows

In the background (chapter 1) section we have talked about the relationship between climate change and armed conflict form environmental security literature and also talked about the previous study attempts, their limitations, about the data and methods used in such study etc. we have also included our research question and research design in chapter 1.

In the methodology and data section (chapter 2) we have talked about our study methods, data preparation and some descriptive study to design our modeling approach

The point process modeling (chapter 3) explains the point process models used in this study to understand the interaction and covariate effect on the conflict distribution and also presents detail modeling approach and results of the point process models.

The lattice approach (chapter 4) explains the spatial and spatio-temporal lattice approach modeling and their result.

In the discussion (chapter 5) part, we have talked about the different model output and relative analysis of the results and answer the research questions. Finally, in chapter 6 we have also talked a little bit about the limitation of our study and future study recommendation.

2 Chapter 2: Methodology and Data

2.1 Methodology

To model the relationship between armed conflict and climate change, our first step was to decide the list of covariates and independent variables, which can explain the relationship between our independent variable and covariates. As this was a spatio-temporal study, we also had to put some extra attention on choosing the covariates because of data unavailability for different time period and extent of our study area.

It was also important to choose a particular spatial area for the study based on data availability for different spatial and temporal period. It was also a major concern about the size of the area for efficient computation, due to time limitation. For our study, based on the data available and efficient processing and computing capability we have chosen the eastern part of the African continent.

In the next step we have constructed a dataset which has the structure of a raster grid. Our unit of observations is subnational "cells" of 0.5 degrees of latitude x 0.5 degree of longitude. All the data were being processed according to that particular resolution.

Our empirical analysis was conducted in the cell and cell by year level. In this study our main dependent variable is events, an integrated measure of conflict indicating the total number of any kind of conflict (have >25 battle-deaths threshold) indicating whether the cell has experienced a conflict related episode of any of the categories included in the UCDP GED dataset over the period of the year 10 years from 1991 to 2000. And our covariates were the number of the total population per cell, The Standardized precipitation Index (SPI) per cell and Weighted Anomaly Standardized soil water index (WASWI) per cell etc.

In order to investigate the local level relationship between climate change and the armed conflict incidence we have estimated several models from point process theory and from lattice approach.

2.2 Point process analyses

From point process theories we have estimated four models for spatial patterns of events analysis and to estimate the covariate effects on event distribution.

Our first model is an Inhomogeneous Poisson Process (IPP) model explained in chapter 3. The second and third model can be termed as Spatial Cluster Process (SCM) model. From SCM we have computed Inhomogeneous Thomas Process

(ITCP) and Inhomogeneous Matern Cluster Process (IMCP). And our final model for point process is an Area Interaction (AI) model which not only considers the covarite effect but also inter point interaction to define an intensity function for point distribution in a study area (for detail see chapter 3). For this entire model we have used the inhomogeneous version of K-Function (for detail chapter 3) to fit our empirical data with theoretical lines of different models.

The behaviors of a point process can be explained through trend (caovariate effect) and dependence (interaction) between the points of a point pattern. The appearance of such interaction or trend consists of either clustering or regularity in the process. A widely used tool for exploring the nature of interaction is a Ripley's K-function (Ripley 1976; Diggle 2003; Cressie 1993). In our study we use an Inhomogeneous version of L-Function to interpret an Inhomogeneous version of K-Function.

In the first part of the point process modelling, we have excluded the temporal dimension of all of the data by aggregating all data into one spatial layer. So, resulting models are based on aggregated event data for the period 1991-2000 as main dependent variable and aggregated WASWI, SPI and number of population as covariates.

In the second part of the study we have fitted all of these four kinds of model with empirical yearly data for 10 years (1991-2000) which explains the relationship between armed conflict and climate by year. while mdelling, besides yearly covariates we have also considered covariates of different temporal lag such as (t-1) and (t-2)

In the third part of the study the Second-order properties are used to analyze the spatio-temporal structure of a point process. The space-time inhomogeneous K-function are used as a measure of spatio-temporal clustering or regularity and as a measure of spatio-temporal interaction (Gabriel and Diggle 2009; Moller and Ghorbani 2012, Illian et al. 2008).

2.3 Lattice approach

Besides the understanding attempt from the point process perspective we have also confirmed the lattice approach for more detailed understanding of the relationship. In the lattice approach we have considered models from two different wings. At the very beginning we have tried Spatial Ordinary least Square (OLS) regression model, which we have termed as model 1.1. Computing ordinary least squares (OLS) regression analyses we tried to model the relationship between the conflict and climate variables, where the basic assumption is all the observations are independent in space which is highly unrealistic. To accommodate the gap we have performed

Simultaneous Autoregressive Lag model (SAR) analyses by considering autocorrelation of the dependent variables in space which was done by using a spatial weight matrix (e.g., considered queen neighbors) (model 1.2). We have also incorporated two other models from Spatial Autoregressive model wing. Those are The Spatial Error Model (SEM); termed as model 1.3, Spatial Durbin Model (SDM); termed as model 1.4 (for detail see chapter 4, spatial modelling section).

In the next step of the lattice approach we have integrated the temporal dimension into the analysis by defiing the spatial, temporal and spatio-temporal neighbors in SAR model (for detail, see chapter 4 spatio-temporal modelling section). We have modified the SAR model for temporal neighbors which has given us one single autocorrelation coefficient to define the correlation both in space and time; we have termed this as model 2. We have also devised the model 2 to fit another single correlation coefficient to describe correlations between all (spatial, temporal, and spatio-temporal) neighbors. For all these models we have also calculated the Nagelkerke R-squared to understand the impact of covariates on our dependent variable, Armed Conflict.

2.4 Data

2.4.1 Sources and dataset construction

For this study we bring together georeferanced data from a variety of sources and constructed a dataset which cover almost 16 countries either partially or completely. The study area consists of the eastern part of Central African Republic, Sudan, Zaire, Eastern part of Angola (app. 25 %), Zambia, northern part of Zimbabwe (app. 75 %), northern part of Mozambique (app. 75 %), Malawi, Tanzania, Burundi, Rwanda, Uganda, Kenya, western part of Ethiopia (app. 60 %), western part of Eritrea (app. 75 %) and a small portion of Botswana. The data set contains the information of every location (cell) over the period 1991-2000, which includes the information of individual conflict episode locations. We have also collected, computed and processed the detailed data on SWI, WASWI, SPI and the number of population. All these data are processed according to 0.5 degree latitude x 0.5 degree longitude degree raster grid.

2.4.1.1 Armed conflict

According to the event definition in Sundberg et al. 2010, Uppsala Conflict Data Program's (UCDP) has been developed the comprehensive Geo-referenced Event Dataset (GED) (version 1.5) on organized violence for the African continent over the time period 1989-2010. This data set consists of UCDP's categories of organized violence which are termed as state-based armed conflict, non-state conflict and one-

sided violence. This data set is the most disaggregated datasets which indicated the location with 10 meter accuracy. In this dataset, the number of deaths at the event location, start time and end time of the event, these kinds of attributes are attributed against the event location. For such local level study this dataset is being used in different study (Melander et al. 2011) and was able to explain the nature of conflict in SSA.

In our study data on armed conflict episodes over the period 1991-2000 are collected from comprehensive version 1.5 of Geo-referenced Event Dataset (GED) dataset developed by UCDP. In our study area, we have considered all kinds of conflicts recorded in GED datasets. The number of total events is 3289 where the total number of state-based armed conflict is 1558, 299 non-state conflict and 1432 one-sided violence. For computational effectiveness we have spread the overlapped points up to 50 meters for aggregated point pattern analysis and for lattice process analyses that was not necessary as the data were prepared in cell level based on event count.

2.4.1.2 Weighted anomaly standardized soil water index

Surface Soil Moisture (SSM): The Surface Soil Moisture (SSM) data from Research Groups Photogrammetry & Remote Sensing, Department of Geodesy and Geoinformation, Vienna University of Technology (TUWIEN). SSM is a time series of the topsoil which indicates a relative measure of the water content in the surface layer (<5 cm from the surface) ranging between 0 and 1. SSM data were derived from scatterometers on-board in the The European Remote-sensing Satellites (ERS-1 and ERS-2) by considering microwave frequencies 1-10 GHz domain as the dielectric properties of soil and water are distinctly different in these frequencies (Pradhan and Saunders 2011). The collected data resolution is like following

- Spatial Resolution 50 sq.km
- Temporal Resolution = daily

The IPCC Fourth Assessment Report states, European Environment Agency and so others also have proved and consider that Soil moisture is an important factor that influences the climate (Weaveret and Avissar 2001; Gregory et al. 1997; Boix-Fayos et al. 1998; Komescu et al. 1998)

But, an understanding of the water in the soil up to 1 meter will be more important than ever with a changing Climate where SSM only provides the water information of soil up to 5 cm. If climate change brings bigger rain events then we could be faced with an increased risk of saturated soils and erosion. Increased temperatures could mean that the soil dries out more often. We can, therefore, expect the soil to be

pushed harder in the years ahead. So here TUWINE comes with a solution by the computing Soil Water Index (SWI).

Soil Water Index (SWI): The retrieved SSM, being a topsoil signature, may change significantly within a few hours whose magnitude depends on the amount of rainfall, evaporation rate and the time lapse since the rainfall event. The Soil Water Index (SWI) for the top 1 meter layer thus estimated from the topsoil moisture content adjusted with precipitation, evaporation etc. (Pradhan et al. 2011). So, the retrieved information is generally in good agreement with general climate regimes and gridded precipitation data. (Scipal, 2002). SWI is generally in good agreement with general climate factors like Precipitation, temperature, evaporation and has been used in Several climate change Studies. For an instance for use of SWI for drought indices in climate change impact assessment (Mavromatis 2010), soil moisture datasets for unravelling climate change impacts on water resources (Wagner et al. 2011), water cycle changes and CMIP3 simulations (Mariotti et al. 2008), monitoring water availability and precipitation distribution at three different scales (Zhao et al. 2007) etc. A plot of yearly SWI is shown in figure 2-1, where the 0 value represents relatively (the value is spatially relative) dry region.

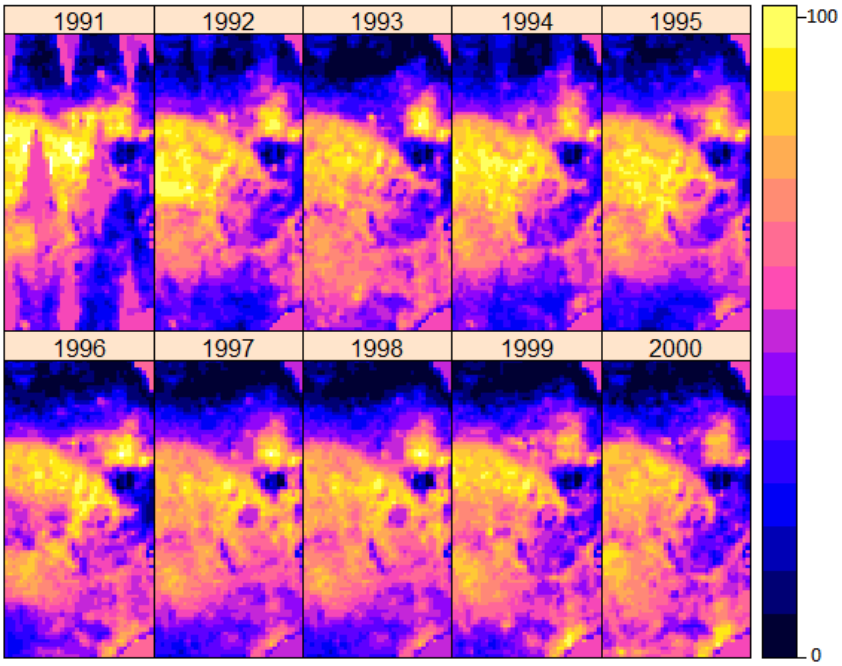


Figure 2-1 space time plot of Soil Water Index (SWI)

A global validation of the ASCAT Soil Water Index (SWI): soil moisture is an essential climate variable. Surface Soil Moisture (SSM) can be estimated from measurements taken by ASCAT on-board Metop-A and have been successfully validated by several studies (Albergel et al. 2009 and 2012; Parrens et al. 2012). Profile soil moisture rather ten SSM (of 5 cm depth) cannot be measured directly by remote sensing. The near real-time Soil Water Index (SWI) product, developed

within the framework of the Global Monitoring for Environment and Security (GMES) project geoland2 aims to minimize this gap.

SWI data from January 1st 2007 until the end of 2010 were compared to in situ soil moisture data from 420 stations belonging to 22 observation networks which are available through the International Soil Moisture Network. These stations delivered 1331 station/depth combinations which were compared to the SWI values. After excluding observations made during freezing conditions the average significant correlation coefficients were 0.564 (min -0.684, max 0.955) while being greater than 0.3 for 88% of all stations/depth combinations (Albergel et.al. 2009 and 2012; Parrens et.al. 2012).

WASWI estimation: Suppose $WASWI$ is the Weighted Anomaly Weighted Anomaly Standardized soil water index which represents a dimensionless measure of the relative severity of the Soil Water Index (SWI) surplus or deficit in a grid cell *x* and according to (Lyon and Barnston 2005) that can be defined as:

$$WASWI_N = \sum_{i=1}^{N} \left(\frac{(swi_i - \overline{swi_i})}{\sigma_i}\right) \frac{\overline{swi_i}}{\overline{swi_A}} \quad \ldots\ldots\ldots (1)$$

Where swi_i = is the observed value of SWI for the i^{th} month;

$\overline{swi_i}$ = represent long term (1991-2000) mean of monthly SWI for the i^{th} month;

σ_i = standard deviation of the anomalies of monthly SWI for the i^{th} month;

$\overline{swi_A}$ = mean annual SWI and

$\frac{\overline{swi_i}}{\overline{swi_A}}$ = Weighting factor representing the monthly fraction of annual SWI to reduce large standardized SWI anomalies that might result from small precipitation amounts or higher temperature and evaporation, occurring near the start or end of dry seasons and to emphasize anomalies during the heart of rainy seasons.

For our study, according to equation (1) we have calculated the WASWI for the month January like following

$$WASWI_{jan\ 1991} = \left(\frac{(swi_{jan\ 1991} - \overline{swi_{(jan\ 1991,\ jan\ 1992 \ldots jan\ 2000)}})}{\sigma_{jan\ 1991}}\right) \frac{\overline{swi_{(jan\ 1991,\ jan\ 1992 \ldots jan\ 2000)}}}{\overline{swi_{(Jan\ 1991,\ feb1991 \ldots dec1991)}}}$$

To make the WASWI data well-suited to our annual Armed Conflict (AC) data, we converted the monthly WASWI index into an annualized index. To do that, the WASWI for each month of a year, are computed and these weighted monthly anomalies are then summed over 12 months to get a 12-month (1.e., year 1991) WASW index. A plot of annualized index is shown in figure 2-1, where negative values and positive values indicate the unusually dry and unusually wet condition,

respectively. From the plot we can observe that in 1992, 1994 and 1998 some region of our study area have experienced relatively extreme dry condition.

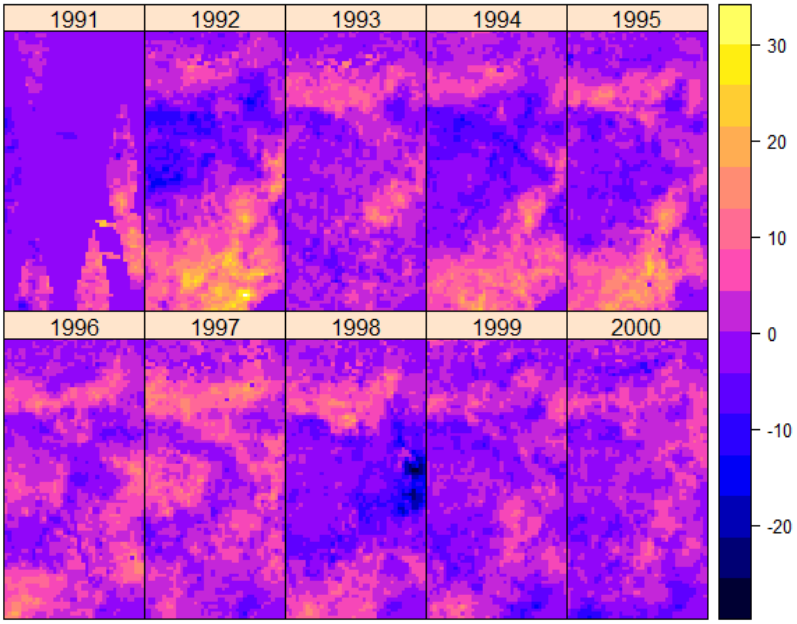


Figure 2-2 spatial time series of annualized Weighted Anomaly Weighted Anomaly Standardized soil water index (WASWI)

2.4.1.3 Standardized precipitation index

The Standardized Precipitation Index (SPI) is a probability index which has been developed by McKee et al. (1993 and 1995) to give a better depiction of irregular wetness and dryness than the conventional Palmer indices (Palmer 1965). The index is standardized by transforming into the probability of the observed precipitation, which enables all users to have a common basis for both spatial and temporal comparison of index values. SPI is a probability based invariant indicator of drought that recognizes the importance of time scales in the analysis of water availability and water use (Guttman 1999).

To calculate SPI first a probability density function which describes the long term time series of observed precipitation. In our case the series is for 1 year time duration. In the next step the cumulative probability of an observed precipitation amount is computed. Then by applying the inverse normal (Gaussian) function, with mean zero and variance one to the cumulative probability we can get the SPI for 1 year time duration. SPI values can be positive or negative where the magnitude of the departure from zero in negative direction considered as a probabilistic measure of the severity of a dry event.

For our study we have collected the data from "SPI-UEA_12-Month SPI-UEA_12-month from IRI Analyses SPI: Standardized Precipitation Index analyses of multiple global precipitation data sets." Where the grid size was 0.5 N. The data sets can be

downloaded for the whole world from this site. http://iridl.ldeo.columbia.edu/SOURCES/.IRI/.Analyses/.SPI/.dataset_documentation.html. To make the data compatible with our annual armed conflict data, we converted the monthly SPI index into an annual index. A plot of annualized SPI is shown in figure 2-2, where negative values and positive values indicate the unusually dry and unusually wet condition, respectively.

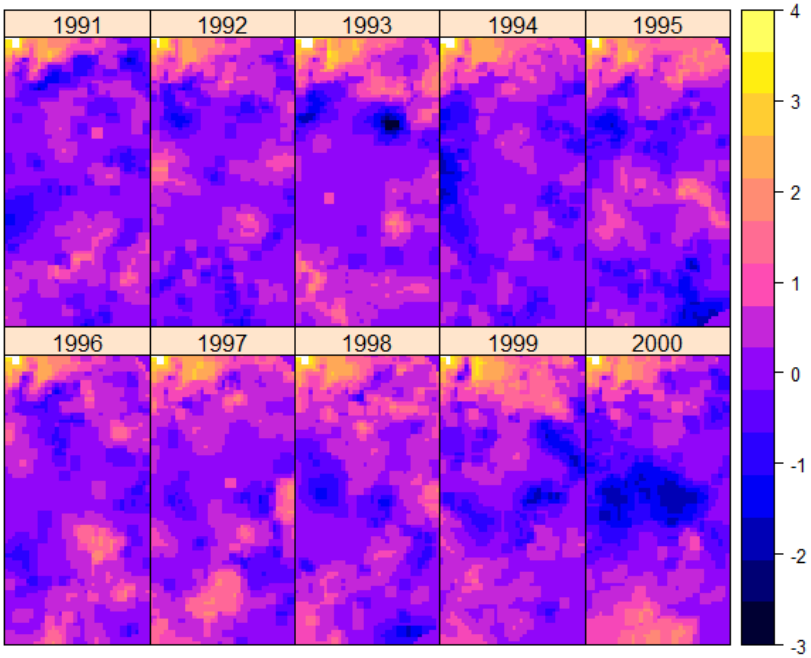


Figure 2-3 spatial time series of annualized Standardized precipitation index (SPI)

2.4.1.4 Population

Center for International Earth Science Information Network (CIESIN) has developed the datasets titled Gridded Population of the World (GPW), which is a collection of subnational administrative boundary data and corresponding population estimates for the world. The data set covers twenty year period, from the late 1980s to the present. The first version of GPW was based on 19,000 subnational units and version three incorporates more than 350,000 subnational units, each with georeferenced boundaries and at least one corresponding population estimate (Balk et al. 2003). The spatial resolutions of the data sets are 2.5 minutes latitude by 2.5 minutes longitude, which is approximately 21 sq. km at the equator. For our study purpose we have collected the data for available years for our study period (1991, 1995 and 2000) regrided the data into 0.5 degree latitude x 0.5 degree longitude grid.

2.5 Descriptive statistics

Our study area is highly dependent on the agriculture and agricultural growth depends on the growing season. If there is less water in soil during the growing

season there might appear an economic shock due to decrease in agricultural productive. By instinct we assumed that there might be some pattern in armed conflict which might be derived from the hampered agricultural growth. To understand this phenomenon we tried to look for some seasonal pattern in our study area for the period 1991 to 2000. From the plot 2-4 we can see that there were no seasonal pattern in armed conflict in the study area but we have observed the increase in the number of armed conflicts in that period (1991-2000). So we have decided to go with the yearly study then the seasonal study.

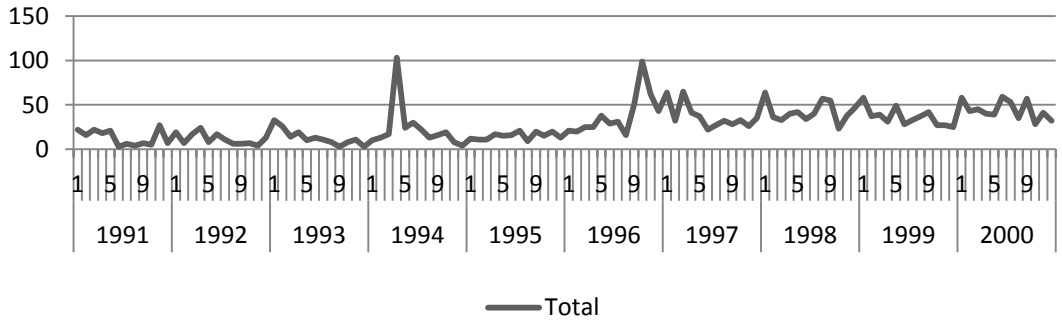


Figure 2-4 season pattern of armed conflict in the study region.

We have divided the whole study area into 3200 sub-region (cells of 0.5 degree longitude x 0.5lattitude). Table 2-1 presents the descriptive statistics of both dependent and independent variables in the study of cells. From the table we can observe that the mean armed conflict was higher in 1996 and 2000 and standard deviation also follows the trend. On the other hand WASWI was highest in 1995

Table 2-1 Descriptive statistics of Armed Conflict, weight) Anomaly Standardized soil water index (WASWI) and Standardized Precipitation Index (SPI)

Year	AC (Mean)	AC (SD)	WASWI (Mean)	WASWI (SD)	SPI (mean)	SPI (SD)
1991	0.08	0.59	3.37	5.85	0.12	0.68
1992	0.06	0.47	2.99	**7.95**	0.18	0.66
1993	0.08	0.67	1.32	3.59	0.33	0.79
1994	0.12	1.73	1.46	5.77	0.08	0.76
1995	0.08	1.00	**5.79**	2.52	0.18	0.84
1996	**0.20**	**1.91**	4.18	2.74	0.28	0.67
1997	0.20	1.45	3.57	4.18	0.41	0.71
1998	0.24	1.66	**-0.41**	5.80	0.33	0.74
1999	0.19	1.51	0.30	3.66	0.27	0.81
2000	**0.22**	**1.86**	0.35	3.41	**0.02**	**0.96**
Total	1.48	12.84	22.91	45.46	2.21	7.60

(SWI anomaly in a positive direction, means relatively more water in the soil but anomaly was higher as well) and in 1998 (SWI anomaly negative direction, means less water in the soil). The lowest mean value of SPI was observed in 2000 and highest in 1997.

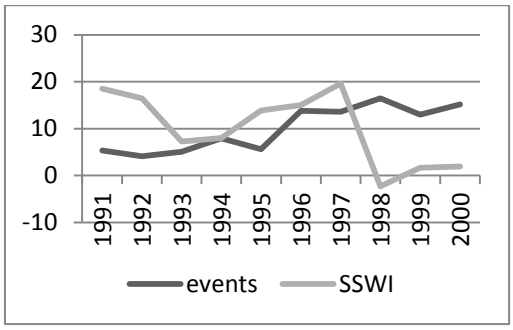


Figure 2-5 standardized trend plot of Armed Conflict and WASWI

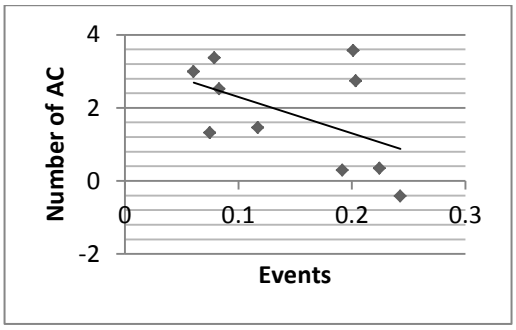


Figure 2-6 regression line of AC and WASWI

Following the aim of our study we have considered WASWI as our main independent variable which considered as a climate change indicator and armed conflict as the dependent variable. We have plotted regression line in figure 2-5 and 2-6. From the figures we can see that the events and WASWI have a negative relationship. In the period 1991 to 1994 the value of WASWI decreased and the number of armed conflicts increased. But in 1995 it's the contrary. Then again the first statement is also true for the period 1998 to 2000. So the relationship doesn't follow a linear trend. We have also plotted the regression line between these two variables in figure 2-6. There we can observe the negative relationship more clearly and the relationship can be expressed through $y = -9.9635x + 3.2927$ where $R^2 = 0.2508$.

3 Chapter 3: Armed Conflict and Point Process Modelling

A point process "...a stochastic process in which we observe the locations of some events of interest within a bounded region A." (Bivand et al. 2007). In the definition the events refer to the actual observations of points, while the region A is usually considered as the window of observation (Baddeley and Turner 2006). In another word, a point process is a collection of random points in a n-dimensional space falling in any bounded set, where for spatial point process n =2 and for spatio-temporal case n =3 (Ripley 1952). According to the definition of point process, the location of armed conflicts is realized as a set of random points in a 2-dimensional space (spatial point process case) falling in our study window and 3-dimensional space for spatio-temporal point process case. Location (e.g., longitude and latitude of an event) of armed conflict can be characterized by a broad range of heterogeneous explanatory variables such as geographical, political and socioeconomic variables in several formats (spatial, non-spatial or spatio-temporal). And these make the modelling and prediction of conflict challenging due to heterogeneous and dynamic nature of the data available.

In this section, the goal of armed conflict location study based on point process theory to identify and understand the complex underlying process in conflict such as interaction, diffusion, heterogeneous growth and hotspot of armed conflict based on the location properties of covariates and interaction between the conflict events. Such study of conflict dynamics, insights from point process theory can provide a predictive framework which helps us to understand the dynamic process of conflict based on the dependence between points and covariates and it can also provide the level of confidence in terms of prediction. In this section, we have studied the spatial or spatio-temporal dependence between points, spreading phenomenon and transformation and covariates effect on event's distributions with statistical accuracy.

As the basic point process is a Poisson process (Ripley 1977), the starting point of any point process study is the homogeneous Poisson process or Complete Spatial Randomness (CSR) (Schabenberger and Gotway 2005), where the intensity in even across the study space. If the intensity does vary spatially then the process is referred to the inhomogeneous Poisson process.

In a point process dependence between points and covariates can be investigated by realizing the interactions in the form of independence, regularity and clustering. The form independence identifies the process as a Poisson process. In regularity case the points tend to avoid each other and clustering refers where points tend to be close to each other. Based on these distances methods several summary functions can be derived such as nearest neighbor distance function which examines the distance between each point and the closest points to it and then compares these to the

expected value of a point process (e.g., Poisson process). The empirical nearest neighborhood function (G-Function) explains the probability of an observed nearest neighborhood of a point appearing at any given distance, which describes the degree of clustering of regularity in any point process. There are several summary functions used in point process studies but all these summary functions (e.g., G-Function, F-functions) considers only the nearest neighbor for each event in a process which can be identified as major drawback but K-Functions are based on all the distances between events in a study region. So, as suggested in several studies, in our study we fit our point process model with the data with K-Function (Ripley 1977)

Simulation process lies in the heart of the point process study. After a model fitted to the data the simulation can be done and based on the simulation we can create the simulation envelops which is the basic tool to estimate the confidence in modelling and prediction (Schabenberger and Gotway 2005). In this study, the term event represents an armed conflict episode and point of sets refers to an arbitrary location

3.1 Intensity

One of the basic properties of point process modeling is intensity. Exploring intensity which also can be termed as the average density of points, we can estimate the expected number of points per unit. Then again, intensity may be homogeneous or inhomogeneous. In the first step of analysis we have investigated the intensity of events. In the study region the event intensity is 4.11 means the expected number of events in each 100 sq. km is 4.11. To check the homogeneity or inhomogeneity we have conducted Quadret Count Test and we have found that the intensity is not homogeneous (see Figure 3-1 and 3-2**).**

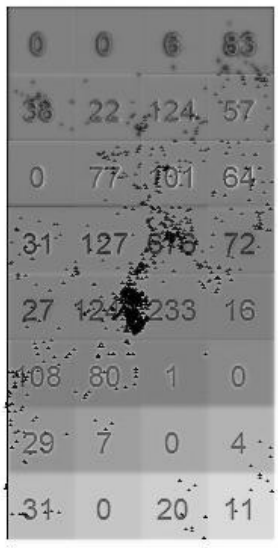


Figure 3-1 Quadrat count of Event data

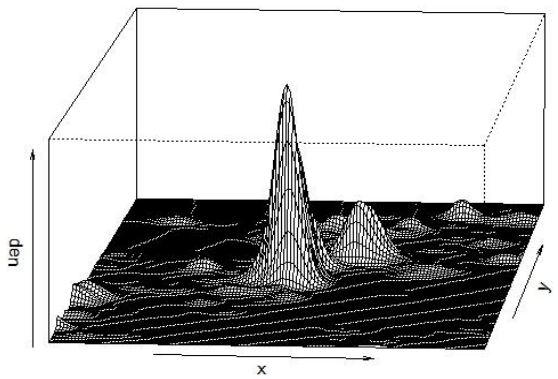


Figure 3-2 Perspective plot of event density

3.1.1 Dependence of intensity on a covariate

To explore the dependence of event intensity on covariates we have estimated the relative distribution of events as a function of different covariates. Let us assume that the intensity of the conflcit point process is a function of the covariate *Z* (Population, WASWI and SPI). Let, *Z(u)* be the value of the covariate then at any spatial location *u*, the intensity of the point process will be

$$\lambda(u) = \rho Z\big((u)\big) \dots \dots (2)$$

Where ρ is a function which explains how the intensity of AC depends on the value of the covariates. In our case, we have used Kernel smoothing to estimate the function ρ, using methods of relative distribution or relative risk, as explained in Baddeley and Turner (2006)

In the figures 3-3 and 3-4 the plots are some estimate of the intensity *ρ(z)* as a function of different covariates. It indicates that the events are relatively unlikely to be found where the number of the population is low or less than 1000 (see figure 3-3) on the other hand events are likely to be found where the WASWI value is low or dry region (maximum number of events are likely to be found in the range of -1 to -2). This Relative distribution estimate gives a signal that the intensity of the events depends on the values of a covariate. In both cases the pper and lower limit were of pointwise 95% confidence interval.

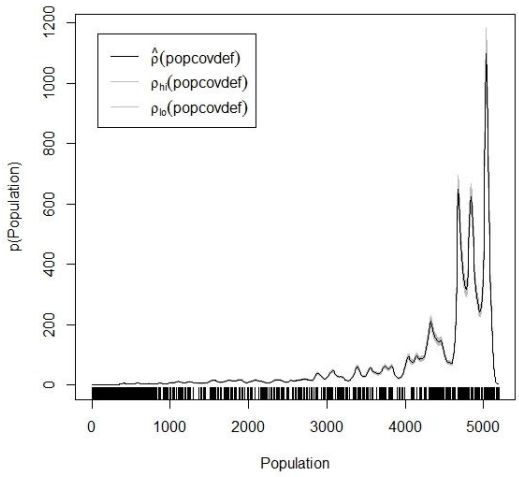


Figure 3-3 Events Intensity as a functions of Spatial Covariate population

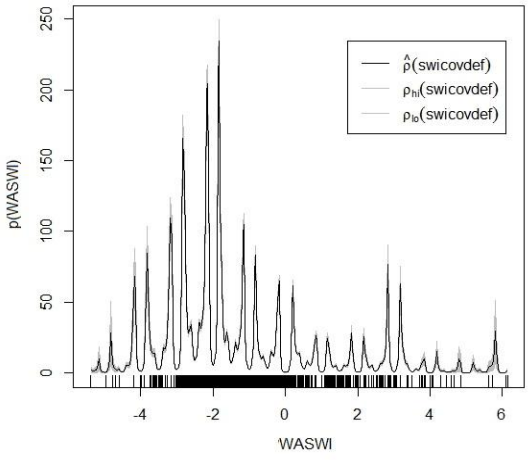


Figure 3-4 Events Intensity as a functions of Spatial Covariate WASWI

Following the event intensity dependence on the value of particular covariates, we have estimated the relative distribution of event based on two covariates . The result was an estimate of the intensity of event point process, as a function of two given spatial covariates WASWI and SPI (see Figure no 3-5) and we have found that in

both cases for SPI and WASWI, it is most likely that the events are most likely to be found where the values of covariates are relatively low on the other hand the events are most likely to be found where the number of population is high and values of WASWI is low (see figure 3-6).

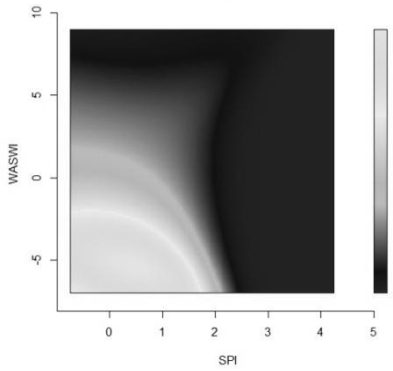


Figure 3-5 Events Intensity as a functions of two Spatial Covariate SPI and WASWI

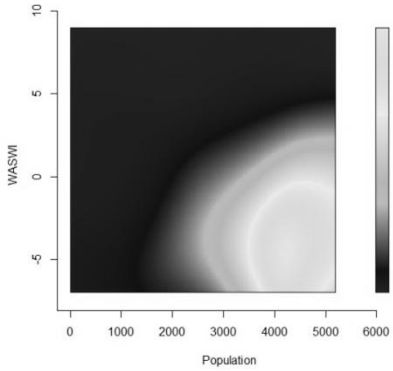


Figure 3-6 Events Intensity as a functions of two Spatial Covariate population and WASWI

3.2 Test for Complete Spatial Randomness

The basic benchmark model of a random point pattern is the uniform Poisson point process with homogeneous intensity λ, can be termed as Complete Spatial Randomness (CSR). If the point pattern is completely random then the points are completely unpredictable and have no trend. In our study our null model was a homogeneous poison process. If our null model is true then our points are independent of each other and have the same propensity to be found at any location. To find the evidence against CSR one of the classical tests of the null hypothesis of CSR is Chi-squared test of CSR using quadrat counts. So, we have estimated the Chi-squared test and the p-value was less than 0.001. Inspecting the p-value, we see that the test rejects the null hypothesis of CSR for the event data. As there are so many criticisms of chi-squared test in classical literature (see Baddeley and Turner 2005) in the next step we have conducted Kolmogorov-Smirnov test of CSR, where we have used population, WASWI and SPI as spatial covariates. The test output has been plotted in figure 3-7,3-8 and 3-9. In the plot we can see that the test reject our null model of CSR for the event data. So we continued our study of the Inhomogeneous point process.

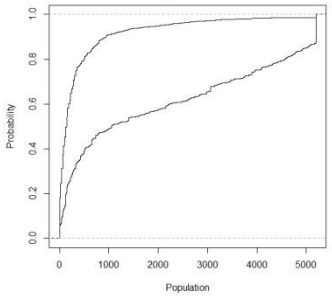


Figure 3-7 Spatial Kolmogorov-Smirnov test of CSR with population

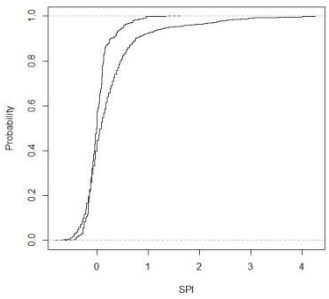


Figure 3-8 Spatial Kolmogorov-Smirnov test of CSR with SPI

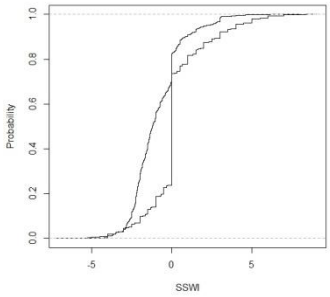


Figure 3-9 Spatial Kolmogorov-Smirnov test of CSR with WASWI

3.3 Inhomogeneous poisson process

The rejection of the null hypothesis for Complete Spatial Randomness (CSR) leads to have a closer look with more fine-grained analysis of point processes. So as suggested in Schabenberger and Gotway (2005) we have constructed our second hypothesis that maybe Inhomogeneous Poisson processes lead to clustering of events because we have observed that the intensity varies spatially (Figure 3-2).

More fine-grained analysis was led by four models. The first model has been chosen for the events data set is an Inhomogeneous Poisson process, the second is an inhomogeneous Thomas process, third one is an inhomogeneous Matern Cluster process and another one model which is considered in this study for our data sets is inhomogeneous (non-stationary) Area Interaction Point Process model.

3.3.1 Model I: Inhomogeneous poisson process

The ihomogeneous Poisson process of intensity $\lambda > 0$ has some particular properties. Such as under CSR the expected number of points falling in any region A is $E[N(A)] = \lambda.area(A)$ and n points which can be represented as $N(A)$ are independent and uniformly distributed in window A but in inhomogeneous Poisson process the intensity function $\lambda(u)$ are replaced by inhomogeneous intensity function. So now the number of point n falling in a region A has expectation

$$\mathbb{E}[N(A)] = \int_A \lambda(u)du \dots\dots (3)$$

where u is a particular location in region A. Then again, n points are independent and have unequal success probability density

$$p(u) = \frac{\lambda(u)}{I} \; where \; I = \int_A \lambda(u)dA$$

The inhomogeneous Poisson process is a credible model for point patterns under several scenarios. One is random thinning. Under such scenario the probability of expecting a point at the location u is $p(u)$. Then the resulting process of expecting points is inhomogeneous Poisson, with intensity $\lambda(u) = \beta p(u)$.

Inhomogeneous Poisson processes generate often clustered patterns. An area in the observed window, where intensity $\lambda(u)$ is high, obtains a greater density of events than the area where intensity $\lambda(u)$ is low. So an inhomogeneous Poisson process is sensitive to be selected as a model for our study where event intensity varies spatially (Diggle 2003).In our study a model of events expectation in a particular location, which assumes that all events are independent of each other, with an outbreak probability that depends on the local climate conditions like WASWI or number of Population. The resulting pattern of events is an inhomogeneous Poisson process.

3.3.2 Intensity estimation for inhomogeneous poisson process.

The estimate of intensity *λ(u)* at the location u is denoted by $\hat{\lambda}(u)$ and it is calculated by estimation of density at the location *u*. Suppose *x* is a set of a point pattern where $x = \{x_1 \dots \dots x_n\}$ in a compact window $A \subset \mathbb{R}^2$ then the density estimator function *f(x)* at x_0 (which is defined as the number of samples within distance *d* from x_0) is defined to be

$$\hat{f}(x_0) = \frac{1}{nd}\sum_{i=1}^{n} k\left(\frac{x_i - x_0}{d}\right) \dots \dots (4)$$

where *k(s)* is the uniform density on-1 ≤ s ≤1

$$k(s) = \begin{cases} 1 & \text{if } |x_i = x_0| \le d \\ 0 & \text{if otherwise} \end{cases}$$

Estimation of $f(x_0)$ is unbiased for a small neighborhood *d* but, it suffers from large variability. To minimize the large variability we have used Gaussian kernel instead of uniform kernel function, which does not refer to equal weight for all points inside the region $x_0 \pm d$, and that is defined to be

$$k(s) = \frac{1}{\sqrt{\pi}} exp\left(-\frac{s^2}{2}\right) \dots \dots \dots (5)$$

The probability estimate of an event at *u* location can be estimated from density estimation and the density estimation can integrate to one over the region *A*. The intensity $\lambda(u)$ and the density $f_A(u)$ in region *A* are related as

$$\lambda(u) = f_A(u) \int_A \lambda(v) dv$$

The product of two univariate kernel functions finds a kernel function product for a process in $\mathbb{R}^2$. Suppose the co-ordinates of x are y_i and z_i, the intensity estimator is defined by the product-kernel functions (Schabenberger and Gotway 2005) as follows:

$$\hat{\lambda}(u_0) = \frac{1}{v(A) d_y d_z} \sum_{i=1}^{n} k\left(\frac{y_i - y_0}{d_y}\right)\left(\frac{z_i - z_0}{d_z}\right) \ldots\ldots (6)$$

where d_y and d_z are the bandwidths in the respective directions of the co-ordinate system. The edge corrected kernel intensity estimator with a single bandwidth is given by

$$\hat{\lambda}(u) = \frac{1}{P_d(u)} \sum_{i=1}^{n} \frac{1}{d^2} k\left(\frac{u - u_i}{d}\right)$$

Where $P_d(u) = \int_A \frac{1}{d^2} k\left(\frac{u-v}{d}\right) dv$

$P_d(u)$ is played the role as the edge correction.

3.3.3 Inhomogeneous K-function

The k-function known as Reply's K-function or reduced second moment function of a stationary point process *x* is defined to estimate the expected number of additional random points within a distance *r* of a random point x_1, which was first introduced by Ripley (1977). $\lambda K(r)$ equals to the expected number of random points within a distance *r* of x_1 where lambda is the intensity of the process and $K(r) = \pi r^2$ Deviations between the theoretical and empirical K curves may suggest spatial clustering or spatial regularity(Diggle 1983).

Ripley's K function is defined only for stationary point processes. A modification of the K-function can be used in inhomogeneous processes to the aggregation in events which was proposed by Baddeley et al. (2000). In inhomogeneous Poisson processes, events are independent in the subregion, but the intensity $\lambda(x)$ varies spatially throughout the region

The inhomogeneous K function $K_{inhom}(r)$ is a direct generalization to non-stationary point processes. Following Baddeley et al. (2000) an inhomogeneous Poisson process is also second-order intensity-reweighted stationary if

$$\lambda_2(x,y) = \lambda(x)\lambda(y)g(x-y)$$

where $g(x-y)$ is a function and which is defined by the spatial lag and the interaction between the arbitrary events x and y. So, the product of the first-order intensities at x and y multiplied by a spatial correlation factor refers to the second-order intensity of the inhomogeneous Poisson processes. If the spatial interaction between the points of the process at the location x and y is 0 then $\lambda_2(x,y) = \lambda(x)\lambda(y)g(x-y)$ as $g(x-y) = 1$. And if ||.|| is the Euclidean norm then the pair correlation function is defined as

$$g(||x-y||) = \frac{\lambda_2(x,y)}{\lambda(x)\lambda(y)}$$

The corresponding intensity reweighted K-function is

$$K_{inhom}(d) = 2\pi \int_0^d ug(u), d >)0$$

For an inhomogeneous Poisson process where there is no spatial interaction between events then the inhomogeneous K-function is $K_{inhom}(d) = \pi d^2$ as for the homogeneous case. In the spatstat package the standard estimators of K-function can be extended to the inhomogeneous K-function (Baddeley et al., 2000) as below

$$\widehat{K_I}(d,\lambda) = \frac{1}{v(A)} \sum_{i=1}^{n} \sum_{j \neq i} \frac{w_{ij} I(d_{ij} \leq d)}{\hat{\lambda}(x_i)\hat{\lambda}(x_j)} \dots \dots (7)$$

where $v(A)$ is an area of the study region, d_{ij} denote distance between the i^{th} and j^{th} observed points and

$$I(d_{ij}) \begin{cases} 1 & \text{if } d_{ij} \leq \text{d} \\ 0 & \text{if otherwise} \end{cases}$$

and w_{ij} is an edge-correction weight and $\hat{\lambda}(x)$ is an estimate of the intensity function $\lambda(x)$.

3.3.4 Interpretation of K_{inhom} with L_{inhom}-function

Analogously to the case of homogeneous K-function, we can set

$$\hat{L}_{inhom} = \sqrt{\frac{\widehat{K}_{inhom}(d)}{\pi}} \dots \dots (8)$$

where the function is linearized dividing by π and the square root transformation also approximately stabilizes the variance of the estimator. In the inhomogeneous passion process, under the null hypothesis of no spatial dependence, we have

$$\hat{L}(d) = d$$

3.3.5 Model fitting and simulation

The inhomogeneous version of L - function for the event dataset with covariates has been estimated under inhomogeneous Poisson model and plotted in figure 3-11. The plot of L-function shows that the events are clustered up to 300 km. This result also complies with the inhomogeneous intensity, can be observed in figure 3-10. The intensity surface plot indicates that 20-25 distinct point are of high density in the study window.

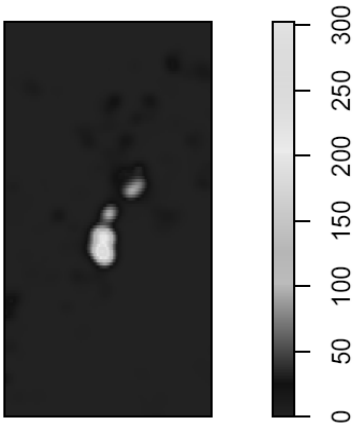


Figure 3-10 Estimation of the inhomogeneous intensity surface

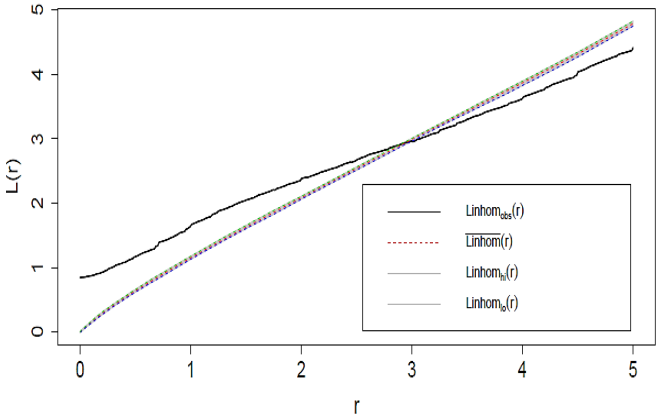


Figure 3-11 estimate of inhomogeneous K-function (right and black solid line) for the crime data sets.

The estimate of the inhomogeneous L - function exhibits that our empirical data do not fit well with the IPP model. The fit seems not to be satisfactory as the estimator $L_{inhom}(d)$ do not lie between it's 99 simulations envelope. This result implies that the data may not come from the inhomogeneous Poisson processes but the event distribution is affected by the independent variable's distribution and so clustered up to 300 km. As we know that the inhomogeneous Poisson process does not accommodate interaction between the points so this model only explains that the event distribution are inhomogeneous depending on the inhomogeneous distribution of covariates but the empirical line lie outside of the simulated envelope. So, this model is not sufficient enough to describe the event occurrence process or event distribution in the study area. So, we can suspect there might be some sort of interaction between points. So, we needed to check other models which accommodated the interaction in the point process modelling besides covariates. Such as Area-interaction model. But before we go for the area interaction model, in

the next stage we have estimated L-function for the Inhomogeneous Poisson Cluster models to see if there is any cluster to cluster interactions.

3.4 Model II: Inhomogeneous poisson cluster process

Assessing the result of the Inhomogeneous Poisson Process model, we have concluded that our event data set might not only depend on the trend. So there might be some sort of interaction or dependence between the points. Non Poisson process models are formulated to exhibits interaction or dependence between the points. Some simple Poisson process cluster models are derived from Poisson process models such as the Marten Cluster Model and Thomas Cluster Models, what we are going to implement to extract armed conflict and climate change related information; Inhomogeneous Thomas Cluster Process Model and Matern Cluster Process Model.

Neyman-Scott process consists of clusters of offspring points which are centered on a set of parent points which was first introduced in Neyman and Scott (1958). Each parent points produces an expected number (μ_0) of offspring points. According to a dispersion density function $\kappa(.)$ those points are dispersed within a radius of r of the parent points where perimeter σ controls the dispersion.

Poisson point process Matern cluster process and Thomas Cluster process are a special case of a Neyman-Scott process where the parent process comes from a Poisson process.

In the Poisson cluster process, the parent point set is y where each parent point $y_i \in y$. According to some stochastic mechanism the parent point set gives birth to a finite set of offspring points Z_i and the point process set x are then constructed comprising all of these offspring. In both homogeneous Thomas Process and Matern process the parent points come from a Poisson Process with intensity κ and each cluster consists of a Poison number (μ) of offspring points. The dispersion of the offspring depends on the dispersion density function. In Thomas process the offspring has an isotropic Gaussian $N(0,\sigma^2 I)$ distribution of its parent and in Matern process the offspring being placed independently and uniformly inside a disc of radius r centered on the parent point..

In this study the cluster processes with homogeneous parent intensity κ and inhomogeneous cluster reference density $\mu(x, y)$, which has the overall intensity $\lambda(x, y) = \kappa\mu(x, y)$ is referenced as an inhomogeneous Poisson cluster process

At this stage of our study we assume that there are interactions between points so there are clustering and each cluster are inhomogeneous. So we have introduced

cluster inhomogeneity in the cluster process model suggested by Waagepetersen (2008).

3.4.1 Modelling procedure:

Let $A \subset \mathbb{R}^2$ be the study window and a homogeneous Poisson process. The first order intensity of that process is κ where $\kappa \geq 0$. At the location u of the study window, $S_{1:p}(u)$ is the $1 \times p$ and $p \geq 1$ are vector of covariates. An offspring event at the location u is obtained with a probability $\frac{exp(s_{1:p}(u)\beta_{1:p}^T)}{M}$, where $M = \max\{\exp[s_{1:p}(u)\beta_{1:p}^T]\}$ and $\beta_{1:p}^T$ is the $1 \times p$ vector of unknown parameters (Møller, 1999).

x is the set of point process consists of offspring event only with parents' points in a stationary Poisson point process of intensity κ. According to the definition of the inhomogeneous point process the intensity function of x is given by

$$\lambda(u) = \kappa\mu \exp\big(s_{1:p}(u)\beta_{1:p}^T\big) \dots \dots (9)$$

Where, in Matern process the offsprings are independent and uniformly distributed and in Thomas process the offspring has isotropic Gaussian $N(0,\sigma^2 I)$ distribution. The equation (2.1) can be rewritten as

$$\lambda(u) = \exp(s(u)\beta^T) \dots \dots (10)$$

where

$$s(u) = \big(1,\ s_{1:p}(u)\big)$$

and

$$\beta = \big(\beta_0, \beta_{1:p}\big) = (\log(\kappa\mu)\,, \beta_{1:p}\,)$$

For parameter estimation, we may obtain an estimate of the parameter using an estimating function. We have

$$l(\beta) = \sum_{u \in X \cap A} s(u)\beta^T - \int_A \exp(s(u)\beta^T) du \dots \dots (11)$$

It corresponds to the log-likelihood. The estimation is given by

$$f(\beta) = \frac{d}{d\beta} l(\beta) = \sum_{u \in X \cap A} s(u) - \int_A \exp(s(u)\beta^T)\, du \dots \dots (12)$$

With sensitivity $t(\beta) = \int_A s(u)^T s(u) \exp((s(u))\beta^T) du$

An estimate of the K-function for x can be acquired using inhomogeneous K-function (discussed in section 3.3.3) by substituting the intensity in 12. So

$$\widehat{K}_{inhom}(d) = \sum_{u_1,u_2 \in X \cap A} e(u_1,u_2) \frac{1[0 < ||u_1 - u_2|| < d}{exp((s(u_1) + s(u_2))\hat{\beta}^T} \beta^T \dots\dots (13)$$

Where $e(u_1,u_2)$ is an edge correction. In our study we have used and inhomogeneous L-function is like following (see section 3.3.4)

$$\hat{L}_{inhom} = \sqrt{\frac{\widehat{K}_{inhom}(d)}{\pi}}$$

3.4.2 Modelling fitting and simulation:

In our study of fitting an Inhomogeneous Thomas model to the final dataset plotted in figure 01, the summary statistic inhomogeneous L-function has been estimated. A general algorithm for fitting theoretical point process models to point process data by the Method of Minimum Contrast described in Diggle and Gratton (1984)**.** In this method, estimates of parameters $\theta = (\kappa, \mu, \sigma)$ can be obtained by the best matching between L-function, $L_\theta(d)$ and the estimated L function of the data, $\hat{L}(d)$ (Baddeley 2010). The following equation can give the best match by minimizing the divergence between the two functions over the interval [a, b]:

$$D(\theta) = \int_a^b \left|\hat{L}^q(d) - L^p(d,\theta)\right|^p d(d);$$

For $0 \leq a \leq b \; and \; p,q > 0, where \; d(d)$ denote integration with respect to distance d.

To fit inhomogeneous Cluster Process (both Matern and Thomas process) we have estimated the inhomogeneous intensity of the process and then obtained an estimate of the inhomogeneous L-function. Minimum contrast method was used to estimate the parent intensity κ and the Gaussian standard deviation σ (for Thomas process) and uniform distribution for Matern process. A long term trend is included in the model as a trend formula

$$log\lambda_\theta(x,y,z) = \theta_0 + \theta_1 x + \theta_2 y + \theta_3 z \dots\dots (14)$$

Where *x,y* and *z* are the covariates which represent population, WASWI and SPI subsequently. The density was predicted and passed to inhomogeneous L-function for the inhomogeneous Thomas process. We can see the inhomogeneous L-function of the model in Figure 3-13.

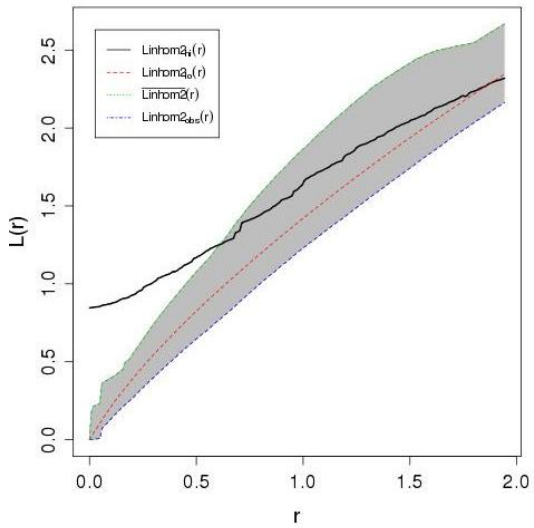


Figure 3-12 Point-wise critical envelopes for inhomogeneous version of the L-function in Inhomogeneous Matern Process; ; the data line (black), theoretical line (red)

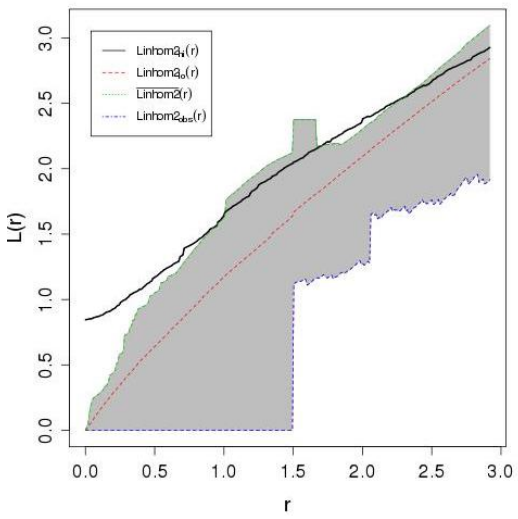


Figure 3-13 Thomas Process, obtained from 99 simulations where the Upper envelope is point-wise maximum of simulated curves and Lower envelope is point-wise minimum of simulated curves; the data line (black), theoretical line (red)

The parameters of the inhomogeneous cluster processes are estimated from the inhomogeneous L-function using minimum contrast method. The estimated parameters for Thomas process are $\kappa = 0.00288579\ and\ \sigma = 5.13038348$ and $\kappa = 0.003042635$ and $\mathrm{R} = 4.112731596$ for the Matern Cluster process. The plot of the inhomogeneous L - function in Figure 3-12 shows a little bit better fit than the Inhomogeneous Poisson process. To explain the adequacy of the model in describing spatial trends of the data based on the convariates and parameters, the model test is carried out by summary functions of the event data comparing with those of 99 simulations from the model. For the Thomas process, the L-function curve does not completely lies within its envelope but after 120 km the empirical line fall inside the envelope but again it lies outside the envelop from 200 to 230 km. This validity test suggests that the model is not good enough as we expected. On the other hand for the Matern Cluster the empirical line fall outside the envelop 0- 60 km. The fitted coefficients of the trend formulas are shown in table 3-1

Table 3-1 Fitted coefficients for trend formula: ~population + WASWI + SPI

	(Intercept)	**Population**	**WASWI**	**SPI**
Thomas Cluster	0.7757	0.0006	-0.2668	-0.4187
Matern Cluster	0.7757	0.0006	-0.26683	-0.41879

From the coefficient it can be interpreted that, the events are positively related with the population and negatively with the WASWI and SPI.

3.5 Gibbs mode model IV: area-interaction process

In this study, Inhomogeneous Area-Interaction model will be specified in terms of its conditional intensity (means the local intensity will depend on the spatial location so it will represent a spatial trend or spatial covariate effects) rather than its likelihood and stochastic interactions or dependence between the points of the random point process. So it's more appealing way to formulate point process models for such study.

In Standard case, in the area-interaction process or Widom-Rowlinson penetrable spheres model model (Widom and Rowlinson 1970; Baddeley and Lieshout 1995), the probability density of a point pattern $x = \{x_1 \dots \dots x_n\} (n \geq 0)$ in a compact window $A \subset \mathbb{R}^d$ is defined to be

$$p(x) = \alpha \beta^{n(x)} \gamma^{-m(U_r(x))} \dots \dots (15)$$

where the disc radius is *r*, intensity parameter β and interaction parameter γ. Here $\beta, \gamma, r > 0$ are parameters and α is the normalizing constant, *m* is Lebesgue measure, and $U_r(x)$ is the discs of radius r centered at the points of the realization $B(x_i, r) = \{\alpha \in \mathbb{R}^d : ||a - x_i|| \leq r\}$. $U_r(x)$ is defined to be

$$U_r(x) = \bigcup_{i-1}^{n} B(x_i, r)$$

Densities reduce to a Poisson process with intensity β_μ when $\gamma = 1$, exhibit ordered patterns for $0 > \gamma > 1$ and for $\gamma > 1$ 'clustered'. The clustered case $\gamma > 1$ of (15) is identical to tile 'penetrable sphere model' of liquid-vapor equilibrium introduced by Widom and Rowlinson (1970). Therefore, both clustered and regular point patterns can be modeled using an area interaction process.

The standard form of the model (15) is a little complicated to interpret in practical applications. For example, each isolated point of the pattern *x* contributes a $\beta\gamma^{-\pi r^2}$ to the probability density. (Baddeley and Turner 2005). In spatstat package, the model (15) is parameterized in a different form, which is easier to interpret. The probability density in canonical scale-free form is rewritten as

$$p(x) = \alpha \beta^{n(x)} \eta^{-C(x)} \dots \dots (16)$$

Where β is the new intensity parameter, η is the new interaction parameter, and $C(x)$ is the interaction potential can be referred to $B(x) - n(x)$. If C(x) = 0 that means the point pattern x does not contain any points that lie close together. $B(x)$ is the normalized area and which devise the discs to have unit area, is defined to be

$$B(x) = \frac{A(x)}{\pi r^2} \dots\dots (17)$$

In canonical scale-free form parameters β and η are defined to be

$$\beta = \kappa\gamma^{-\pi r^2} = \frac{\kappa}{\eta} \quad and\ \eta = \gamma^{\pi r^2}$$

The parameter η can take any nonnegative value. When η = 1, it corresponds to a Poisson process, with intensity β and for $\eta < 1$ the process is regular and in other cases the process is clustered.

The non-stationary area interaction process is similar except that the contribution of each individual point $x_{(i)}$ is a function $\beta\, x_{(i)}$ of location, rather than a constant β.

3.5.1 Model fitting and simulation:

The interaction in the point pattern in terms, of the area of the union covered by the area of influence, of the points with radius r can be profiled by the area interaction process. In this study, the points represent the location of the conflicts and the 'area of influence' is the area in which the conflicts get inspired by the spatial change in independent variables/covariates.

To fit the Area Interaction Point process model we have used the profilepl functionality of spatstat package. This model fitting function fits point process models to point pattern data by using profile maximum Pseudolikelihood method. This function is a binder which finds the values of the irregular parameters (in our cases we have used the sequence of 100 to 200 Km, by 10 Km) that give the best fit. We have found the best fit for 100 km radius. See Figure 3-15

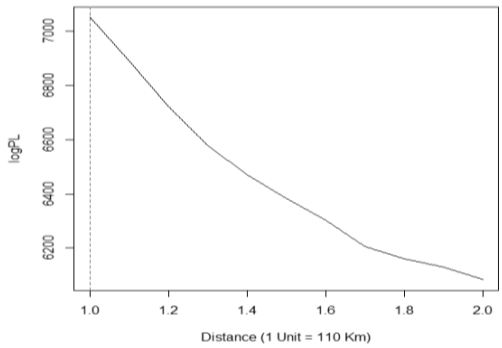


Figure 3-14 Profile log pseudolikelihood values for the Trend formula: Armed Conflict ~population+ WASWI + SPI; fitted with rbord= 4; Interaction: Area Interaction with irregular parameter 'r' in [100, 200 km]. Optimum value of irregular parameter: r = 100 km

We have simulated Inhomogeneous Area Interaction point process model with trend and interaction, where the interaction radius was set as 110 km. The specific trend function and parameter specification are taken from Baddeley and Turner (2005). A realization from this point process was generated on a unit square. In our case, the inhomogeneous L - function was computed to test for departure from Complete Spatial Randomness (CSR). The intensity was computed at each location using trend function and with the point process was simulated. The inhomogeneous L - function of Area Interaction Process is plotted in Figure 3-15.

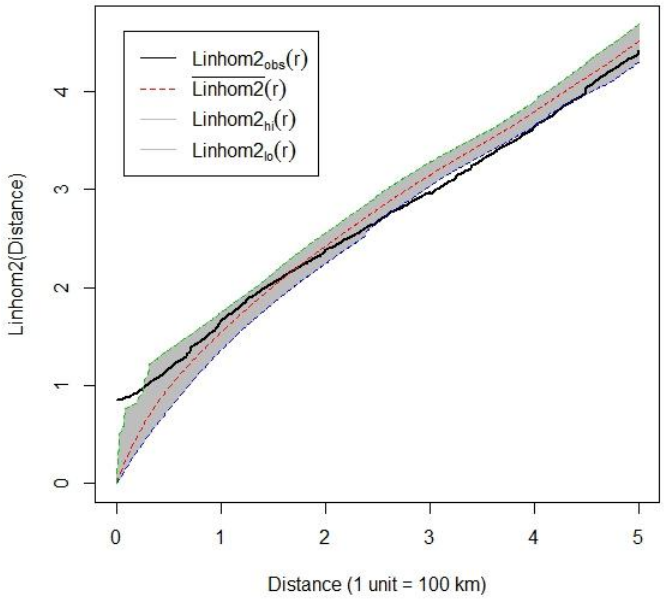


Figure 3-15 : Pointwise critical envelopes for inhomogeneous version of the L-function in Inhomogeneous Area Interaction Process, obtained from 99 simulations of fitted Gibbs model where the Upper envelope is point-wise maximum of simulated curves and Lower envelope is point-wise minimum of simulated curves; Significance level of Monte Carlo test: 2/100 = 0.02' Data: Inhomogeneous Area Interaction with fitted parameter 'r' in [100 km]; Trend formula: Armed Conflict ~population+ WASWI+SPI

The estimate of inhomogeneous L-function function suggests better agreement between the model and the data. The fit seems to be satisfactory as the estimator lies

between its 99 simulations envelope. The values of the coefficients of intensity have evidenced by good matching between the model and the data too. The estimator detects some inhibition at small distances and interprets the trend as clustering up to 200 km. It is the model which represents an association between the points, AC data exhibits clustering, so here is a clear signal that attracts between points and the covariates in the pointer association. The Inhomogeneous cluster point process model Fitted to point pattern dataset 'Armed Conflict' Fitted using Area-interaction process and the trend formula was Armed Conflict ~population+ Weighted Anomaly Standardized soil water index + Standardized Precipitation Index. The Fitted coefficients for trend formula are shown in the able 3-2:

Table 3-2 Fitted coefficients for trend formula Armed Conflict ~population+ Weighted Anomaly Standardized soil water index + Standardized Precipitation Index using Nonstationary Area-interaction process

(Intercept)	Population	WASWI	SPI
-16.30236	0.00044	-0.16475	-0.479

The coefficient exhibits that change of one unit in Weighted Anomaly Standardized soil water index (WASWI), the average change in the mean of Armed Conflict is about -0.1647 units. Thus the WASWI is negatively related to the Armed Conflict incidents. If the value of WASWI decreases the Armed Conflict increases meaning the study region experienced armed conflict due to reduction of water in Soil. On the other hand the SPI is positively related with the AC and population as well.

3.6 Yearly plot of K function

In the next stage of point pattern analysis, we have fitted to characterize yearly conflict data based on the yearly interaction and covariate effect and the plots have been shown in anex . In most of the years the Cluster Process Model gives best results. The empirical lines are fitted well with the theoretical lines and lie inside the 99 simulation envelop. For some year, the Inhomogeneous Matern Cluster and Inhomogeneous Thomas Cluster process are plotted in figure 3-16 to 3-21. According to previous studies, we can assume that the extreme climate events of a particular year can trigger conflict in the next following years. So, we tried to check the fact that the probability to find an Armed Conflict in a particular location might depends on the value of WASWI (negatively related), SPI (negatively related) and Population (positively related) of the previous years. So besides plotting the summary function with the covariates of the same year we have plotted the summary function combining the covariates of previous one and two rears.

In figure 3-16 to 3-24 we have shown the plot of estimation of L-Function for Inhomogeneous Cluster process for the year 1997 and 1999 with the temporal lag of "0" year "one" year and "two" year respectively. From the plot we can observe that

the cluster process of 1997 with a time "0" lag year fitted well and lie inside the envelope on the other hand the data of 1997 with a one year lag and a two year lag are not fitted well which means there are less significant relation between the conflict events and covariates of the previous year but still it exhibits clustating.

Then again, based on the plot for year 1999 we can see that the event location are significantly related with "0", "one" and "two" lag year covariates. So the data maintain a nonlinear relationship.

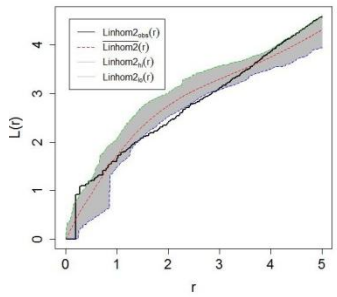


Figure 3-16 Inhomogeneous Thomas cluster process for the year 1999 (with covariates of the year 1999)

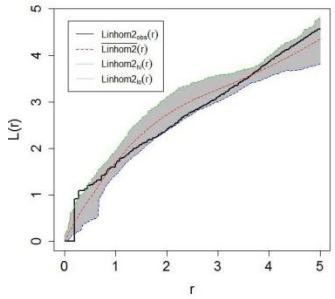


Figure 3-17 Inhomogeneous Thomas cluster process for the year 1999 (with covariates of the year 1998)

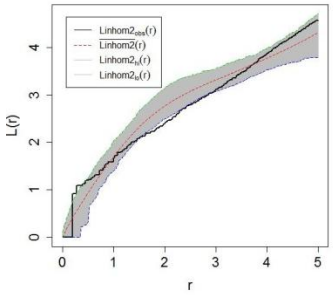


Figure 3-18 Inhomogeneous Thomas cluster process for the year 1999 (with covariates of the year 1997)

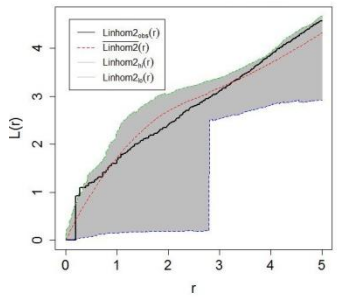


Figure 3-19 Inhomogeneous Matern cluster process for the year 1999 (with covariates of the year 1999)

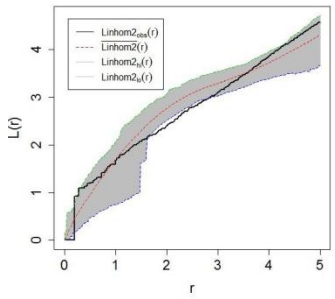


Figure 3-20 Inhomogeneous Matern cluster process for the year 1999 (with covariates of the year 1999)

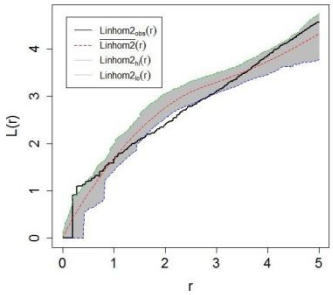


Figure 3-21 Inhomogeneous Matern cluster process for the year 1999 (with covariates of the year 1999)

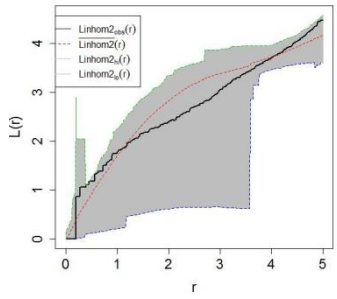


Figure 3-22 Inhomogeneous Matern cluster process for the year 1997 (with covariates of the year 1997)

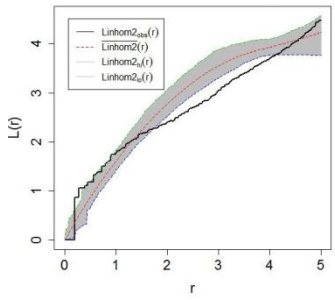


Figure 3-23 Inhomogeneous Matern cluster process for the year 1997 (with covariates of the year 1996)

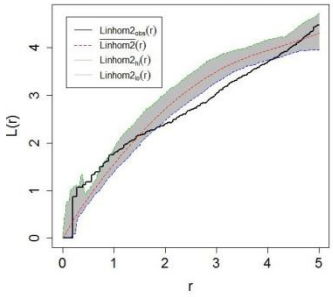


Figure 3-24 Inhomogeneous Matern cluster process for the year 1997 (with covariates of the year 1995)

For different temporal lag from the fitted model we got coefficient for different covariates which explain the relationship between armed conflict and different covariates. To understand the relationship trend we have plotted the coefficient against time in figure 3-25, 3-26 and 3-27. These figures show the coefficient of WASWI for different temporal lag. From the plot its evidenced that most of the time the conflict and WASWI are negatively related but some time it has modeled the positive relationship such as in the year 2000 with temporal lag "0". In 1999 with temporal lag t-1 and t-2 and so. But the positive relationships are less significant. See the L-Function plots in annex.

Then again it's the same for SPI. Though with some specific temporal lag in different years the relationship is positive but most of the time the relationship is negative. Which explains that the negative change of one unit in SPI, the average change in the mean of Armed Conflict is about the value of fitted coefficient units in figure SPI. On the other hand in figure 3-27 we can see that the population is always positively related to the armed conflict events.

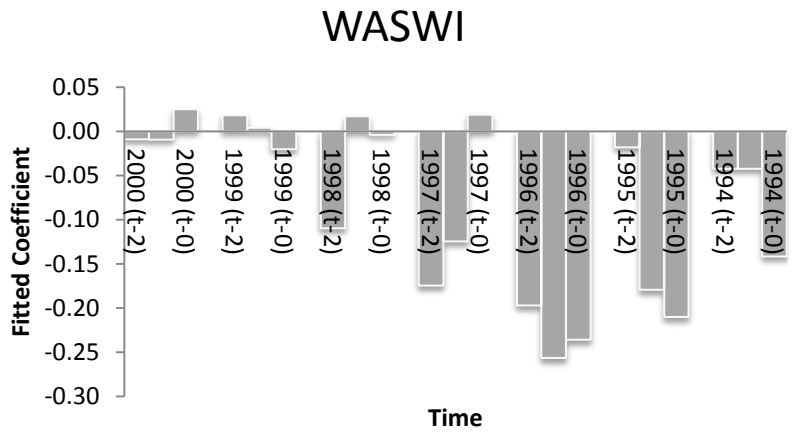


Figure 3-25 Year wise fitted coefficients of WASWI with different temporal lag

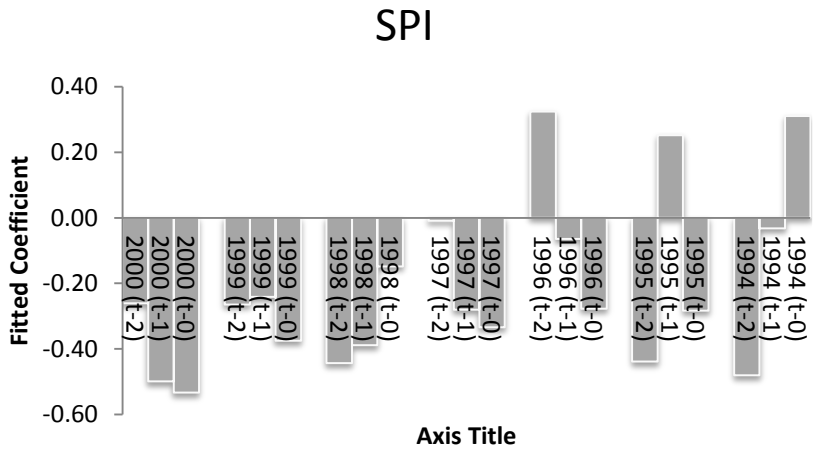


Figure 3-26 Year wise fitted coefficients of SPI with different temporal lag

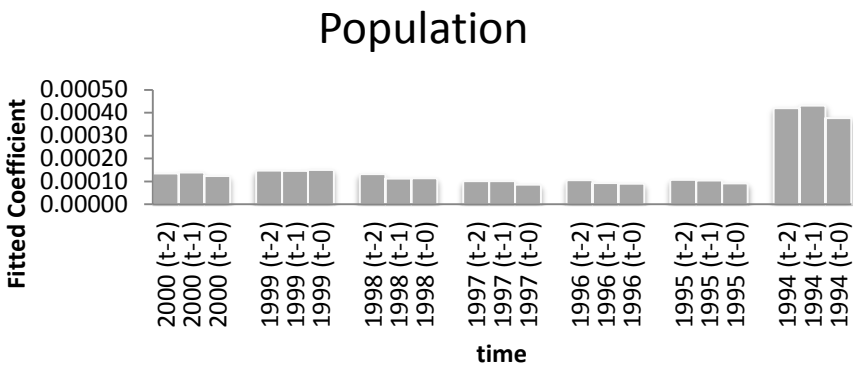


Figure 3-27 Year wise fitted coefficient of Population with different temporal lag

3.7 Space time point process modelling

In our study we have fitted spatial point process models to understand the spatial interaction of events and model the relationship between armed conflict and climate change through the sense of covariates effects in point process modeling. In the next step we want to go a step further by including the temporal dimension in the point process modeling dynamically.

A suitable Spatio-temporal point process models accommodate spatial and temporal inhomogeneity and identify clustering in time besides clustering in time. In other word it explains the probability of finding events in space and time. In our study the spatio-temporal conflict outbreak assumption was that the pattern of conflict incidence depends on both spatial distribution of covariates and temporal variations as well. Spatial distribution and temporal variation can predict the probability of finding events in space and time. In our study we have used Space-Time Inhomogeneous K- function (STIKhat) to model the spatio-temporal relationship between armed conflict and covariates.

3.7.1 Estimation of the space-time inhomogeneous K-function

Suppose x is a set of a point process where $x_i = \{s_i t_i \dots \dots s_n t_n\}$ in a compact window $A \subset \mathbb{R}^2$ of the form $A = S \times T$. Here S is the region, T time interval, s_i is the spatial location of the i^{th} event and t_i is the time of occurrence. If $\lambda\,(s,t)$ denotes the mean number of events per unit volume at location(s,t), according to Gabriel and Diggle (2009), the second order intensity (covariance of events per unit volume) can be defined as

$$\lambda_2\big((s,t),(s',t')\big) = \lim_{|ds\times dt|,|ds'\times dt'|\to 0} \frac{\mathbb{E}[Y(ds\times dt)Y(ds'\times dt')]}{|ds\times dt|,|ds'\times dt'|} \quad (18)$$

Where $(ds \times dt)$ is a cylindrical container with volume $|ds \times dt|$, contain the point (s,t) and $Y(ds \times dt)$ denotes the number of events in the container $(ds \times dt)$.then again, the standardized probability density of an event occurring in each of two volume $(|ds \times dt| \; and \; |ds' \times dt'|)$ can be identified by pair correlation function and according to (Cressie 1993; Diggle 2003) that can be defined as

$$g\big((s,t),(s',t')\big) = \frac{\lambda_2\big((s,t),(s',t')\big)}{\lambda(s,t)\lambda(s',t')} \dots\dots (19)$$

For a spatio-temporal Poisson process case, the covariance density is 0 and $g\big((s,t),(s',t')\big)$ = 1. If the value of $g\big((s,t),(s',t')\big)$ is grater than 1 then it is the indication of how much more likeliness of an event occurance at a specific location than in a poission process.

Suppose (u,v) is the spatio-temporal difference vector where $u = ||s - s'|| \; and \; v = |t - t'|$. Then according to Gabriel and Diggle (2009) for the inhomogeneous space-time point process, the space-time inhomogeneous K-function (STIK-function) can be defined as

$$K_{ST}(u,v) = 2\pi \int_0^v \int_0^u g(u',v')u'du'dv' \dots\dots (20)$$

Which can be used as a measure of spatiotemporal clustering (Gabriel et al. 2012). For an inhomogeneous spatio-temporal Poisson process $K_{ST}(u,v) = \pi u^2 v$. On the other hand, if $K_{ST}(u,v) > \pi u^2 v$ the STIK-function indicates aggregation and $K_{ST}(u,v) < \pi u^2 v$ indicates regularity. Besides, The STIK function also useful to test space-time clustering and space-time interaction (Møller and Ghorbani 2012). In our study we have used STIKhat function from stpp package in R where for the parameter infectious = TRUE, the STIK function can be defined as (21)

$$\widehat{K}_{ST}(u,v) = \frac{1}{|S \times T|} \sum_{i=1}^{n} \sum_{j \neq i} \frac{1}{w_{ij} v_{ij}} \frac{1}{\lambda(s_i,t_i)\lambda\big(s_j,t_j\big)} 1\{||s_i - s_j|| \leq u; t_j - t_i \leq v\}$$

Where, w_{ij} denotes the Ripley's spatial edge correction factor and v_{ij} denotes the temporal edge correction factor (the one-dimensional analogue of the Ripley's edge correction factor)) (see Gabriel 2012.)

And Space-time inhomogeneous pair correlation function can be defined as

$$\hat{g}(u,v) = \frac{1}{|S \times T|} \sum_{i=1}^{n} \sum_{j \neq i} \frac{1}{w_{ij} v_{ij}} \frac{k_s(u - ||s_i - s_j||)k_t(v - ||t_i - t_j||)}{\lambda(s_i,t_i)\lambda(s_j,t_j)} \; (22)$$

Where $k_s \; and \; k_t$ are the kernel functions with bandwidth $h_s \; and \; h_t$

3.7.1.1 Simulation study

We have estimated the spatial intensity by the command *msd2d* of the *splancs* package (Rowlingson and Diggle 2010) which is based on minimizing a mean square error (for detail see Diggle 1985). Then the spatial intensity was normalized based on the total area and number of events. We have calculated the spatial intensity for two cases (spatial intensity based on karnel and spatial intensity based on covariates) and those are shown in figure 3-28 and 3-29. The intensity image shows the inhomogeneous intensity over the region in both cases. We have used the Image of the spatial intensity based on kernel for the STIkhat without covariates and Image of the spatial intensity based on covariates was used in STIKhat with covariate case. We have also calculated the temporal intensity using Gaussian kernel.

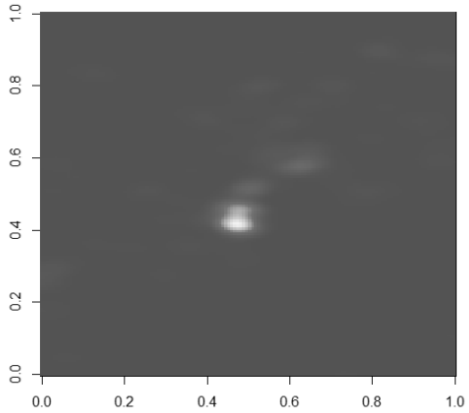


Figure 3-28 Image of the spatial intensity based on kernel

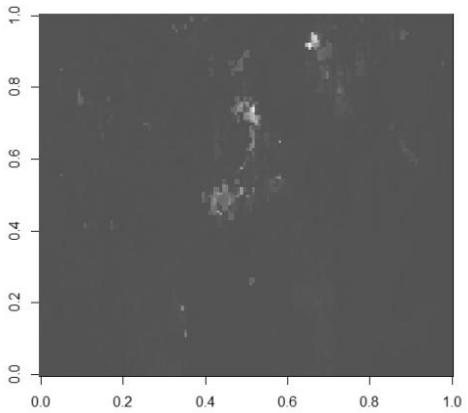


Figure 3-29 Image of the spatial intensity based on covariates

Based on the spatial and temporal intensity we have calculated the STIKhat for two cases. STIKhat without covariates and STIKhat with covariates. STIKhat was estimated using stikhat function of the stpp package (Gabriel E., Diggle P. (2009)). While estimating STIkhat we have used several compositions of sequences of time and distance. Some of them are plotted in figure 3-30 and figure 3-36 for both cases. By looking at the STIKhat functions it is clear that the event data do not behave as an inhomogeneous spatio-temporal Poisson process and on the contrary they show space-time interaction. When we control the variability of the covariate information, this interaction is in form of aggregation in the following sense that events tend to come in clusters for spatial distances up to 10 units and temporal distances up to 20 days. Note that the values of the STIKhat are positive and far from zero indicating that the part coming from the spatio-temporal K function is stronger than that coming from the theoretical value under the Poisson case.

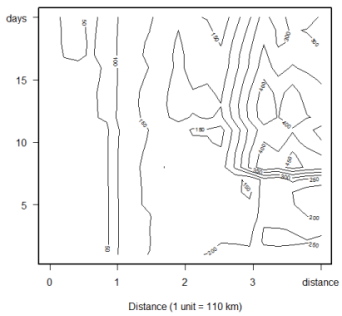


Figure 3-30 $K_{ST}(u,v) - \pi u^2 v$ with small u (up to 440 km) and v (up to 20 days) for the case events without covariates

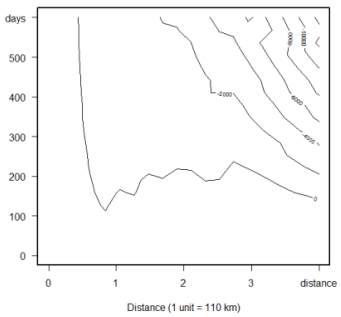


Figure 3-31 $K_{ST}(u,v) - \pi u^2 v$ with small u (up to 440 km) and larger v (up to 2 years app.) for the case events without covariates

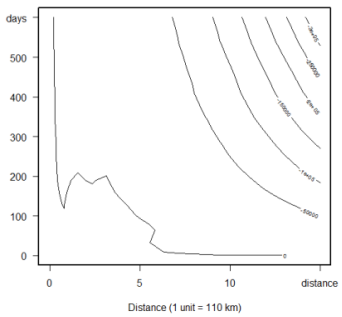


Figure 3-32 $K_{ST}(u,v) - \pi u^2 v$ with larger u (up to 1500 km) and v (up to 2 years app.) for the case events without covariates

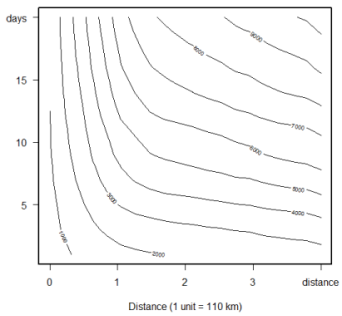


Figure 3-33 $K_{ST}(u,v) - \pi u^2 v$ with small u (up to 440 km) and v (up to 20 days) for the case events with covariates

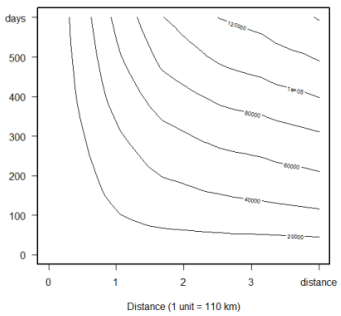


Figure 3-34 $K_{ST}(u,v) - \pi u^2 v$ with small u (up to 440 km) and larger v (up to 2 years app.) for the case events with covariates

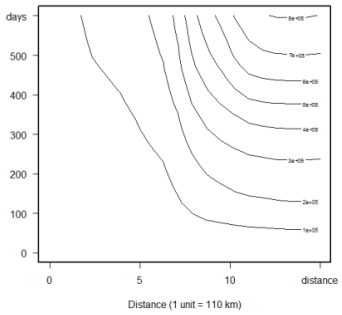


Figure 3-35 $K_{ST}(u,v) - \pi u^2 v$ with larger u (up to 1500 km) and v (up to 2 years app.) for the case events with covariates

We can also see that the simulated behavior superimposed with the events in figure 3-38, 3-39, 3-40 and 3-41. The simulated events behave as an inhomogeneous Poisson case. In the case of covariates, they are more scattered than when using kernel, a consequence of having covariate information.

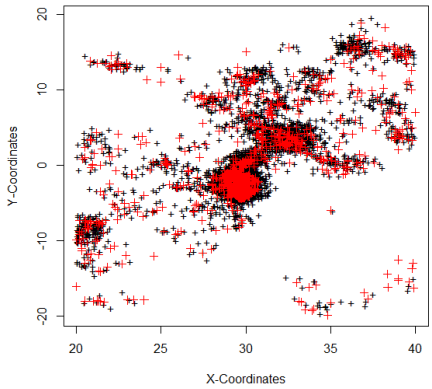


Figure 3-36 Superimposed events (red) with simulated (black) rpp using karnel (1st cases)

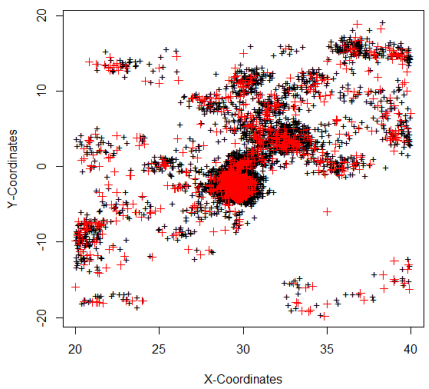


Figure 3-37 Superimposed events (red) with simulated (black) rpp using karnel (2nd cases)

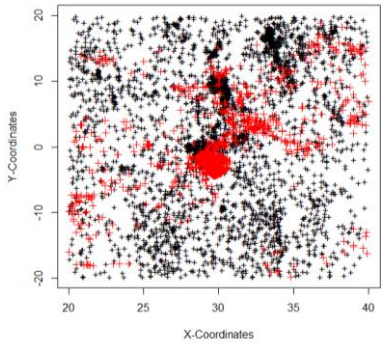


Figure 3-38 Superimposed events (red) with simulated (black) rpp using Covariates (1st cases)

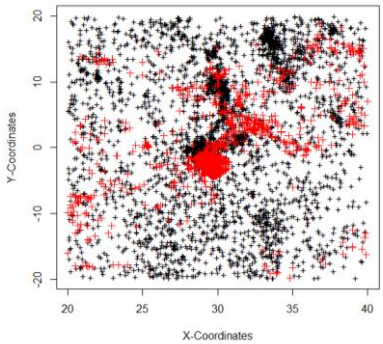


Figure 3-39 Superimposed events (red) with simulated (black) rpp using Covariates (2nd cases)

In the next stage we have Simulated IPP with spatial and temporal intensity used in STIKhat for 100 times and calculated STIKhat of those IPP for both with covariates and without covariate case and calculated the P-value from that and found TRUE for all the cases means for any of the considered distances, events are not Poisson but clustered at a significance level of 5%. The p-value plot is shown in figure 3-40 and 3-45 for the both cases.

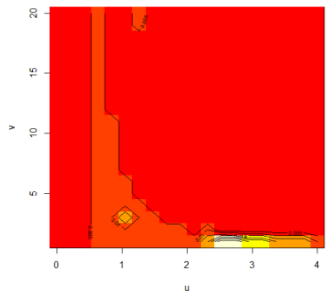


Figure 3-40 P-value for STIKhat with small u (up to 440 km) and v (up to 20 days) for the case events without covariates from 100 simulations

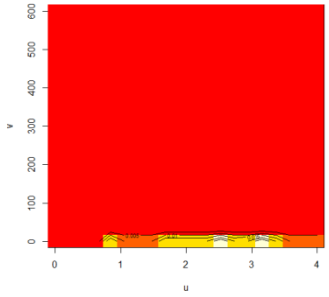


Figure 3-41 P-value STIKhat for small u (up to 440 km) and larger v (up to 2 years app.) for the case events without covariates from 100 simulations

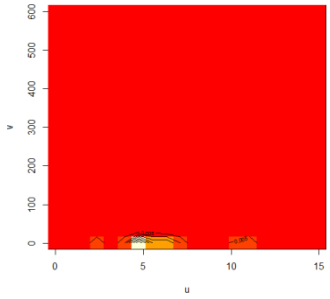


Figure 3-42 P-value for STIKhat with larger u (up to 1500 km) and v (up to 2 years app.) for the case events without covariates from 100 simulations

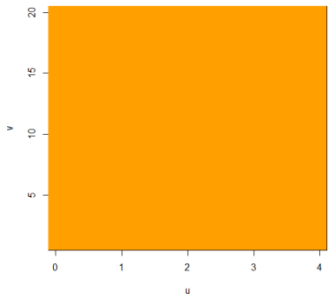


Figure 3-43 P-value for STIKhat with small u (up to 440 km) and v (up to 20 days) for the case events with covariates from 100 simulations

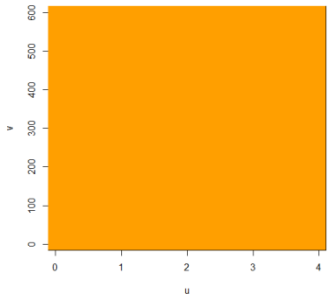


Figure 3-44 P-value STIKhat for small u (up to 440 km) and larger v (up to 2 years app.) for the case events with covariates from 100 simulations

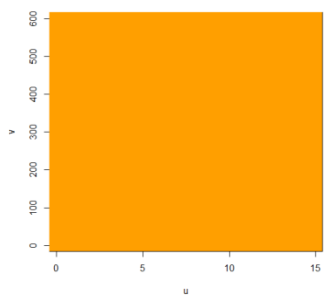


Figure 3-45 P-value for STIKhat with larger u (up to 1500 km) and v (up to 2 years app.) for the case events with covariates from 100 simulations

4 Chapter 4: Lattice Approach

4.1 Spatial cross-sectional models

Our starting point is the linear-in-parameters cross-sectional model. The number of study area's conflict experience *y* on change in covariates in the study area *X* is can be defined as

$$y = X\beta + u \dots \dots (23)$$

Where *u* is the classic error term and $u \sim N(0, \sigma^2)$. Then again, according to Tobler (1979) the first law of geography states the near things are more related than the distant things. So, the starting point of spatial modeling is based on the concept that near things are related. Based on this, our dependent variable, armed conflict will be spatially auto correlated if there is a dependency between armed conflicts and this dependency may diminish as the distance separating the conflict increases. The Spatial Autoregressive Model (SAR) captures this spatial change and dependency.

4.1.1 Spatial autoregressive model (SAR)

If the levels of the dependent variable *y* depend on the values of *y* in neighboring regions (for an instance the conflict in a particular cell depends on the covariate values of neighboring cell) a general model, incorporating spatial lags is formulated as (Cliff and Ord 1973):

$$y = \lambda W u + X\beta + u \dots \dots (24)$$

Where in each *N* location, *y* is a $(N \times 1)$ vector of observation on a dependent variable and *X* is a $(N \times k)$ matrix of exogenous variable. β is a $(k \times 1)$ vector of parameters and λ is a scalar spatial error parameter and *u* is a spatially auto correlated disturbance vector with content variance and covariance terms specified by a fixed spatial weights matrix and a single coefficient

$$\lambda = u \sim N(0, \sigma^2 (I - \lambda W)^{-1} (I - \lambda W')^{-1});\ I \text{ is the identity matrix.}$$

4.1.2 The spatial error model (SEM)

Spatial Error Model (SEM) omits variable bias including the autcorrelation term in spatial modelling. In our study another motivation for the spatial error model is spatial heterogeneity**.** The main idea is to include the direct, most important spatial variables in the model and leave the indirect spatial features to the residuals. This is done by specifying a global spatial autoregressive process in the error. According to Anselin (2003) the spatial influence comes only through the error terms and a general model for that can be defined as:

$$y = X\beta + u \text{ and } u = \rho Wu + v \ldots\ldots (25)$$

Where W is the weight matrix and λ is a spatial autoregressive parameter to be estimated jointly with the regression coefficients. With v assumed to be normal with $E(v) = 0, E(vv' = \sigma^2 I)$. The two vectors of errors are assumed to be uncorrelated. Solving the error specification for *u* we find that

$$(I - \rho W)u = v$$
$$u = (I - \rho W)^{-1} v$$

So, the final error model may be expressed as

$$y = X\beta + (I - \rho W)^{-1} v$$

According to Ord (1975) we have used maximum likelihood method for estimating the spatial error model and the log-likelihood function for the spatial error model can be defined as:

$$l(\beta, \lambda, \sigma^2) = -\frac{N}{2} ln2\pi - \frac{N}{2} ln\sigma^2 + \ln|I - \lambda W|$$
$$-\frac{1}{2\sigma^2}[(y - X\beta)'(I - \lambda W)'(I - \lambda W)(y - X\beta)]$$

4.1.3 Spatial durbin model (SDM):

The spatial lagging of the explanatory variables is added so that the characteristics of neighboring covariates could have an influence on the conflict in the sample (Brasington and Hite 2005). In this way the spatial Durbin model allows for neighboring covariates to determine the conflict in a particular cell, in addition to the characteristics of neighboring conflict.

Spatial Durbin Model which has a time-series equivalent and motivated by concern over spatial heterogeneity, could be developed from the spatial error model (Anselin 2006). This model includes a spatial lagging of the dependent variable besides the spatial lagging in independent variable and which allows capturing the effects for spatial autoregressive models. So the spatial lagging in covariates let us assess the effect of neighboring covariates values (Brasington and Hite 2005). If *a* is a vector of intercepts, which follows a spatial autoregressive process and suppose *X* and *u* in the SEM model are correlated then

$$a = X\gamma + v$$

Substituting this back into the SEM we get

$$y = \lambda Wy + X(\beta + \gamma) + WX(-\beta\gamma) + v \ldots\ldots (26)$$

So by putting such additional constraints on the parameters, the spatial Durbin model is specified as:

$$y = \rho Wy + \alpha ln + X\beta + WX\theta + u \dots \dots (27)$$

In SDM, the log-likelihood has a similar form to the SEM

4.2 Spatial cross-sectional modeling:

To find the relation between our independent variable armed conflict and dependent variable we have applied different spatial regression models discussed above. The regression models, applied in this study can be categorized in two groups. First, we have considered only spatial variability. This was done by aggregating all the data in one spatial layer. For an instance aggregated long term Soil Water Anomaly; Weighted Anomaly Standardized soil water index (WASWI) as the independent variable. These spatial models are being considered as **Model 1**. Under model 1 we have considered four different kinds of spatial models. Percent of cell's Armed Conflict Y on covariates (WASWI, SWI, SPI, and Population) of that particular cell X was estimated by following models (details discussed in the previous section):

1. Model 1.1: Ordinary least squares (OLS)
2. Model 1.2: Spatial autoregressive model (SAR)
3. Model 1.3: Spatial error model (SEM)
4. Model 1.4: Spatial Durbin Model: Likelihood function (SDM)

The output of different spatial autoregressive models is presented in table 4-1. From the table we observe that, model 1.1 (OLS) has found high correlation between armed conflict and all other covariates. When we have considered the neighboring cell characteristics, then we have found that SWI and SPI don't have any significant correlation with the dependent variable but WASWI and Population have significant correlation. Other models (model 1.3 and model 1.4) also give the same result. So from here we can say that WASWI and Population have significant correlation with armed conflict. WASWI is negatively correlated with WASWI means if the WASWI value reduces the armed conflict increases and for the population it's contrary. From the SDM model its observable that lagged armed conflict has also significant relation which means if a particular cell is more likely to experience armed conflict, the neighboring cell also likely to experiences armed conflict. On the other hand the lagged covariates do not have a significant effect on an armed conflict outbreak.

Table 4-1 Spatial autoregressive model output for the aggregated data for the year 1991 to 2000

	OLS	**SAR**	**SEM**	**SDM**
(Intercept)	0.00000	-0.0002	-0.0005	-0.0003
	0.01705	0.0139	0.0475	0.0139
WASWI (1991-2000)	-0.4717***	-0.1898***	-0.3374***	-0.1981**
	(0.03887)	(0.0327)	(0.0657)	(0.0865)
SWI (1991-2000)	0.08683***	0.0268	0.0611	-0.0455
	(0.0212)	0.0174	0.0447	0.0923
SPI (1991-2000)	0.04134*	0.0182	0.0163	0.0109
	(0.02117)	0.0173	0.0355	0.0463
POP (1991-2000)	0.5525***	0.21939***	0.3820***	0.2089**
	(0.03886)	(0.0331)	(0.0645)	(0.0839)
Lagged Armed Conflict (1991-2000)		0.7034***		0.7027***
		(0.0180)		(0.0182)
Lagged error (1991-2000)			0.70608***	
			(0.0181)	
Lagged WASWI (1991-2000)				0.0055
				(0.0995)
Lagged SWI (1991-2000)				0.0778
				(0.0966)
Lagged SPI (1991-2000)				0.0116
				(0.0521)
Lagged POP (1991-2000)				0.0163
				(0.0979)
AIC		7862.7	7875.9	7869.9
Number of observations	3200	3200	3200	3200
Moran's I Residuals	0.3557	-0.0035	-0.0051	-0.0030
Nagelkerke pseudo-R-squared	60.91[#]	0.3194	0.3166	0.3196
Moran's I Std. Deviate	42.4686	-0.4007	-0.598	-0.3356

the result derived from F-Test 'p ≤ :1; *p ≤ :05; **p ≤ :01; ***p ≤ :001

We have also calculated *Nagelkerke pseudo-R-squared* and which indicated that all of these models except OLS explains the relationship between dependent variable and independent variable explains on an average 32 percent. For all of these models Moran's I statistics were computed and we can see that in OLS models the data are clustered but in other models there are no clustering in residuals. From all these models The SAR results in a better fit.

4.3 Impacts in Spatial Lag models

In the Autoregressive models if there is neighborhood's covariates effect on the *i* cell which can be realized by λ then the summary measure of the impact β arising from changes in the observation in *i* cell of explanatory variables cannot capture the

complete impact. So, Lesage and Pace (2007) suggested using direct or indirect effect for better model interpretation. Direct effect composed of the estimated covariate's coefficient and the feedback effect. Here feedback effect can be realized through λ, which explains the covariate effect of *i* cell on the neighboring cell and this effect impose some additional effect on the *i* cell through the spatially autoregressive term, which is termed as feedback effect. On the other hand indirect impact estimates the spillover effect, means this impact measures the change of dependent variable in *i* cell due to the change in covariates in neighboring cells j ($i \neq j$) (LeSage and Pace 2009). So direct effect is the one that comes from the same region *i* and indirect effect is the one that comes from neighboring cells; in our case neighboring 8 cells as we considered queen neighbor. We have estimated the direct and indirect effect in the Spatial Durbin model (SDM) and the estimates have been reported in table 4-2

Table 4-2 impacofn spatial autoregressive model (SDM) output for the aggregated data for the year 1991 to 2000

	Coefficient	Direct	Indirect	Total
WASWI (1991-2000)	-0.1981**	-0.2193***	-0.4285***	-0.6479***
SWI (1991-2000)	-0.0455	-0.0383	0.1467	0.1084
SPI (1991-2000)	0.0463	0.0140	0.0621	0.0761
Population (1991-2000)	0.2089**	0.2348***	0.5229***	0.7577***

'p ≤ :1; *p ≤ :05; **p ≤ :01; ***p ≤ :001

In the table, we can see that both the direct and indirect effects are significant for the independent variables WASWI means the WASWI and population variables are affected by the impact measure and population and effects are not significant for SPI and SWI. And in both cases the indirect effects are significant and that implies the necessity of accounting for neighboring effect. If we excluded the neighbor weight we would interpret a biased result, which we got from the OLS model.

The direct impact of WASWI in any particular cell is -0.2113 means in a particular cell, if the WASWI value decreases by one unit that will result more conflict outbreak and will increase by 21 percent. Then again the indirect impact of the WASWI and Population are greater than the direct effect. So we interpret that neighborhood's WASWI and population value has a larger impact on the *I* cell's Armed conflict outbreak. But one important thing to note that this indirect effect is coming from neighboring 8 cells so if we say that one neighboring cell's WASWI value has -0.05362 measure of the impact on the dependent variable of *i* cell and which is -0.4290 for all neighboring cells. The neighboring cell effect also complies with the point process clustering results, as it was identified that the events are clustered in space up to 200 km (see chapter 3). On the other hand number of the population has positive direct and indirect impact on the conflict outbreak means if

the amount of population increase in a particular cell and in neighboring cells that has significant impact on the conflict outbreak in the cell *i*.

4.4 Spatio-temporal lattice approach modeling:

Suppose *y* is a set of dependent variable where $y = (y_1, \dots, y_n)'$ with n number of observations p is a set of covariates where$x_j = (x_{1j}, \dots, x_{nj})'$. Then regression models approximate dependent variable to the set of covariates by a linear function, can be defined as

$$y = \sum_{j=1}^{p} \beta_j x_j + \epsilon = X\beta + \epsilon \quad \dots\dots (28)$$

Where *X* is the design matrix where the location of x_{ij} is in row *i* and column *j* and β is the regression coefficient vector which can be estimated by minimizing the redidual sum of square, $\epsilon'\epsilon$ .Cressie and Wikle (2011) define the residual process $y - X\beta$ of a SAR model which follow an autoregressive process, is defined to be

$$y - X\beta = B(y - X\beta) + v$$
$$y = X\beta + (1 - B)^{-1}v \dots\dots (29)$$

where v assumed to be normal with $E(v) = 0, E(vv' = \sigma^2 I)$ and *B* defines which residuals are correlated and degree of correlation as well. Typically the value of B_{ii} = *0* but if y_i and y_j are neighbors then B_{ij} possess a non-zero value where the parameters of B defines the degree of autocorrelation which can be termed as λ. So for any non-zero B_{ij} , cells *i* and *j* are neighbors and $B_{ij} = \lambda$.

In our study, to define the spatial neighbors in **model 1.2** (spatial SAR model) we have used queen neighbors; 8 cells adjacent to each grid cell and we have defined the spatio-temporal neighbors according to Espindola et al. (2011).

For spatio-temporal regression modeling, the set of dependent variable observation are stamped with time where we have denoted $y_{[t]} = (y_{1,t}, \dots, y_{n,t})$ as the observation in grid cell *i* and time step $t \in \{1, \dots, m\}$. If *B* only denoted spatial neighbors only then a temporally lagged observation $y_{[t-1]}$ into the regression can be defined as

$$y_{[t]} = X\beta + \gamma y_{[t-1]} + (1 - B)^{-1}v) \dots\dots (30)$$

According to the temporally lagged SAR model defined above can be detailed for all time steps $t \in \{1, \dots, m\}$ according to Espindola, Pebesma, et al. (2011), where the *B* matrix will consider two temporal neighbors of $y_{i,t-1}$and $y_{i,t+1}$ and will also

consider spatial neighbors $y_{i,t}$ and $y_{j,t}$ with $i \neq j$. In our study this spatio temporal SAR model considered as **Model 2**. In this model the assumption is that, a single autocorrelation coefficient will define the correlation both in space and time.

Now, if we want to approximate a dependent variable to the set of covariate by a linear function which will fit coefficient to describe corrections between space, time and space time that can be done by extending model 2 with spatio-temporal neighbors. For an instance in model 3 the residuals $y_{i,t}$ and $y_{j,t+1}$ will be correlated when grid cells *i* and *j* are neighbors. In this model the simplifying assumption is, a single correlation coefficient will be fitted to describe correlations between all (spatial, temporal, and spatio-temporal) neighbors. In our study this model is considered as **model 3** to model the relationship between armed conflict and climate change. Figure 3 explains the concepts of space time neighbors of the temporally lagged SAR model, model 2 and model 3. We have used a maximum likelihood method (explained in SEM section) for estimating model 2 and model 3. For model 2 and 3 the regressions were carried out with the R function spautolm in R package spdep (Bivand et al. 2008) by "…defining neighbors in space and time combined with a weighting factor that defines how neighboring in space compares to neighboring in time, in terms of weights". The output of model 2 and 3 are provided in table 4-3

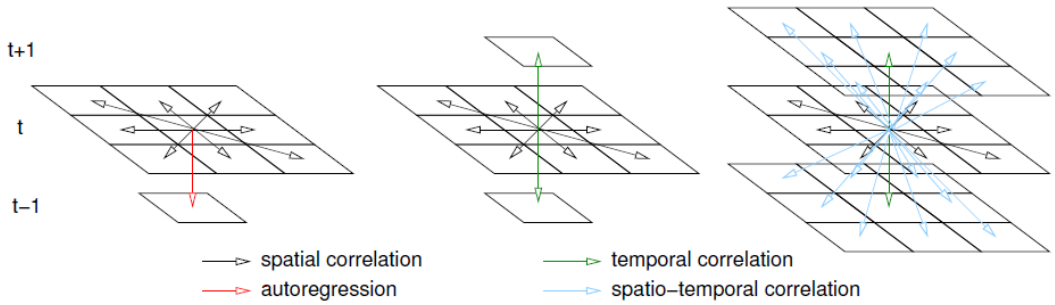


Figure 4-1 Neighbors addressed for of temporally lagged SAR model (left), mode 2 (middle) and model 3 (right). This figure is adapted from Espindola, Pebesma, et al. (2011).

Table 4-3 summarize the results obtained from our two space time regression SAR models devised for the period 1991 to 2000. These two models show that how our independent variable was impacted by the covariates, considering the covariates in space, time and space-time. In both of the model we can observe that dependent variable armed conflict (events) was highly impacted by the predictor armed conflict of the previous year and neighboring cell over space and time. In both of the model the autocorrelation coefficients of events were found to be significant at 0.001 level of significance for the whole period. From the list pf predictors, two predictors have shown a significant relation to the armed conflict; number of population and cumulative preceding three years WASWI sum (WASWI (yearly aggregated) ((t-1) + (t-2) + (t-3))). For both of the models we have also calculated Nagelkerke R-squared

values and that shows that model 2 explains the impact better than model 3 and AIC value also indicated that the model 2 was better fitted then model 3.

Table 4-3 Spatio-temporal autoregressive model (SAR) output for the dis-aggregated data for the year 1991 to 2000

	model 2	**model 3**
Intercept	0.0001	0.0004
	(0.0126)	(0.0121)
WASWI (yearly aggregated)	0.0003	0.0018
	(0.0097)	(0.0074)
WASWI (yearly aggregated) (t-1)	-0.0027	0.0057
	(0.0098)	(0.0075)
WASWI (yearly aggregated) (t-2)	0.0016	0.0016
	(0.0185)	(0.0185)
WASWI (yearly aggregated) ((t-1)+(t-2))	0.0042	0.0042
	(0.0277)	(0.0277)
WASWI (yearly aggregated) ((t-1) + (t-2) + (t-3))	-0.0330**	-0.0329**
	(0.0194)	(0.0194)
SWI (yearly aggregated)	0.0053	-0.0054
	(0.0219)	(0.0157)
SWI (yearly aggregated) (t-1)	0.0184	0.0237
	(0.0219)	(0.0153)
SWI (yearly aggregated) (t-2)	0.0158	0.0157
	(0.0449)	(0.0449)
SWI (yearly aggregated) ((t-1) + (t-2))	0.0849	0.0849
	(0.0656)	(0.0655)
SWI (yearly aggregated) ((t-1) + (t-2) + (t-3))	-0.0615	-0.0614
	(0.0544)	(0.0543)
SPI (yearly aggregated)	-0.0136	-0.01824
	(0.0102)	(0.0076)
SPI (yearly aggregated) (t-1)	0.0106	-0.0058
	(0.0101)	(0.0076)
SPI (yearly aggregated) (t-2)	0.0006	0.0006
	(0.0246)	(0.0245)
SPI (yearly aggregated) ((t-1) + (t-2))	0.0229	0.0229
	(0.0390)	(0.0389)
SPI (yearly aggregated) ((t-1) + (t-2) + (t-3))	-0.0131	-0.0130
	(0.0217)	(0.0216)
POP (yearly aggregated)	0.0617***	0.0660***
	(0.0179)	(0.0193)
Lagged Armed Conflict	0.64363***	0.61093***
	(0.0091)	0.013084)
Number of observations	28800	28800
AIC	67689	69718
R square	0.38637	0.34158

'p ≤ :1; *p ≤ :05; **p ≤ :01; ***p ≤ :001

5 Chapter 5: Discussion and Conclusion

Armed conflict, in general, is a result of social, economical, political or a combined effect of all these determinants. A change in one of these factors or all at the same time can causes armed conflict. Economists have recently started to work on conflicts and political scientists are doing it for many years. The things they have been doing is trying to associate different approximates or determinants of conflict with conflict incidents for better understanding of the cause of armed conflict and prediction. Most studies have been done on SSA and that's because the continent experienced most of the armed conflicts so far. About 2/3 of the total conflicts observed in SSA. Clearly, these conflicts are one of the most important causes of development constrain in this continent. So, if we are trying to develop policy in such area it's very important to understand what are the preconditions, might prevent or stop conflicts and that's how people are getting engaged with this sort of topic.

The climate comes in because scientist has found relatively robust correlation between economic shock induced by weather fluctuations and which have significant impact on conflict growth. Such as Miguel et.al in 2004 has identified the impact of income shocks on GDP growth exploiting the fact that in SSA the vast majority of the population depends on the rain fed agriculture and fluctuation in precipitation have an impact on agriculture, results lower GDP growth. There are also several papers which also done a similar exercise using different variables like temperature change and conflict (Ciccone 2011). But the common features of these studies are, they rely on cross-country evidence. These studies have considered the average of the variables (e.g., temperature, rainfall) across the country over the year and observed how it impacts on conflict growth. Then again, there are arguments that Sub national disaggregated studies may offer more support for the resource conflict nexus. So this study is possibly a step further in studying these relationship taking the analysis into a different scale. This study followed by a geographically disaggregated unit of observation (0.5 degree x 0.5 degree; roughly corresponds to 55 km x 55 km in this part of the world) and temporal disaggregation (yearly) with combining high-resolution conflict and climate datasets deploying various approaches to model the relationship between armed conflict and climate change.

To capture the climate variability, we have introduced a climate indicator termed as Weighted Anomaly Soil Water Index (WASWI), which is derived from the Soil Water Index (SWI). WASWI captures the variation of precipitation, evaporation and temperature at the same time. This index has not been used in climate security literature yet. This index makes us able not to look at temperature or rainfall alone but the interplay of these factors. If we consider soil moister or water containment in

the soil as natural resource then this study is a more organized support for the research on climate and conflict nexus.

Besides disaggregated analysis approach and new climate indicator, this study also contributes a sound method in climate conflict literature as this study focuses on cell specific yearly variation combining two major wings of probability modelling, point process approach and spatial autoregressive modelling. Then again, this study might be the first which considered the disaggregated level of analyses in space and space-time with careful consideration of spatial dependence in modelling and spatio-temporal autocorrelation etc.

5.1 Point process models

Point process models were used to characterize the distribution of conflict location in the study area, exploiting properties of point process. First order properties were used to understand the trend of distribution of points in the pattern and second order properties or model fitting method was used to extract the information on possible interaction. We have fitted the empirical data with a different point process model using an inhomogeneous version of K-function which is basically a distance function, estimates the number of points that were expected to be found within a particular distance, centered in any arbitrary points. If the number of points observed in a particular distance is greater than expected, then we know that there is a clustered pattern in the point dataset and we have found out event observations to be clustered.

5.1.1 Spatial point process model fitting and selection

First order properties show that the conflicts sites are not homogeneously distributed and we can see some hot spots within the study area. This result rejects the null hypothesis of random distribution of events. Under second order properties, several models were fitted to characterize the data and to understand the interaction process between points and covariate effect or the trend of the point distribution. From the CSR test we can see that our data have inhomogeneous distributional characteristics and we basically assumed that conflict distribution is not uniform as the covariate's distribution is not regular. The trend was observed by fitting inhomogeneous Poisson process (IPP) model and inhomogeneous version of K-function successfully classified the conflict data as a clustered pattern. From the IPP plot, we can see that the events are clustered up to 200 km. Though this model explains the effect of covariates (clustering because of these environmental factors) but as the empirical line fall outside the simulated envelope, this model is not adequate to explain the distribution (or clustering) of conflict events. We assume that this could be caused by the interaction (possibly the point attracts each other or push each other) between

points. So, still we needed to fit interaction model which is basically some cluster process models and area interaction model to explain the part of clustering which is not explained by the covariate effect. These higher order models can give information on effect of covariates, interaction between points and interaction between clusters in a point pattern. Besides, these models also can organize prediction to find where we can expect to find at the conflict in the study area.

By following the theory of Poisson Cluster process (discussed in chapter 3), we tried to fit Matern cluster process model and Thomas cluster process model without empirical event data using covarites. Both of these cluster models have given us a better fit than the IPP model but still it's not good enough which can best constitute an intensity function for conflict distribution.

So in the next step, we go for the area interaction (AI) model which included in the covariate affect on point distribution and interaction. The diagnosis of this model was performed by goodness-of-fit; for statistical significance the simulated envelope was recoded for all models. Non-stationary area interaction process gave the fit with 95% confidence level, shows that the inhomogeneous version of the empirical k-function fall inside the simulated envelope and tell us that interaction is a reason that causes the clustering within the pattern. We can also see the interaction coefficient η of this model is far away from 1, that means there is an interaction trend and it's not poisson process. This also explains the cause of the inadequacy of the earlier models. So conflict occurs near to another conflict location and this behavior can be explained by the location of administrative boundary. Such as the conflict intensity is higher in Burundi and Ruanda (the conflict hotspot identified in the middle of the study area) and the conflict distribution in that area are confined by surrounding international boundaries. Besides these, the clustering in that particular area was also followed by the effect of covariates like higher population density and the lower water contained in the soil. This fitted AI model actually gives us some formulas which make us able to predict and the coefficient of these factors, which explains both the direction and strength that these variables has on conflict distribution. For example, the coefficient of WASWI exhibits that change of one unit in WASWI, the average change in the mean of Armed Conflict will be about -0.1647 units. Thus the WASWI is negatively related to the conflict incidents. This model is good enough to make prediction with 95 % confidence level.

Based on the similar type of climatic condition (e.g., relatively dry region) and inter events interaction (surrounded by other conflict events, which could drive a particular location towards experiencing conflict), there were some potential location which might have experienced conflict but in reality they did not. This is one of the limitation of our model, which cannot explain all the reasons of armed conflict . These limitations can be overcome by introducing more explanatory variable like

socioeconomic conditions of that location. For an instance, several authors argued that climate impacts on income and income shock drive to conflict. But if there were means (e.g., a region with higher GDP) to overcome those economic shocks which can stop the conflict may cut the probability of conflict occurrences. Adjustments of such limitations can lead us towards more successful prediction. So, this AI model can be improved by introducing all the reason (e.g., socio-political covariates such as administrative boundary) identified in conflict literature behind armed conflict but that was beyond our study goals.

5.1.2 Spatio-temporal point process modeling

We also have fitted several models for yearly conflict data using yearly covariates of the same year and previous years (e.g., 1997 event data set fitted with Matern Cluster process using covariates of 1997 (t-0), 1996 (t-1) and 1995(t-2)) and considering the yearly inter point interaction . K-function were being observed from 99 simulations for each model. This is being done to discuss the assumption that climate condition of a particular year can trigger the conflict in the next years. For an instance, we tried to check the fact that the probability to find an Armed Conflict in a particular location depends on the value of WASWI of previous years ((t-1) and (t-2)). The cluster models fitted well with the yearly empirical event data and we also can observe that inhomogeneous version of empirical K-function fall inside the envelope, which means the yearly cluster models are good enough to explain the yearly event intensity based on the covariate effect and inter point interaction. After plotting the Standardized fitted coefficients for Inhomogeneous cluster point process models we can observe that the covariates keep up a nonlinear relationship with the events, though the directions are mostly positive for population and negative for WASWI and SPI (Fig 5-1).

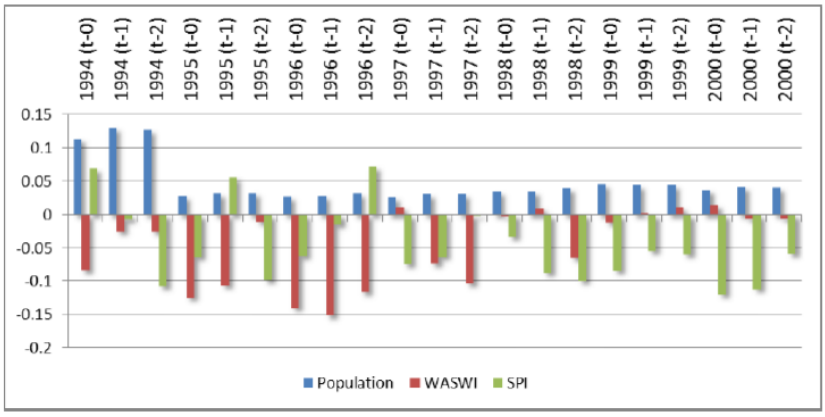


Figure 5-1 Standardized fitted coefficients for Inhomogeneous cluster point process models

The event intensity in a particular location of a particular year depends on the covariates of that year and previous years and they are negatively correlated. From 1994 to 1998 WASWI has greater impact on conflict outbreak and tends to reduce afterwards. The reason can be explained exploiting the fact that the change in

socioeconomic conduction was controlling the conflict outbreak after 1998 but this interpretation can't have enough confidence due to smaller temporal extent (1998-2000). Further study recommended with higher temporal extent. The number population was always positively correlated with conflict and maintained a persistent effect on conflict clustering.

For a better understanding we have gone a step further by characterizing the data and modelling the relationship, using Space time inhomogeneous Point Process Modelling. The Inhomogeneous space time K-Function (STIKhat) was used to classify the data according to pattern (e.g., clustering or regularity), which can be considered as a dynamic space time point process approach. This approach has helped us to establish the relationship based on spatial (which has been explained in the previous section) and introducing the dynamic temporal dimension (which has not been considered while modelling yearly point process). Space time point process modelling has filled the gap in understanding the space time clustering in a dynamic way (cluster in space and cluster in time; relative measure). From STIKhat estimation we can observe that the events do not behave like an inhomogeneous spatio-temporal Poisson process (that support the spatial point process modelling), on the contrary they show space-time interaction. When we control the variability of the covariate information, this interaction is in the form of aggregation in the sense that events tend to come in clusters from spatial distances up to 500 km and temporal distances up to 2 years. Note that the values of the STIKhat are positive and far from zero indicating that the part coming from the spatio-temporal K function is stronger than that coming from the theoretical value under the Poisson case. Our significance test also indicates that for any of the considered distances, events are not Poisson but clustered at a significance level of 5%.

5.2 Lattice approach

From the spatial and spatio-temporal point process modelling we have found that there is a climate conflict relationship in our study area. To cross check the fact in the next step we have considered both spatial and spatio-temporal lattice approaches by taking the advantage of the spatial nature of the data and try to look at from different ways in which correlation can affect the occurrence of conflict. By nature, the climate data are very likely to be correlated in space. We can observe same spatial pattern in the space-time plot of WASWI, SPI etc. Then again, from the point pattern analyses we can observe the spatial pattern of in the conflicts and they are clustered up to 200 km. This result supports the fact that, when we observe conflict in one cells it's very possible to spill over to neighboring areas. So we have fit different models in different spatial and spatio-temporal sense which accommodate the facts like spatial dependence, spatial and temporal autocorrelations etc. This approach also is

to prove the relative importance of the determinant factors of armed conflict for the entire period (1991-2000). Autocorreltion identified the non-randomness in the data.

5.2.1 Spatial regression models

We have considered several spatial regression models to understand the relationship between armed conflict and climate change. Several independent variables were introduced. Such as WASWI, SPI, SWI and population. Where, WASWI is a dimensionless measure of the relative (spatio-temporal) severity of water contained in the soil (surplus or deficit of water in the soil) in a particular cell. Spatial regression models we estimated based on aggregated data for the whole period (1991-2000). For an instance, agreegated WASWI redirects the overall surplus or deficit of water in the soil for 10 years. On the other hand SPI also indicates the same characteristics but the basic difference between SPI and WASWI is WASWI was estimated based on the study extend for 10 years. And SPI was estimated based on the whole world for 100 years, which is a global variable and using this sort of variables it's hard to model the local variation.

Computing ordinary least squares (OLS) regression analyses we tried to model the relationship between the conflict and climate variables. All the independent variables were found to be significant, in OLS regression model, where the basic assumption is all the observations are independent in space which is highly unrealistic. To accommodate the gap we have performed SAR analyses by considering autocorrelation of the dependent variables in space which was done by using a spatial weight matrix (e.g., considered queen neighbors). Then again, by considering queen neighbors we have taken the clustering distance into account.

In SAR model the regression coefficient for variables WASWI and Population were found to be significant for at least .05 significance level but the SPI and SWI is not significant. As discussed before the independent variables are also correlated in space and to accommodate that fact in modeling, Spatial Durbin Model (SDM) was performed which consider the lagged independent variable in the modeling. To compare the relative importance of each predictor we have presented the regression coefficient derived from SDM and corresponding standard errors. The comparison showed how conflict occurrences were impacted by spatial predictors WASWI and Population in own cell and neighboring cells in space. So, spatial autoregressive models conclude that there are significant climate conflict relationship and WASWI is a significant predictor of armed conflict in the study area. Besides, if we observe the Lagged Armed Conflict we can see that the degree of correlation of the dependent variable is very high and around 0.7, which means the conflict of particular cell are affected by the neighboring armed conflict experiences but lagged covariates are not that much significant. Then again if we observe the AIC value of

all three autoregressive models we can see that all the models have a very similar kind of value, means all models are having similar kind of fit. Note, in spatial modelling we have explored the relationship between cumulative 10-years total and that were found to be significant but at this stage we cannot say anything about the change in conflict for a weather shock. For the model adequacy and performance evaluation, we have calculated the Nagelkerke pseudo-R-squared (coefficient of determination) for all of these models and we have found that the models can explain the relationship almost 32 %.

In the next stage we have computed the impact in spatial autoregressive model (SDM) output fo the aggregated data for the year 1991 to 2000. We can see that both the direct and indirect effects are significant for the independent variables WASWI and Population. Which means the WASWI and population variables are affected by the impact measure and effects are not significant for SPI and SWI. And in both cases the indirect effects are significant and that implies the necessity of accounting for neighboring effect. If we excluded the neighbor weight we would interpret a biased result. Note that this indirect effect is coming from neighboring 8 cells. The neighboring cell effect also complies with the point process clustering results, as it was identified that the events are clustered in space up to 200 km.

5.2.2 Spatio-temporal regression models

We have extended the Spatial Autoregressive (SAR) model by defining spatial, temporal and spatio-temporal neighbors in auto regression to accommodate correlation component in space and time (e.g., temporal and spatio-temporal correlations). In addition to the spatial autoregressive effect of the residuals (which explains the difference between the actual value of a dependent variable and the value that was predicted by the statistical models), we have incorporated temporally lagged observation into the regression models. For the space time SAR model, one simple assumption was that, a single autocorrelation coefficient describes the correlation both in space and time neighbors (for model 2) and for model 3, which was an extended version of model 2 by introducing spatio-temporal neighbors in model 2. Again in model 3, a single coefficient was fitted to describe the correlation between spatial, temporal and spatio-temporal neighbors.

To compare the importance of the predictors (considering the space time neighbors than considering only spatial neighbors) we have presented the regression coefficient and the corresponding standard error with associated level of significance. Here, in model 2 and model 3 we have included a 1-year lag, 2-year lag, and a cumulative 2-year totals and cumulative 3-year totals of the independent variable WASWI and SPI than yearly totals. For both model 2 and 3 very few variables were found to be significant for the time period. The simple cumulative 3-year totals (can be

represented by ((t-1)+(t-2)+(t-3))) of WASWI proved the most significant in all comparisons. Which means, soil dryness experience 3 years in a row of a particular cell and in its neighbors, is likely to trigger armed conflict in that cell and in neighboring cells. This significance complies with the result of the spatial autoregressive model output which has considered cumulative 10-years total of WASWI and found to be significant. On the other hand the other alternative lag specifications 1-year lag, 2-year lag and cumulative 2-year totals have found to be insignificant, predicting armed conflict.

The single autocorrelation coefficient for the both model 2 (correlation in space and time) and 3 (correlation in space, time and space-time) are found higher and significant and is around 0.6 for both model. And that gives us a clear picture of how the conflicts in neighbors in space and time are affecting the conflicts of a particular cell this result also complies with the space time point pattern analyses. We have also calculated R square value for model 2 and 3 to see how much variability in the dependent variable remains unexplained by the predictors in the model. We have found that the r-square value for the model 3 was higher and it was around 0.39. So we can conclude that model 3 can better explain the armed conflict variability based on independent variable better than any other model.

5.2.3 Climate change and armed conflict

Point pattern analyses can model the relationship between a dependent variable and independent variable based on the interaction among dependent variable and variability in the explanatory variables (e.g., population, WASWI etc.). On the other hand Regression analysis attempts to model one variable as a function of one or more explanatory variables, such as trying to predict armed conflict occurrences based on climate change indicators (e.g., WASWI, interplay of climate indicators like temperature, evapotranspiration and rainfall etc.). In our study we considered both approaches and found that climate change is a significant predictor of armed conflict and can explain the occurrence of armed conflict very efficiently.

In both point and lattice approaches and in different models we have found that armed conflict has a negative and significant relationship with WASWI which means if the WASWI value decreases the armed conflict increases. In other word, if the anomaly in soil water containment increases the armed conflict increases and as soil water containment is a function of several climate indicators like temperature, rainfall etc., we can say that climate change has a significant impact on the armed conflict outbreak. So there is a likely link between climate change and armed conflict and local resource scarcity (e.g., soil water containment) in terms of climate change offers a better prediction of conflict behavior.

We also see that climatic factors effect locally and that trigger conflict locally and one particular conflict have impact on another conflict up to 200 km and it is less probable that one particular conflict will impact a larger region. So conflicts are very much likely to be triggered by climate anomaly and persist locally, both in space and time.

Sub national disaggregated studies may provide more support for the resource conflict nexus and we find it to be true. Because our cell/point based study can give a very clear picture of climate conflict relations and can predict conflict very competently. Such as, change in WASWI, impacts change in armed conflict by -0.1981 or -0.1657 and Conflict in the own cell associated with a (0.3651) increase in the probability of conflict of the following year. So, climate change indicator; long term WASWI measured at the cell level is a strong local conflict predictor. We have also found that climate measures of a particular year don't have a significant effect on an armed conflict outbreak of the following year but climate change (long term measure) has a significant effect on armed conflict outbreaks. So these results also accept such hypothesis, that in future the conflict situation is going to be worse due to climate change if it is not taken care of.

This kind of study requires a significant amount of time and data to have a critical look using different standard methods. This study was sound from the methodological perspective but it's also true that it was a limited monitoring approach as the time period considered in this study is only 10 years where climate related study requires a longer time period than that. Then again, probability of conflict risk increases by increasing climate induced poverty but strong financial and bureaucratic basis can reduce the probability significantly. This study was able to explain the first part of the argument but later part was unattended. By introducing socioeconomic and political indicators that can be minimized and might provide more clear empirical data dependent sound models which can predict conflict more precisely.

5.3 Conclusion

Disaggregated climate conflict study in subnational level (or regional level) provides more support to the resource-conflict and climate-conflict nexus. This study has established the effect of climate change on conflict occurrence and provides an indication of future conflict scenario incurred by climate change. The conflict scenario is going to be worse due to global warming if the issue left unattended and might question the future security. But modelling the climate conflict relationship we can make a climate change related conflict risk prediction and can answer tomorrow's World Peace today.

Recently, climate change has been realized as a risk multiplier. In particularly the increase of poverty, food and water scarcity risks Incurred by climate change were found to be significant and seems very difficult to address. These unhelpful circumstances are increasing regional instability and regional variability are questioning the international stability and security. To address such issues more efficiently it is very much necessary to understand the relationship between climate and regional instability. By modelling the climate conflict relationship we can start answering the question to regional instability and give the researcher tools to fight with climate incurred security risk.

Space Advisory Group (SAG) is the group that is advising the European Commission on the definition of the space theme within the next Framework Programme for Research and Development, e.g., The Horizon 2020. According to the recommendations of the SAG, it is very much necessary to understand the mechanisms, leading the gradual increasing risks incurred by climate change to address the security issues efficiently (SAG 2012). This study is a start point towards climate related risk understanding and prediction. Such understanding will guide us mitigating the gradual increasing risks incurred by climate change.

6 Referances

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review, 26*(2), 153-166.

Anselin, L. (2006). Spatial Econometrics. In T. C. Mills & K. Patterson (Eds.), *Palgrave Handbook of Econometrics* (Vol. 1): Palgrave Macmillan.

Arbia, G. M., Espa, G., Giuliani, D., & Mazzitelli, A. (2009). Clusters of firms in space and time: Department of Economics, University of Trento, Italia.

Baddeley, A. (2010). *Analysing spatial point patterns in R.* Paper presented at the A detailed set of workshop notes on analysing spatial point patterns using the statistical software package 'R'.

Baddeley, A., & Lieshout, M. (1995). Area-interaction point processes. *Annals of the Institute of Statistical Mathematics, 47*(4), 601-619.

Baddeley, A., & Turner, R. (2005). Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software, 12*(6), 1 - 42.

Baddeley, A., & Turner, R. (2006). Modelling Spatial Point Patterns in R. In A. Baddeley, P. Gregori, J. Mateu, R. Stoica & D. Stoyan (Eds.), *Case Studies in Spatial Point Process Modeling* (Vol. 185, pp. 23-74): Springer New York.

Baddeley, A. J., Møller, J., & Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica, 54*(3), 329-350. doi: 10.1111/1467-9574.00144

Balk, D., Pozzi, F., Yetman, G., Nelson, A., & Deichmann, U. (2004). *Methodologies to improve global population estimates in urban and rural areas.* Paper presented at the Annual Meetings of the Population Association of America, Boston, March-April.

Barron, P., Kaiser, K., & Pradhan, M. (2009). Understanding Variations in Local Conflict: Evidence and Implications from Indonesia. *World Development, 37*(3), 698-713. doi: 10.1016/j.worlddev.2008.08.007

Bernauer, T., Koubi, V., Kalbhenn, A., & Spilker, G. (2012). Climate Variability, Economic Growth, and Civil Conflict. *Journal of Peace Research, 49*(1).

Bivand, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2008). *Applied spatial data analysis with R*: Springer.

Bloom, D. E., & Sachs, J. D. (1998). Geography, Demography, and Economic Growth in Africa. *Brookings Papers on Economic Activity, 29*(2), 207-296.

Bossdorf, O., Schurr, F., & Schumacher, J. (2000). Spatial patterns of plant association in grazed and ungrazed shrublands in the semi-arid Karoo, South Africa. *Journal of Vegetation Science, 11*(2), 253-258. doi: 10.2307/3236804

Brancati, D. (2007). Political Aftershocks: The Impact of Earthquakes on Intrastate Conflict. *Journal of Conflict Resolution, 51*(5), 715-743. doi: 10.1177/0022002707305234

Brasington, D. M., & Hite, D. (2005). Demand for environmental quality: a spatial hedonic analysis. *Regional Science and Urban Economics, 35*(1), 57-82.

Buhaug, H. (2010). Climate not to blame for African civil wars. *Proceedings of the National Academy of Sciences, 107*(38), 16477-16482.

Burkea, M. B., Miguelc, E., Satyanathd, S., Dykemae, J. A., & Lobellb, D. B. (2009). Warming increases the risk of civil war in Africa. *National Academy of Sciences of the United States of America, 106*(49), 20670-20674.

Burkea, M. B., Miguelc, E., Satyanathd, S., Dykemae, J. A., & Lobellb, D. B. (2010). Climate robustly linked to African civil war. *Proceedings of the National Academy of Sciences, 107*(51), E185.

Burnley, C., Buda, D., & Kayitakire, F. (2008). Quantitative global model for armed conflict risk assessment.

Cliff, A. D., & Ord, J. K. (1973). *Spatial autocorrelation*. London: Pion.

Collier, P., Conway, G., & Venables, T. (2008). Climate change and Africa. *Oxford Review of Economic Policy, 24*(2), 337-353.

Collier, P., & Hoeffler, A. (2004). Greed and Grievance in Civil War. *Oxford Economic Papers, 56*, 563-595.

Cressie, N., & Wikle, C. K. (2011). *Statistics for spatio-temporal data* (Vol. 465): Wiley.

Cressie, N. A. C. (1993). *Statistics for Spatial Data, revised edition* (Vol. 928): Wiley, New York.

Diggle, P. (1985). A kernel method for smoothing point process data. *Applied Statistics*, 138-147.

Diggle, P., Chetwynd, A., Häggkvist, R., & Morris, S. (1995). Second-order analysis of space-time clustering. *Statistical Methods in Medical Research, 4*(2), 124-136. doi: 10.1177/096228029500400203

Diggle, P. J. (1983). *Statistical analysis of spatial point patterns*: Academic Press (London and New York).

Diggle, P. J. (2003). *Spatial Analysis of Spatial Point Patterns* (2nd ed.). London: Edward Arnold.

Diggle, P. J., & Gratton, R. J. (1984). Monte Carlo Methods of Inference for Implicit Statistical Models. *Journal of the Royal Statistical Society. Series B (Methodological), 46*(2), 193-227. doi: 10.2307/2345504

Espindola, G. M. d., Pebesma, E., & Cˆamara, a. G. (2011). *Spatio-temporal regression models for deforestation in the Brazilian Amazon.* Paper presented at the The International Symposium on Spatial-Temporal Analysis and Data Mining, University College London.

Fearon, J. D. (2005). Primary Commodity Exports and Civil War. *Journal of Conflict Resolution 49*(4), 483-507.

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review, 97*(01), 75-90. doi: http://dx.doi.org/10.1017/S0003055403000534

Gabriel, E., & Diggle, P. J. (2009). Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica, 63*(1), 43-51. doi: 10.1111/j.1467-9574.2008.00407.x

Gabriel, E., Rowlingson, B., & Diggle, P. J. (2012). stpp: An R Package for Plotting, Simulating and Analysing Spatio Temporal Point Patterns. *Journal of Statistical Software*.

Giuseppe, A., Giuseppe, E., Diego, G., & Andrea, M. (2009). Clusters of firms in space and time: Department of Economics, University of Trento, Italia.

Gleditsch, N. P. (1998). Armed Conflict and The Environment: A Critique of the Literature. *Journal of Peace Research, 35*(3), 381- 400.

Gleditsch, N. P., Buhaug, H., & Theisen, O. M. (2011). *Climate Change and Armed Conflict.* Paper presented at the Sixth General Conference, European Consortium for Political Research, University of Iceland.

Guttman, N. B. (1999). ACCEPTING THE STANDARDIZED PRECIPITATION INDEX: A CALCULATION ALGORITHM1. *JAWRA Journal of the American Water Resources Association, 35*(2), 311-322.

Hauge, W., & Ellingsen, T. (1998). Beyond Environmental Scarcity: Causal Pathways to Conflict. *Journal of Peace Research, 35*(3), 299 - 317.

Hegre, H., Ellingsen, T., Gates, S., & Gleditsch, N. P. (2001). Toward a democratic civil peace? Democracy, political change, civil war, 1816-1992. . *American Political Science Review, 95*(1), 33-48.

Hendrix, C. S., & Salehyan, I. (2012). Climate change, rainfall, and social conflict in Africa. *Journal of Peace Research, 49*(1), 35-50.

Homer-Dixon, T. F. (1991). On the threshold: environmental changes as causes of acute conflict. *International security*, 76-116.

Homer-Dixon, T. F. (2001). *Environment, scarcity, and violence*: Princeton University Press.

Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns*. London: John Wiley & Sons.

Kahl, C. H. (2006). *States, scarcity, and civil strife in the developing world*: Princeton University Press.

Lakshmi, V., Jackson, T. J., & Zehrfuhs, D. (2003). Soil moisture–temperature relationships: results from two field experiments. *Hydrological Processes, 17*(15), 3041-3057. doi: 10.1002/hyp.1275

Le Billon, P. (2001). The political ecology of war: natural resources and armed conflicts. *Political Geography, 20*(5), 561-584. doi: 10.1016/s0962-6298(01)00015-4

Lecoutere, E., D'Exelle, B., & Van Campenhout, B. (2010). Who Engages in Water Scarcity Conflicts? A Field Experiment with Irrigators in Semi-arid Africa *MICROCON Research Working Papers* (pp. 1-27): MICROCON: A Micro Level Analysis of Violent Conflict, Institute of Development Studies at the University of Sussex.

Lee, J. R. (2009). *Climate Change and Armed Conflict: Hot and Cold Wars*: Routledge.

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics* (Vol. 196): Chapman & Hall/CRC.

LeSage, J. P., & Kelley Pace, R. (2007). A matrix exponential spatial specification. *Journal of Econometrics, 140*(1), 190-214.

Levy, M. A., Thorkelson, C., Vörösmarty, C., Douglas, E., Humphreys, M., & Hampshire, N. (2005). *Freshwater availability anomalies and outbreak of internal war: Results from a global spatial time series analysis.* Paper presented at the international workshop on 'Human Security and Climate Change, Holmen, Norway.

Low, P. S. (2005). *Climate change and Africa*: Cambridge University Press.

Lynch, H. J., & Moorcroft, P. R. (2008). A spatiotemporal Ripley's K-function to analyze interactions between spruce budworm and fire in British Columbia, Canada. *Canadian Journal of Forest Research, 38*(12), 3112-3119. doi: 10.1139/x08-143

Lyon, B., & Barnston, A. G. (2005). ENSO and the spatial extent of interannual precipitation extremes in tropical land areas. *Journal of Climate, 18*(23), 5095-5109.

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). *The relationship of drought frequency and duration to time scales.* Paper presented at the Proceedings of the 8th Conference on Applied Climatology.

McKee, T. B., Doesken, N. J., & Kleist, J. (1995). *Drought monitoring with multiple time scales.* Paper presented at the Ninth Conference on Applied Climatology. American Meteorological Society, Boston.

Melander, E., & Sundberg, R. (2011). Climate change, environmental stress, and violent conflict: tests introducing the UCDP Georeferenced Event Dataset. *International Studies Association, March*, 16-19.

Miguel, E. (2005). Poverty and Witch Killing. *The Review of Economic Studies, 72*(4), 1153-1172.

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy, 112*(4), 725-753.

Møller, J. (1999). Markov chain Monte Carlo and spatial point processes. In O. E. Barndorff-Nielsen, W. S. Kendall & M. N. M. v. Lieshout (Eds.), *Stochastic Geometry. Likelihood and Computation.* (pp. 141-172). Boca Raton: Chapman and Hall/CRC.

Møller, J., & Ghorbani, M. (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes. *Statistica Neerlandica, 66*(4), 472-491. doi: 10.1111/j.1467-9574.2012.00526.x

Nel, P., & Righarts, M. (2008). Natural Disasters and the Risk of Violent Civil Conflict. *International Studies Quarterly, 52*(1), 159-185. doi: 10.1111/j.1468-2478.2007.00495.x

Neyman, J., & Scott, E. L. (1958). Statistical Approach to Problems of Cosmology. *Journal of the Royal Statistical Society. Series B (Methodological), 20*(1), 1-43. doi: 10.2307/2983905

Onda, C. (2008). The Effect of the Spatial Resolution of Conflict Data on the Analysis of Drought as a Determinant of Civil War Onset: Africa, 1980-2001. *Consilience - The Journal of Sustainable Development, 1*(1), 52-64.

Oneal, J. R., & Russett, B. (2005). Rule of Three, Let It Be? When More Really Is Better. *Conflict Management and Peace Science, 22*(4), 293 - 310.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association, 70*(349), 120-126.

Palmer, W. C. (1965). *Meteorological drought*: US Department of Commerce, Weather Bureau Washington, DC, USA.

Posner, D. N. (2004). Measuring Ethnic Fractionalization in Africa. *American Journal of Political Science, 48*(4), 849-863. doi: 10.1111/j.0092-5853.2004.00105.x

Pradhan, Y., Brooks, M., & Saunders, R. (2011). ESCAT soil moisture climatology (Vol. 551). United Kingdom: Met Office.

Raleigh, C., & Kniveton, D. (2012). Come rain or shine: An analysis of conflict and climate variability. *Journal of Peace research, 49*(1).

Raleigh, C., & Urdal, H. (2007). Climate change, environmental degradation and armed conflict. *Political Geography, 26*(6), 674-694. doi: 10.1016/j.polgeo.2007.06.005

Reuveny, R. (2007). Climate change-induced migration and violent conflict. *Political Geography, 26*(6), 656-673. doi: 10.1016/j.polgeo.2007.05.001

Reuveny, R., & Moore, W. H. (2009). Does Environmental Degradation Influence Migration? Emigration to Developed Countries in the Late 1980s and 1990s*. *Social Science Quarterly, 90*(3), 461-479. doi: 10.1111/j.1540-6237.2009.00569.x

Ripley, B. D. (1976). The second-order analysis of stationary point processes. *Journal of applied probability*, 255-266.

Ripley, B. D. (1977). Modelling Spatial Patterns. *Journal of the Royal Statistical Society, 39*(2), 172-212.

Root, E. D. (2011). Spatial Point Patterns Retrieved 12/05/2012, 2012, from http://www.colorado.edu/geography/class_homepages/geog_4023_s11/lectures.html

Rowlingson, B., & Diggle, P. (2010). splancs: Spatial and Space-Time Point Pattern Analysis. *R package version*, 2.01-27.

Schabenberger, O., & Gotway, C. A. (2005). *Statistical methods for spatial data analysis*. Boca Raton, FL [etc.]: Chapman & Hall/CRC.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., . . . Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews, 99*(3–4), 125-161. doi: http://dx.doi.org/10.1016/j.earscirev.2010.02.004

Sommers, M. (2003). *Urbanization, War, and Africa's Youth at Risk: Towards Understanding and Addressing Future Challenges.* Paper presented at the Youth explosion in developing world cities: Approaches to reducing poverty and conflict in an urban age, Washington, DC: Woodrow Wilson International Center for Scholars.

Strandow, D., Findley, M., Nielson, D., & Powell, J. (2011). The UCDP and AidData codebook on georeferencing aid : Version 1.1 (pp. 30). Uppsala: Department of Peace and Conflict Research, Uppsala University.

Sundberg, R., Lindgren, M., & Padskocimaite, A. (2010). UCDP GED Codebook Version 1.0-2011. *Department of Peace and Conflict Research, Uppsala University*.

Tao, F., Yokozawa, M., Hayashi, Y., & Lin, E. (2003). Future climate change, the agricultural water cycle, and agricultural production in China. *Agriculture, Ecosystems & Environment, 95*(1), 203-215. doi: http://dx.doi.org/10.1016/S0167-8809(02)00093-2

Theisen, O. M. (2012). Climate Clashes? Weather Variability, Land Pressure, and Organized Violence in Kenya 1989-2004. *Journal of Peace Research, 49*(1).

Theisen, O. M., Holtermann, H., & Buhaug, H. (2010). *Drought, political exclusion, and civil war.* Paper presented at the International Studies Association Annual Convention, New Orleans, LA.

Tobler, W. (1979). Cellular geography. *Philosophy in geography, 9*, 379-386.

Waagepetersen, R. (2008). Estimating functions for inhomogeneous spatial point processes with incomplete covariate data. *Biometrika, 95*(2), 351-363. doi: 10.1093/biomet/asn020

Wallensteen, P., & Sollenberg, M. (2001). Armed Conflict, 1989-2000. *Journal of Peace Research, 38*(5), 629-644. doi: 10.1177/0022343301038005008

Widom, B., & Rowlinson, J. S. (1970). New Model for the Study of Liquid--Vapor Phase Transitions. *The Journal of Chemical Physics, 52*(4), 1670-1684.

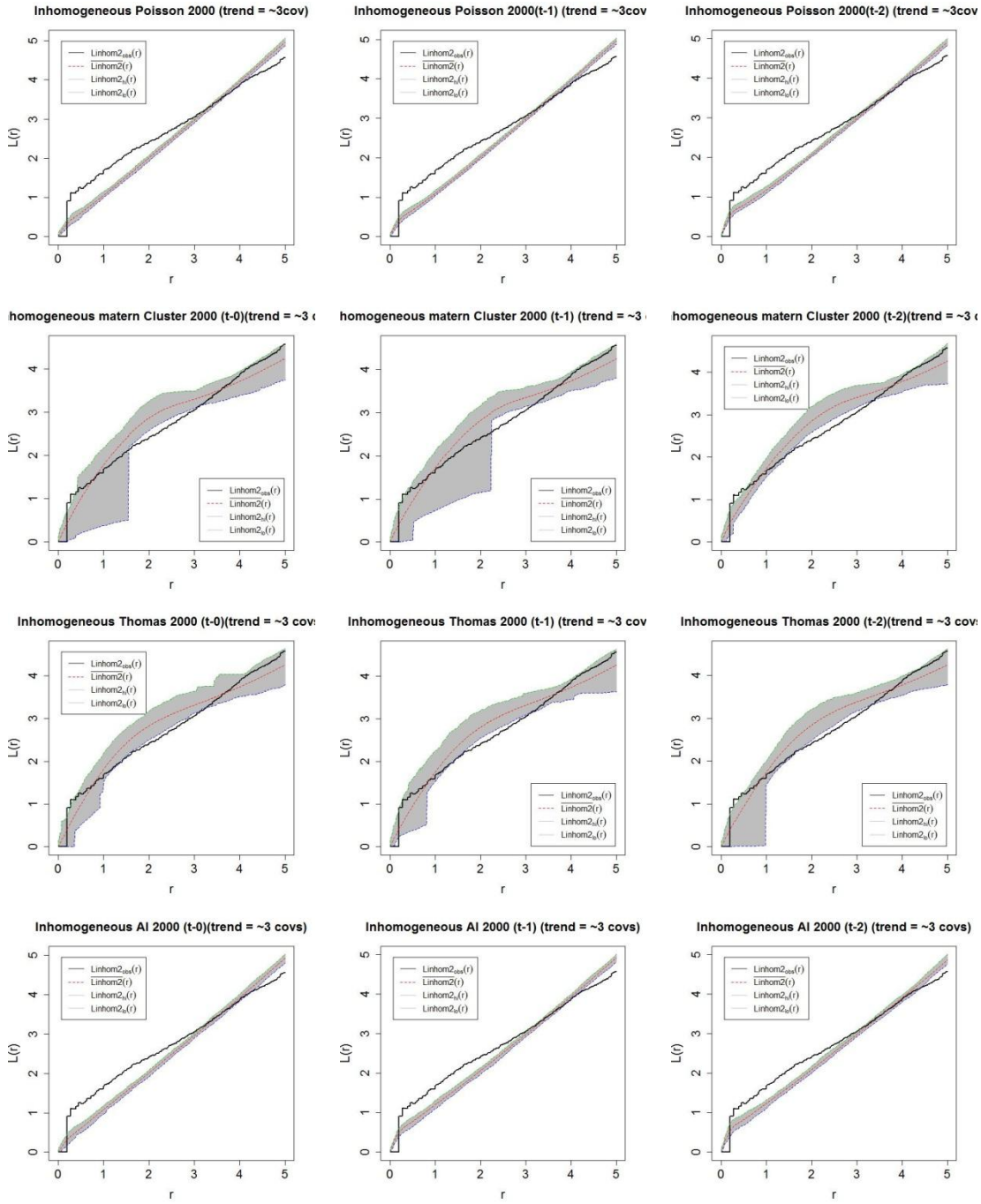
Wiegand, T., & A. Moloney, K. (2004). Rings, circles, and null-models for point pattern analysis in ecology. *Oikos, 104*(2), 209-229. doi: 10.1111/j.0030-1299.2004.12497.x

Xu, Q., Liu, S., Wan, X., Jiang, C., Song, X., & Wang, J. (2012). Effects of rainfall on soil moisture and water movement in a subalpine dark coniferous forest in southwestern China. *Hydrological Processes, 26*(25), 3800-3809. doi: 10.1002/hyp.8400

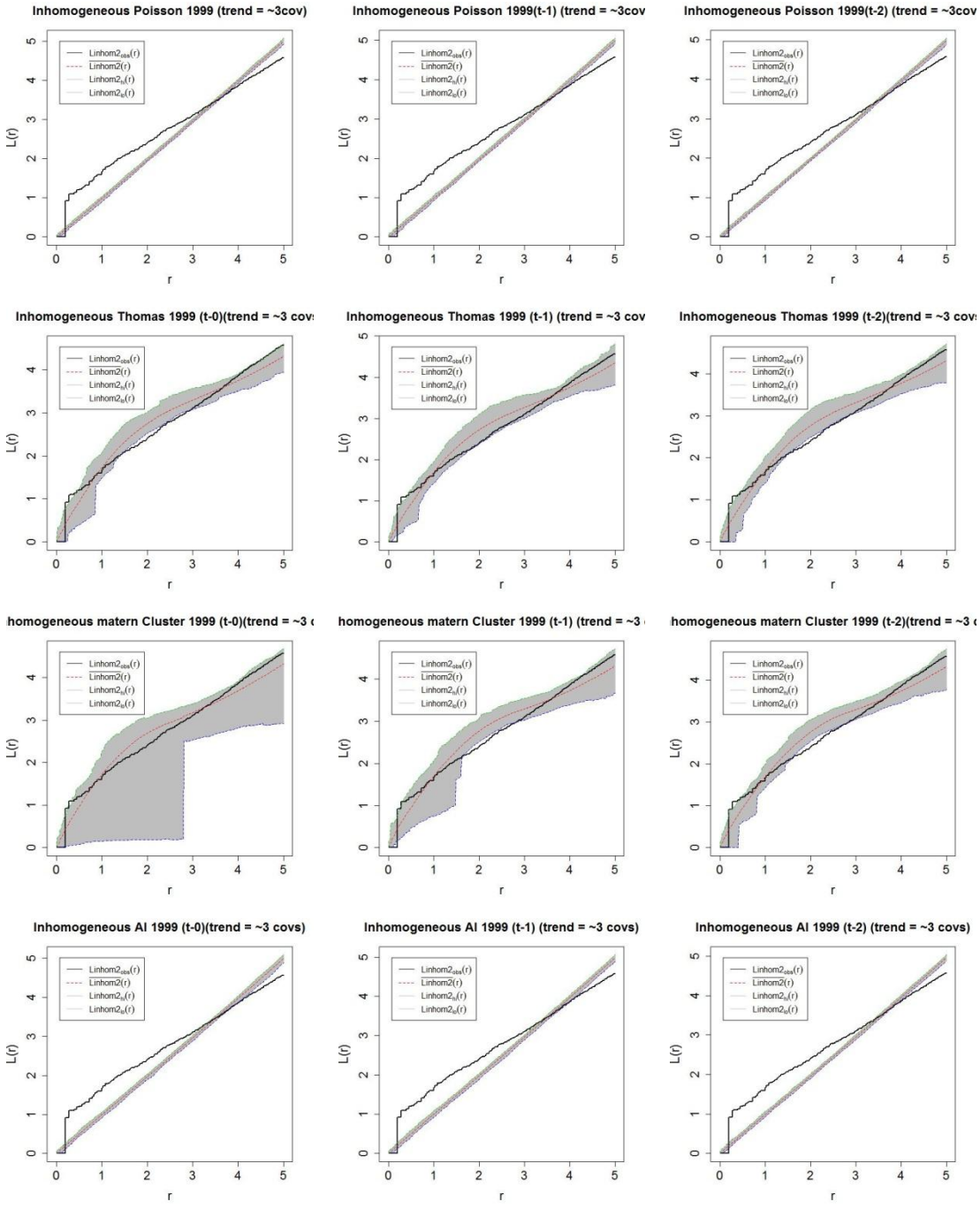
Zammit-Mangion, A., Dewar, M., Kadirkamanathan, V., & Sanguinetti, G. (2012). Point process modelling of the Afghan War Diary. *Proceedings of the National Academy of Sciences, 109*(31), 12414-12419.

7 ANNEX 1(yearly K-function plot)

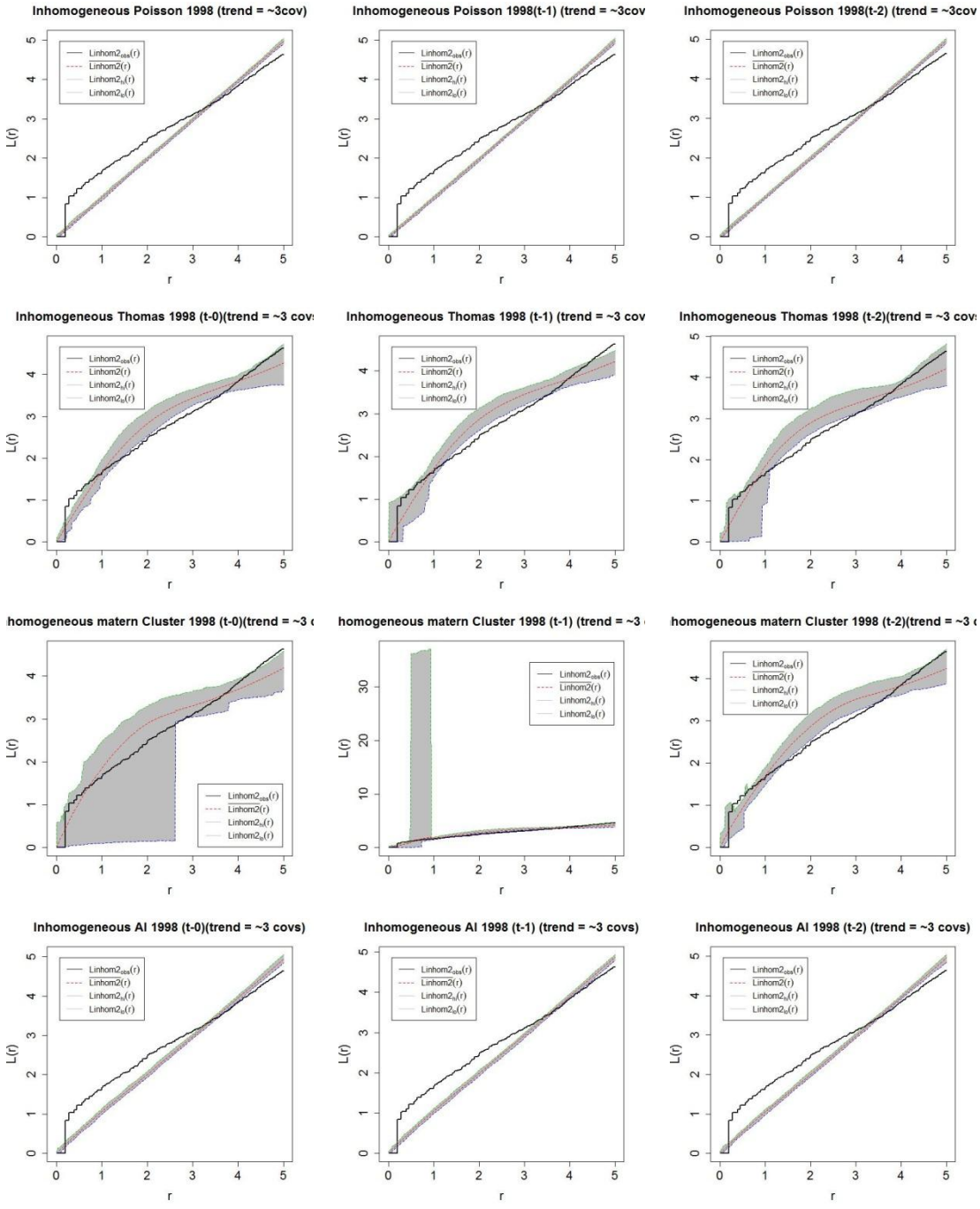
Year 2000



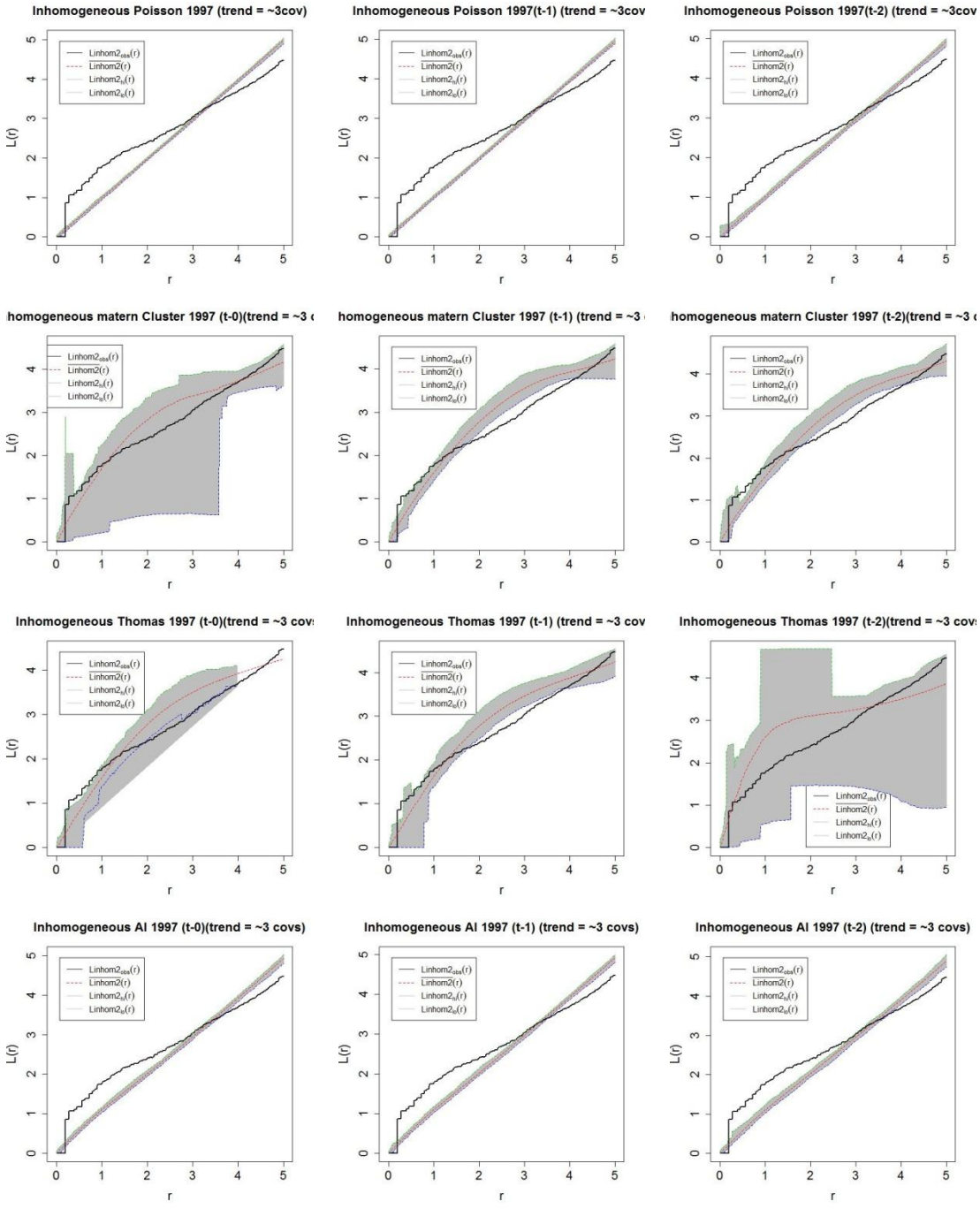
Year 1999


Inhomogeneous Poisson 1999 (trend = ~3cov)
Inhomogeneous Poisson 1999(t-1) (trend = ~3cov
Inhomogeneous Poisson 1999(t-2) (trend = ~3cov
Inhomogeneous Thomas 1999 (t-0)(trend = ~3 cov
Inhomogeneous Thomas 1999 (t-1) (trend = ~3 cov
Inhomogeneous Thomas 1999 (t-2)(trend = ~3 cov
homogeneous matern Cluster 1999 (t-0)(trend = ~3
homogeneous matern Cluster 1999 (t-1) (trend = ~3
homogeneous matern Cluster 1999 (t-2)(trend = ~3
Inhomogeneous AI 1999 (t-0)(trend = ~3 covs)
Inhomogeneous AI 1999 (t-1) (trend = ~3 covs)
Inhomogeneous AI 1999 (t-2) (trend = ~3 covs)
L(r)
r

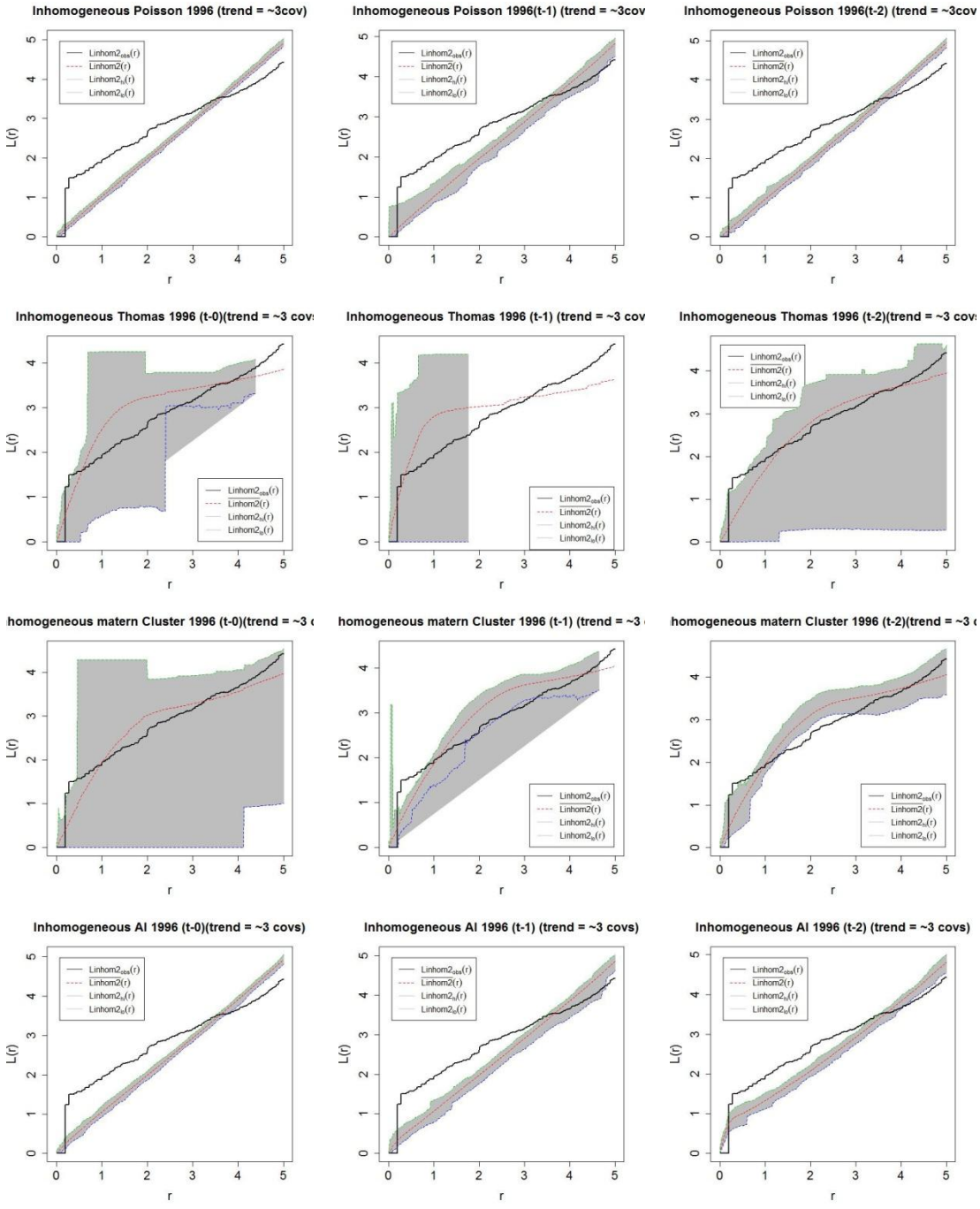
Year 1998


Inhomogeneous Poisson 1998 (trend = ~3cov)
Inhomogeneous Poisson 1998(t-1) (trend = ~3cov
Inhomogeneous Poisson 1998(t-2) (trend = ~3cov
Inhomogeneous Thomas 1998 (t-0)(trend = ~3 cov
Inhomogeneous Thomas 1998 (t-1) (trend = ~3 cov
Inhomogeneous Thomas 1998 (t-2)(trend = ~3 cov
homogeneous matern Cluster 1998 (t-0)(trend = ~3
homogeneous matern Cluster 1998 (t-1) (trend = ~3
homogeneous matern Cluster 1998 (t-2)(trend = ~3
Inhomogeneous AI 1998 (t-0)(trend = ~3 covs)
Inhomogeneous AI 1998 (t-1) (trend = ~3 covs)
Inhomogeneous AI 1998 (t-2) (trend = ~3 covs)
L(r)
r

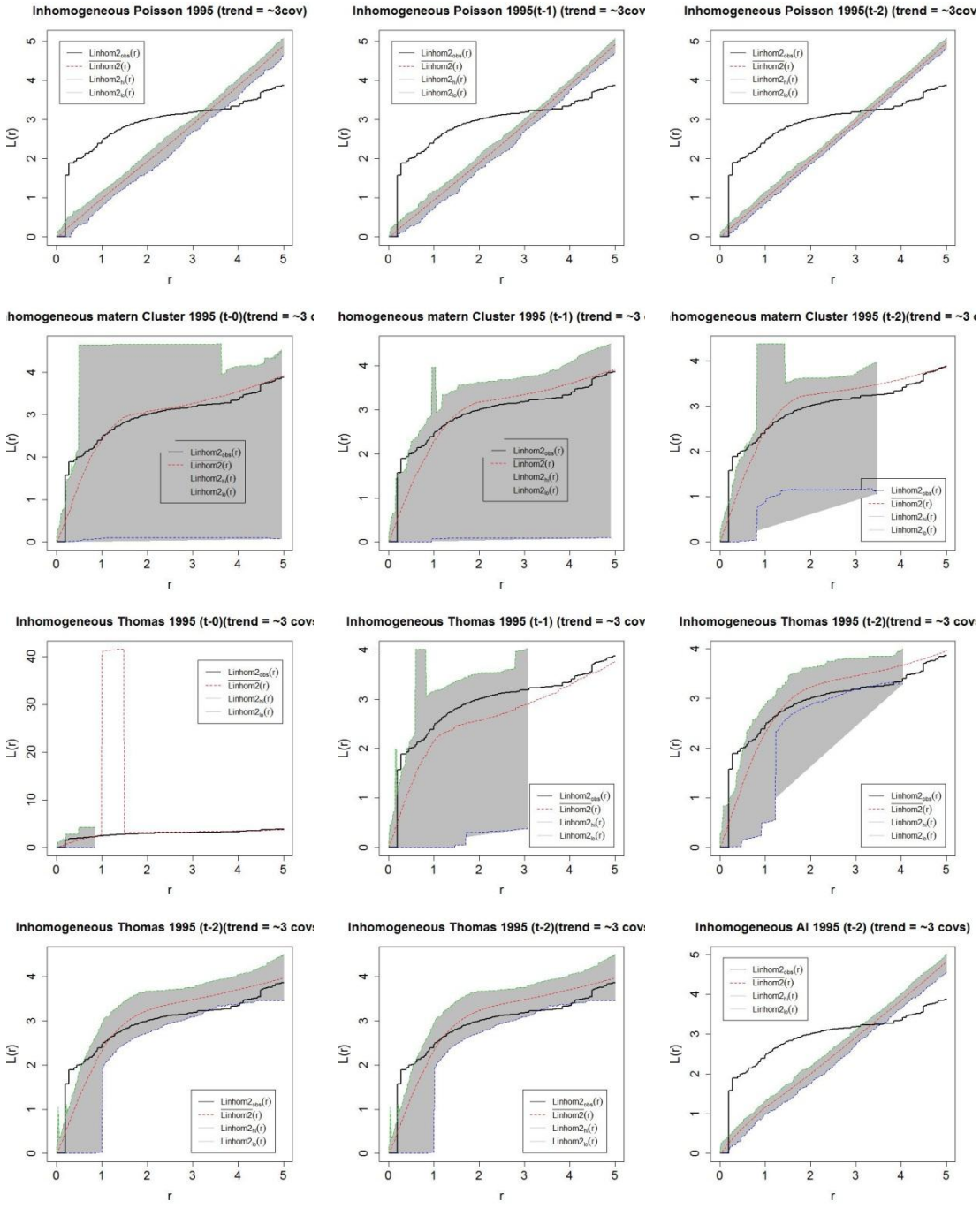
Year 1997


Inhomogeneous Poisson 1997 (trend = ~3cov)
Inhomogeneous Poisson 1997(t-1) (trend = ~3cov
Inhomogeneous Poisson 1997(t-2) (trend = ~3cov
homogeneous matern Cluster 1997 (t-0)(trend = ~3
homogeneous matern Cluster 1997 (t-1) (trend = ~3
homogeneous matern Cluster 1997 (t-2)(trend = ~3
Inhomogeneous Thomas 1997 (t-0)(trend = ~3 cov
Inhomogeneous Thomas 1997 (t-1) (trend = ~3 cov
Inhomogeneous Thomas 1997 (t-2)(trend = ~3 cov
Inhomogeneous Al 1997 (t-0)(trend = ~3 covs)
Inhomogeneous Al 1997 (t-1) (trend = ~3 covs)
Inhomogeneous Al 1997 (t-2) (trend = ~3 covs)
L(r)
r

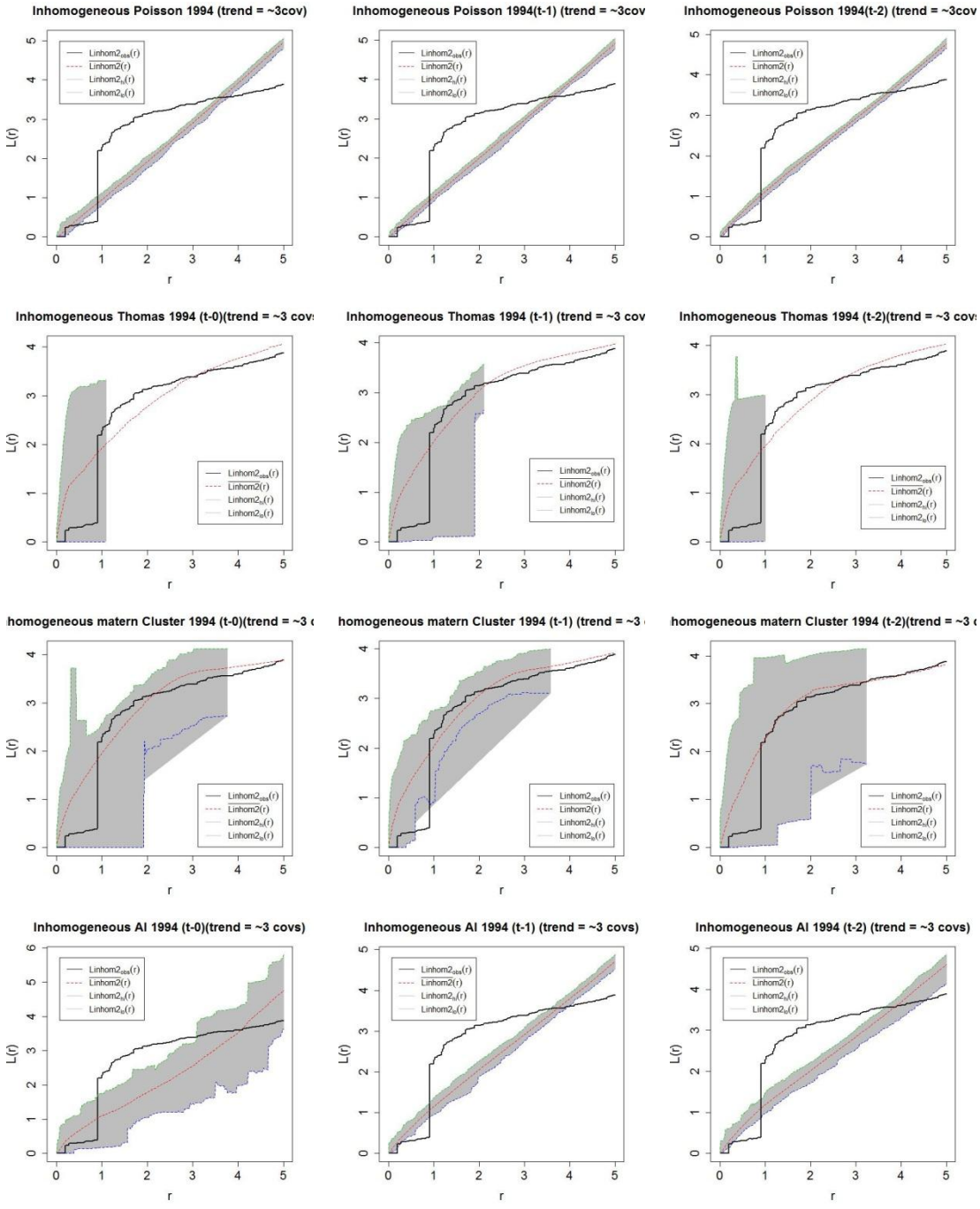
Year 1996


Inhomogeneous Poisson 1996 (trend = ~3cov)
Inhomogeneous Poisson 1996(t-1) (trend = ~3cov
Inhomogeneous Poisson 1996(t-2) (trend = ~3cov
Inhomogeneous Thomas 1996 (t-0)(trend = ~3 cov
Inhomogeneous Thomas 1996 (t-1) (trend = ~3 cov
Inhomogeneous Thomas 1996 (t-2)(trend = ~3 cov
homogeneous matern Cluster 1996 (t-0)(trend = ~3
homogeneous matern Cluster 1996 (t-1) (trend = ~3
homogeneous matern Cluster 1996 (t-2)(trend = ~3
Inhomogeneous Al 1996 (t-0)(trend = ~3 covs)
Inhomogeneous Al 1996 (t-1) (trend = ~3 covs)
Inhomogeneous Al 1996 (t-2) (trend = ~3 covs)
L(r)
r

Year 1995


Inhomogeneous Poisson 1995 (trend = ~3cov)
Inhomogeneous Poisson 1995(t-1) (trend = ~3cov
Inhomogeneous Poisson 1995(t-2) (trend = ~3cov
homogeneous matern Cluster 1995 (t-0)(trend = ~3
homogeneous matern Cluster 1995 (t-1) (trend = ~3
homogeneous matern Cluster 1995 (t-2)(trend = ~3
Inhomogeneous Thomas 1995 (t-0)(trend = ~3 cov
Inhomogeneous Thomas 1995 (t-1) (trend = ~3 cov
Inhomogeneous Thomas 1995 (t-2)(trend = ~3 cov
Inhomogeneous Thomas 1995 (t-2)(trend = ~3 cov
Inhomogeneous Thomas 1995 (t-2)(trend = ~3 cov
Inhomogeneous AI 1995 (t-2) (trend = ~3 covs)
L(r)
r

Year 1994


Inhomogeneous Poisson 1994 (trend = ~3cov)
Inhomogeneous Poisson 1994(t-1) (trend = ~3cov
Inhomogeneous Poisson 1994(t-2) (trend = ~3cov
Inhomogeneous Thomas 1994 (t-0)(trend = ~3 cov
Inhomogeneous Thomas 1994 (t-1) (trend = ~3 cov
Inhomogeneous Thomas 1994 (t-2)(trend = ~3 cov
homogeneous matern Cluster 1994 (t-0)(trend = ~3
homogeneous matern Cluster 1994 (t-1) (trend = ~3
homogeneous matern Cluster 1994 (t-2)(trend = ~3
Inhomogeneous Al 1994 (t-0)(trend = ~3 covs)
Inhomogeneous Al 1994 (t-1) (trend = ~3 covs)
Inhomogeneous Al 1994 (t-2) (trend = ~3 covs)
L(r)
r

8 ANNEX 2 (R-code)

```
#.....................................................................
# Title: Riazuddin_MSc_Thesis.R
# Author: Kawsar, Riazuddin
# Date: February 2013
# Topic: Spatio-temporal analyses of the relationship between armed
conflict and climate change in Eastern Africa
# Dependent Variable: Armed Conflict Events
# Covariates: SWI, WASWI, SPI, Population
# Note: This section only contains the modelling codes not the code
for data processing and exploratory analyses.
#.....................................................................

####################################
## Part 1: Point Process Analyses ##
####################################

library(splancs)
library(spatstat)
library(stpp)

# Read study area polygon into R
region=read.csv("region.csv",header=T)

#Read dependent variable data into R
events=read.csv("events.csv",header=T)

# Read covariates into R and convert to class "im" for spatstat
#population
popvector<-as.real(covariates_05x05d$pop)
popcov<-matrix(popvector, nrow = 200, ncol = 100,byrow=T)
popcovdef <-im(popcov, xcol =
seq(min(covariates_05x05d$x),max(covariates_05x05d$x),length = 100),
yrow = seq(min(covariates_05x05d$y),max(covariates_05x05d$y),length
=200))
#SPI
spivector<-as.real(covariates_05x05d$spi)
spicov<-matrix(spivector, nrow = 200, ncol = 100,byrow=T)
spicovdef <-im(spicov, xcol =
seq(min(covariates_05x05d$x),max(covariates_05x05d$x),length = 100),
yrow = seq(min(covariates_05x05d$y),max(covariates_05x05d$y),length
=200))
#WASWI
swivector<-as.real(covariates_05x05d$sswi)
swicov<-matrix(swivector, nrow = 200, ncol = 100,byrow=T)
swicovdef <-im(swicov, xcol =
seq(min(covariates_05x05d$x),max(covariates_05x05d$x),length = 100),
yrow = seq(min(covariates_05x05d$y),max(covariates_05x05d$y),length
=200))
# Contour:
polyowinRegion=owin(poly=list(x=region$x,y= region$y),
unitname=c("Unit (one Unit = 110 km app)"))

# Point patterns of the data events
Datappp=ppp(x= events$x,y=events$y,window=polyowinRegion)
```

```
########################################
## Spatial Point Process Model Fitting ##
########################################

# Inhomogeneous poisson models
source("kinhom2.txt")
covRegion=list(popcovdef =
popcovdef,swicovdef=swicovdef,spicovdef=spicovdef)
Model_Poisson_3cov=ppm(QRegion,~popcovdef+swicovdef+spicovdef,Poisso
n(),covariates= covRegion)
Model_Poisson_3cov_envelopes=envelope(Model_Poisson_3cov,Linhom,nsim
=99, correction="border",global=F)
plot(Model_Poisson_3cov_envelopes,ylab="L(r)",xlab="r",cex.lab=1.6,c
ex.axis=1.5,cex.main=1.5,main="Inhomogeneous Poisson (trend =
~3cov)")

# Inhomogeneous Thomas Cluster Process
Model_Thomas_3cov=kppm(Datappp,
~popcovdef+swicovdef+spicovdef,"Thomas",covariates = covRegion)
Model_Thomas_3cov_envelopes=envelope(Model_Thomas_3cov,Linhom,nsim=9
9, correction="border",global=F)
plot(Model_Thomas_3cov_envelopes,ylab="L(r)",xlab="r",cex.lab=1.6,ce
x.axis=1.5,cex.main=1.5,main="Inhomogeneous Thomas (trend = ~3
covs)")

# Inhomogeneous Matern Cluster Process
Model_MatClust_3cov=kppm(Datappp,
~popcovdef+swicovdef+spicovdef,"MatClust",covariates = covRegion)
Model_MatClust_3cov_envelopes=envelope(Model_MatClust_3cov,Linhom2,n
sim=99, correction="border")
plot(Model_MatClust_3cov_envelopes[1:200],ylab="L(r)",xlab="r",cex.l
ab=1.6,cex.axis=1.5,cex.main=1.5,main="Inhomogeneous MatClust (trend
= ~3 covs)")

#Estimating irregular parameters for interaction models
s <- data.frame(r=seq(0.5,2, by=0.1))
ratioAIa <-
profilepl(s,AreaInter,QRegion,~popcovdef+swicovdef+spicovdef,covaria
tes = covRegion)
if(interactive()) plot(ratioAIa,ylab="logPL",xlab="Distance (1 Unit
= 110 Km)",cex.lab=1.6,cex.axis=1.5,cex.main=1.5,main="Profile
Maximum Pseudolikelihood")
#best fit was found for 1 unit distance

# Inhomogeneous Area-Interaction Process
Model_AI_3cov=ppm(QRegion,~popcovdef+swicovdef+spicovdef,AreaInter(r
=1),covariates= covRegion)
Model_AI_3cov_envelopes_a=envelope(Model_AI_3cov,Linhom2,nsim=99,
correction="border",global=F)
plot(Model_AI_3cov_envelopes_a,ylab="Linhom2(Distance)",xlab="Distan
ce (1 unit = 100
km)",cex.lab=1,cex.axis=1,cex.main=1,main="Inhomogeneous Area -
Interaction Process (trend = ~3 Covariates)")

# end of spatial point pattern analysis
####################################################
```

```
################################################
## Spatio-temporal Point Process Model fitting ##
################################################

#Events Data (space-time format)
events1b=read.table("events1b.txt",header=T)
eventsc=cbind(events1b$x,events1b$y,events1b$days)
EVENTSC=eventsc
EVENTS.in=EVENTSC[inpip(EVENTSC[,1:2],region),1:3]
EVENTSC[723 ,1]=20.02
EVENTSC=as.3dpoints(EVENTS.in[, 1] , EVENTS.in[, 2] , EVENTS.in[,3])
region=region

###STIKhat Estimation for EVENTS with intensity coming from kernel
# estimation of the temporal intensity
Mt <- density(EVENTSC[ ,3], n = 1000)
mut <- Mt$y[findInterval(EVENTSC[ ,3], Mt$x)] * dim(EVENTSC)[1]
h <- mse2d(as.points(EVENTSC[,1:2]), region, nsmse = 50, range = 4)
h <- h$h[which.min(h$mse)]
# estimation of the spatial intensity
Ms <- kernel2d(as.points(EVENTSC[ ,1:2]), region, h = 1, nx = 140,
ny =140)
atx <- findInterval(x = EVENTSC[ ,1], vec = Ms$x)
aty <- findInterval(x = EVENTSC[ ,2], vec = Ms$y)
# check if any atx or aty are zero; if 0 replace it with 1.
atx[which(atx==0)]=1
aty[which(aty==0)]=1
mhat <- NULL
for(i in 1:length(atx)) mhat <- c(mhat, Ms$z[atx[i],aty[i]])
#normalizing the intensity
aa=(areapl(region)/(length(Ms$x)*length(Ms$x)))*sum(Ms$z)
bb=((range(EVENTS.in[,3])[2]-
range(EVENTS.in[,3])[1])/length(Mt$y))*sum(Mt$y*dim(EVENTSC)[1])
Mt.new=Mt$y*dim(EVENTSC)[1]*(dim(EVENTSC)[1]/bb)
Ms.new=Ms$z*(dim(EVENTSC)[1]/aa)

Lst=array(0, dim = c(140, 140, 1000))
for(k in 1:1000) Lst[,,k] <- Ms.new*Mt.new[k]/dim(EVENTSC)[1]
cc=((800*3652)/(140*140*1000))*sum(Lst)
mut.new=Mt.new[findInterval(EVENTSC[ ,3], Mt$x)]
mhat.new=NULL

for(i in 1:length(atx)) mhat.new=c(mhat.new, Ms.new[atx[i],aty[i]])
l.in=mhat.new * mut.new / dim(EVENTSC)[1]

u=seq(0.001, 4, len = 20)
v=seq(1,20,len=20)
stik.without.cov=STIKhat(xyt = EVENTSC, s.region = region, t.region
= c(1, 3653),lambda = l.in, dist = u, times = v, infectious = TRUE)

plotK(stik.without.cov)

### rpp
Lst=array(0, dim = c(140, 140, 1000))
for(k in 1:1000) Lst[,,k] <- Ms.new*Mt.new[k]/dim(EVENTSC)[1]
ipp2Events=rpp(lambda=Lst, s.region = region, t.region =c(1,3653),
discrete.time = TRUE)
```

```
###STIKhat Estimation for EVENTS with spatial intensity coming from
covariates
#spatial intensity comming fromt he covaraites
polyowinRegion=owin(poly=list(x=region[,1],y= region[,2]))
Datappp=ppp(x= EVENTSC[,1],y=EVENTSC[,2],window=polyowinRegion)
QRegion=quadscheme(data= Datappp, dummy=list(x=covariates_05x05d$x,
y=covariates_05x05d$y))
covRegion=list(popcovdef =
popcovdef,swicovdef=swicovdef,spicovdef=spicovdef)
Model_Poisson_3cov=ppm(QRegion,~popcovdef+swicovdef+spicovdef,Poisso
n(),covariates= covRegion)
predict.Model_Poisson_3cov=predict(Model_Poisson_3cov,ngrid=140,type
="trend")

Mt=density(EVENTSC[ ,3], n = 1000)
Ms.cov=predict.Model_Poisson_3cov
atx <- findInterval(x = EVENTSC[ ,1], vec =
predict.Model_Poisson_3cov$xcol)
aty <- findInterval(x = EVENTSC[ ,2], vec =
predict.Model_Poisson_3cov$yrow)
atx[which(atx==0)]=1
aty[which(aty==0)]=1
aa=(areapl(region)/(length(Ms.cov$xcol)*length(Ms.cov$xcol)))*sum(Ms
.cov$v)
bb=((range(EVENTS.in[,3])[2]-
range(EVENTS.in[,3])[1])/length(Mt$y))*sum(Mt$y*dim(EVENTSC)[1])
Mt.new=Mt$y*dim(EVENTSC)[1]*(dim(EVENTSC)[1]/bb)
Ms.cov.new=Ms.cov$v*(dim(EVENTSC)[1]/aa)

Lst.cov=array(0, dim = c(140, 140, 1000))
for(k in 1:1000) Lst.cov[,,k] <-
Ms.cov.new*Mt.new[k]/dim(EVENTSC)[1]
cc=((800*3652)/(140*140*1000))*sum(Lst.cov)
mut.new=Mt.new[findInterval(EVENTSC[ ,3], Mt$x)]
mhat.new=NULL

for(i in 1:length(atx)) mhat.new=c(mhat.new,
Ms.cov.new[atx[i],aty[i]])
l.covs.in=mhat.new * mut.new / dim(EVENTSC)[1]

u=seq(0.001, 4, len = 20)
v=seq(1,20,len=20)
stik.with.cov=STIKhat(xyt = EVENTSC, s.region = region, t.region =
c(1, 3653),lambda = l.covs.in, dist = u, times = v, infectious =
TRUE)
plotK(stik.with.cov)

## rpp ####
Lst=array(0, dim = c(140, 140, 1000))
for(k in 1:1000) Lst.cov[,,k]=Ms.cov.new*Mt.new[k]/dim(EVENTSC)[1]
ipp3Events=rpp(lambda=Lst.cov, s.region = region, t.region
=c(1,3653), discrete.time = TRUE)

# P-value Estimation for EVENTS without covs (with Kernel)
nsim=100
res.new=array(0,c(20,20,nsim))
for(k in 1:nsim){
      cat("simul=",k,"\n")
      simulated=NULL
```

```
      stik.simulated=NULL
      Mt=NULL;mut=NULL;Ms=NULL;atx=NULL;aty=NULL
      simulated=rpp(lambda=Lst, s.region = region, t.region
=c(1,3653), discrete.time = TRUE)
      Mt <- density(simulated$xyt[ ,3], n = 1000)
      mut <- Mt$y[findInterval(simulated$xyt[ ,3], Mt$x)] *
dim(simulated$xyt)[1]
      Ms <- kernel2d(as.points(simulated$xyt[ ,1:2]), region, h = 1,
nx = 140, ny =140)
      atx <- findInterval(x = simulated$xyt[ ,1], vec = Ms$x)
      aty <- findInterval(x = simulated$xyt[ ,2], vec = Ms$y)
      atx[which(atx==0)]=1
      aty[which(aty==0)]=1
      aa=(areapl(region)/(length(Ms$x)*length(Ms$x)))*sum(Ms$z)
      bb=((range(simulated$xyt[,3])[2]-
range(simulated$xyt[,3])[1])/length(Mt$y))*sum(Mt$y*dim(simulated$xy
t)[1])
      Mt.new=Mt$y*dim(simulated$xyt)[1]*(dim(simulated$xyt)[1]/bb)
      Ms.new=Ms$z*(dim(simulated$xyt)[1]/aa)
      mut.new=Mt.new[findInterval(simulated$xyt[ ,3], Mt$x)]
      mhat.new=NULL
      for(i in 1:length(atx)) mhat.new=c(mhat.new,
Ms.new[atx[i],aty[i]])
      l.in=mhat.new*mut.new/dim(simulated$xyt)[1]
      stik.simulated=STIKhat(xyt = simulated$xyt, s.region = region,
t.region = c(1, 3653),lambda = l.in, dist = u, times = v, infectious
= TRUE)
      tt=stik.simulated$Khat-stik.simulated$Ktheo
      res.new[,,k]=tt
}

# P-value checking and plotting for EVENTS without covs (with
Kernel)
lu=length(u)
lv=length(v)
empirical.covs=sqrt(stik.without.cov$Khat-stik.without.cov$Ktheo)
       pvalue.covs=matrix(0,lu,lv)
        for(i in 1:lu){
          for(j in 1:lv){
            aaa=NULL
            aaa=c(empirical.covs[i,j],res.new[i,j,])
            pvalue.covs[i,j]=1-(rank(aaa)[1]/length(aaa))
            }
        }
    pvalues.covs.TF=pvalue.covs<0.05
    print(pvalues.covs.TF)

image(u,v,pvalue.covs)
contour(u,v,pvalue.covs,add=T)

# P-value EstimationEVENTS with covariates
nsim=100
res.new.cov=array(0,c(20,20,nsim))
for(k in 1:nsim){
      cat("simul=",k,"\n")
      simulated=NULL
      stik.simulated=NULL
      Mt=NULL;mut=NULL;Ms=NULL;atx=NULL;aty=NULL
```

```
      simulated=rpp(lambda=Lst.cov, s.region = region, t.region
=c(1,3653), discrete.time = TRUE)

      polyowinRegion=owin(poly=list(x=region[,1],y= region[,2]))
      Datappp=ppp(x=
simulated$xyt[,1],y=simulated$xyt[,2],window=polyowinRegion)
      QRegion=quadscheme(data= Datappp,
dummy=list(x=covariates_05x05d$x, y=covariates_05x05d$y))
      covRegion=list(popcovdef =
popcovdef,swicovdef=swicovdef,spicovdef=spicovdef)
      Model_Poisson_3cov=ppm(QRegion,~popcovdef+swicovdef+spicovdef,
Poisson(),covariates= covRegion)
      predict.Model_Poisson_3cov=predict(Model_Poisson_3cov,ngrid=14
0,type="trend")

      Mt=density(simulated$xyt[ ,3], n = 1000)
      Ms.cov=predict.Model_Poisson_3cov
      atx <- findInterval(x = simulated$xyt[ ,1], vec =
predict.Model_Poisson_3cov$xcol)
      aty <- findInterval(x = simulated$xyt[ ,2], vec =
predict.Model_Poisson_3cov$yrow)
      atx[which(atx==0)]=1
      aty[which(aty==0)]=1
      aa=(areapl(region)/(length(Ms.cov$xcol)*length(Ms.cov$xcol)))*
sum(Ms.cov$v)
      bb=((range(simulated$xyt[,3])[2]-
range(simulated$xyt[,3])[1])/length(Mt$y))*sum(Mt$y*dim(simulated$xy
t)[1])
      Mt.new=Mt$y*dim(simulated$xyt)[1]*(dim(simulated$xyt)[1]/bb)
      Ms.cov.new=Ms.cov$v*(dim(simulated$xyt)[1]/aa)
      mut.new=Mt.new[findInterval(simulated$xyt[ ,3], Mt$x)]
      mhat.new=NULL
      for(i in 1:length(atx)) mhat.new=c(mhat.new,
Ms.cov.new[atx[i],aty[i]])
      l.covs.in=mhat.new*mut.new/dim(simulated$xyt)[1]
      stik.simulated=STIKhat(xyt = simulated$xyt, s.region = region,
t.region = c(1, 3653),lambda = l.covs.in, dist = u, times = v,
infectious = TRUE)
      res.new.cov[,,k]=stik.simulated$Khat-stik.simulated$Ktheo
}
lu=length(u)
lv=length(v)
empirical.covs=stik.with.cov$Khat-stik.with.cov$Ktheo
       pvalue.covs=matrix(0,lu,lv)
        for(i in 1:lu){
          for(j in 1:lv){
            aaa=NULL
            aaa=c(empirical.covs[i,j],res.new.cov[i,j,])
            pvalue.covs[i,j]=1-(rank(aaa)[1]/length(aaa))
            }
        }
    pvalues.covs.TF=pvalue.covs<0.005
    print(pvalues.covs.TF)

image(u,v,pvalue.covs)
contour(u,v,pvalue.covs,add=T)

# end of spatio temporal point pattern analysis
```

```
#########################
##### Lattice Approch ###
#########################

library(maptools)
library(zoo)
library(spacetime)
library(spdep)

# read the year wise shapefiles
 shp1=readShapePoly("All_Data_50x50grid_1991.shp")
 shp2=readShapePoly("All_Data_50x50grid_1992.shp")
 shp3=readShapePoly("All_Data_50x50grid_1993.shp")
 shp4=readShapePoly("All_Data_50x50grid_1994.shp")
 shp5=readShapePoly("All_Data_50x50grid_1995.shp")
 shp6=readShapePoly("All_Data_50x50grid_1996.shp")
 shp7=readShapePoly("All_Data_50x50grid_1997.shp")
 shp8=readShapePoly("All_Data_50x50grid_1998.shp")
 shp9=readShapePoly("All_Data_50x50grid_1999.shp")
 shp10=readShapePoly("All_Data_50x50grid_2000.shp")

### Independent and Dependent Variables: Aggregated data (EVENTS,
SSWI, SWI, SPI, POP)
 all <- readShapePoly("alldata50x50.shp")

### Dependent Variables: number of EVENTS for each cell in each year
 EVENTS1991=(shp1$eventcount)
 EVENTS1992=(shp2$eventcount)
 EVENTS1993=(shp3$eventcount)
 EVENTS1994=(shp4$eventcount)
 EVENTS1995=(shp5$eventcount)
 EVENTS1996=(shp6$eventcount)
 EVENTS1997=(shp7$eventcount)
 EVENTS1998=(shp8$eventcount)
 EVENTS1999=(shp9$eventcount)
 EVENTS2000=(shp10$eventcount)

 EVENTS_T=c(EVENTS1992, EVENTS1993, EVENTS1994, EVENTS1995,
EVENTS1996, EVENTS1997, EVENTS1998, EVENTS1999, EVENTS2000) #time T
 EVENTS_TB=c(EVENTS1991, EVENTS1992, EVENTS1993, EVENTS1994,
EVENTS1995, EVENTS1996, EVENTS1997, EVENTS1998, EVENTS1999) #time T-
1: AR(1)
 EVENTS_TB_1=c(EVENTS1991, EVENTS1991, EVENTS1992, EVENTS1993,
EVENTS1994, EVENTS1995, EVENTS1996, EVENTS1997, EVENTS1998)

 ### Independent Variables: WEIGHTED ANOMALY SOIL WATER INDEX
(WASWI) for each cell in each year
 swi1=readShapePoly("SWI50x501991.shp")
 swi2=readShapePoly("SWI50x501992.shp")
 swi3=readShapePoly("SWI50x501993.shp")
 swi4=readShapePoly("SWI50x501994.shp")
 swi5=readShapePoly("SWI50x501995.shp")
 swi6=readShapePoly("SWI50x501996.shp")
 swi7=readShapePoly("SWI50x501997.shp")
 swi8=readShapePoly("SWI50x501998.shp")
 swi9=readShapePoly("SWI50x501999.shp")
 swi10=readShapePoly("SWI50x502000.shp")

 SWI1991=(swi1$swi)
```

```
 SWI1992=(swi2$swi)
 SWI1993=(swi3$swi)
 SWI1994=(swi4$swi)
 SWI1995=(swi5$swi)
 SWI1996=(swi6$swi)
 SWI1997=(swi7$swi)
 SWI1998=(swi8$swi)
 SWI1999=(swi9$swi)
 SWI2000=(swi10$swi)

 SWI_T=c(SWI1992, SWI1993, SWI1994, SWI1995, SWI1996, SWI1997,
SWI1998, SWI1999, SWI2000) #time T
 SWI_TB=c(SWI1991, SWI1992, SWI1993, SWI1994, SWI1995, SWI1996,
SWI1997, SWI1998, SWI1999) #time T-1: AR(1)
 SWI_TB_1=c(SWI1991, SWI1991, SWI1992, SWI1993, SWI1994, SWI1995,
SWI1996, SWI1997, SWI1998) #time T-2: AR(1)

 SWI_TB_2=c((SWI1991+SWI1992), (SWI1991+SWI1992), (SWI1992+SWI1993),
 (SWI1993+SWI1994), (SWI1994+SWI1995), (SWI1995+SWI1996),
(SWI1996+SWI1997), (SWI1997+SWI1998), (SWI1998+SWI1999))

 SWI_TB_3=c((SWI1991+SWI1991+SWI1991), (SWI1991+SWI1991+SWI1992),
(SWI1991+SWI1992+SWI1993), (SWI1992+SWI1993+SWI1994),
  (SWI1993+SWI1994+SWI1995), (SWI1994+SWI1995+SWI1996),
(SWI1995+SWI1996+SWI1997), (SWI1996+SWI1997+SWI1998),
(SWI1997+SWI1998+SWI1999))


 ### Independent Variables: STANDARIZED SOIL WATER INDEX (SWI) for
each cell in each year
 SSWI1991=(shp1$sswi)
 SSWI1992=(shp2$sswi)
 SSWI1993=(shp3$sswi)
 SSWI1994=(shp4$sswi)
 SSWI1995=(shp5$sswi)
 SSWI1996=(shp6$sswi)
 SSWI1997=(shp7$sswi)
 SSWI1998=(shp8$sswi)
 SSWI1999=(shp9$sswi)
 SSWI2000=(shp10$sswi)

 SSWI_T=c(SSWI1992, SSWI1993, SSWI1994, SSWI1995, SSWI1996,
SSWI1997, SSWI1998, SSWI1999, SSWI2000) #time T
 SSWI_TB=c(SSWI1991, SSWI1992, SSWI1993, SSWI1994, SSWI1995,
SSWI1996, SSWI1997, SSWI1998, SSWI1999) #time T-1: AR(1)

 SSWI_TB_1=c(SSWI1991, SSWI1991, SSWI1992, SSWI1993, SSWI1994,
SSWI1995, SSWI1996, SSWI1997, SSWI1998) #time T-2: AR(1)

 SSWI_TB_2=c((SSWI1991+SSWI1992), (SSWI1991+SSWI1992),
(SSWI1992+SSWI1993),
 (SSWI1993+SSWI1994), (SSWI1994+SSWI1995), (SSWI1995+SSWI1996),
(SSWI1996+SSWI1997), (SSWI1997+SSWI1998), (SSWI1998+SSWI1999))

 SSWI_TB_3=c((SSWI1991+SSWI1991+SSWI1991),
(SSWI1991+SSWI1991+SSWI1992), (SSWI1991+SSWI1992+SSWI1993),
(SSWI1992+SSWI1993+SSWI1994),
```

```
  (SSWI1993+SSWI1994+SSWI1995), (SSWI1994+SSWI1995+SSWI1996),
(SSWI1995+SSWI1996+SSWI1997), (SSWI1996+SSWI1997+SSWI1998),
(SSWI1997+SSWI1998+SSWI1999))


 ### Independent Variables: STANDARIZED PRECIPITATION INDEX (SPI)
for each cell in each year
 SPI1991=(shp1$spi)
 SPI1992=(shp2$spi)
 SPI1993=(shp3$spi)
 SPI1994=(shp4$spi)
 SPI1995=(shp5$spi)
 SPI1996=(shp6$spi)
 SPI1997=(shp7$spi)
 SPI1998=(shp8$spi)
 SPI1999=(shp9$spi)
 SPI2000=(shp10$spi)

 SPI_T=c(SPI1992, SPI1993, SPI1994, SPI1995, SPI1996, SPI1997,
SPI1998, SPI1999, SPI2000) #time T
 SPI_TB=c(SPI1991, SPI1992, SPI1993, SPI1994, SPI1995, SPI1996,
SPI1997, SPI1998, SPI1999) #time T-1: AR(1)

 SPI_TB_1=c(SPI1991, SPI1991, SPI1992, SPI1993, SPI1994, SPI1995,
SPI1996, SPI1997, SPI1998) #time T-2: AR(1)

 SPI_TB_2=c((SPI1991+SPI1992), (SPI1991+SPI1992), (SPI1992+SPI1993),
 (SPI1993+SPI1994), (SPI1994+SPI1995), (SPI1995+SPI1996),
(SPI1996+SPI1997), (SPI1997+SPI1998), (SPI1998+SPI1999))

 SPI_TB_3=c((SPI1991+SPI1991+SPI1991), (SPI1991+SPI1991+SPI1992),
(SPI1991+SPI1992+SPI1993), (SPI1992+SPI1993+SPI1994),
  (SPI1993+SPI1994+SPI1995), (SPI1994+SPI1995+SPI1996),
(SPI1995+SPI1996+SPI1997), (SPI1996+SPI1997+SPI1998),
(SPI1997+SPI1998+SPI1999))

  ### Independent Variables: POPULATION for each cell in each year
 POP1991=(shp1$pop)
 POP1992=(shp2$pop)
 POP1993=(shp3$pop)
 POP1994=(shp4$pop)
 POP1995=(shp5$pop)
 POP1996=(shp6$pop)
 POP1997=(shp7$pop)
 POP1998=(shp8$pop)
 POP1999=(shp9$pop)
 POP2000=(shp10$pop)

 POP_T=c(POP1992, POP1993, POP1994, POP1995, POP1996, POP1997,
POP1998, POP1999, POP2000) #time T
 POP_TB=c(POP1991, POP1992, POP1993, POP1994, POP1995, POP1996,
POP1997, POP1998, POP1999) #time T-1: AR(1)

 POP_TB_1=c(POP1991, POP1991, POP1992, POP1993, POP1994, POP1995,
POP1996, POP1997, POP1998) #time T-2: AR(1)

 POP_TB_2=c((POP1991+POP1992), (POP1991+POP1992), (POP1992+POP1993),
 (POP1993+POP1994), (POP1994+POP1995), (POP1995+POP1996),
(POP1996+POP1997), (POP1997+POP1998), (POP1998+POP1999))
```

```
 POP_TB_3=c((POP1991+POP1991+POP1991), (POP1991+POP1991+POP1992),
(POP1991+POP1992+POP1993), (POP1992+POP1993+POP1994),
  (POP1993+POP1994+POP1995), (POP1994+POP1995+POP1996),
(POP1995+POP1996+POP1997), (POP1996+POP1997+POP1998),
(POP1997+POP1998+POP1999))

 ### Aggregated Modelling:
 ### Neighbourhood List Creation:
 W_cont_el <- poly2nb(all, queen=T)
 W_cont_el_mat <- nb2listw(W_cont_el, style="W", zero.policy=TRUE)
 formula = scale(eventcount) ~
scale(sswi)+scale(SWI)+scale(spi)+scale(pop)

 # Model 1.1 (Ordinary Least Square)
 mod.lm_agg <- lm(formula, data = all)
 summary(mod.lm_agg, Nagelkerke=TRUE)

  # Model 1.2 (Spatial AutoRegressive Model Version 2)
 mod.sar_agg <- spautolm(formula, data = all, listw=W_cont_el_mat,
family = "SAR", method = "Matrix")
 summary(mod.sar1_agg, Nagelkerke=TRUE)

  # Model 1.3 (Spatial Error Model)
 mod.sem_agg <- errorsarlm(formula, data = all, listw=W_cont_el_mat,
zero.policy=T, tol.solve=1e-15)
 summary(mod.sem_agg, Nagelkerke=TRUE)

  # Model 1.4 (Spatial Durdin Model)
 mod.sdm_agg <- lagsarlm(formula, data = all, listw=W_cont_el_mat,
zero.policy=T, type="mixed", tol.solve=1e-12)
 summary(mod.sdm_agg, Nagelkerke=TRUE)

 # Morans I test
res.lm <- mod.lm_agg$residuals
res.sar <- mod.sar_agg$residuals
res.sem <- mod.sem_agg$residuals
res.sdm <- mod.sdm_agg$residuals

MI.lm = moran.test(res.lm, listw=W_cont_el_mat, zero.policy=T)
MI.sar = moran.test(res.sar, listw=W_cont_el_mat, zero.policy=T)
MI.sem = moran.test(res.sem, listw=W_cont_el_mat, zero.policy=T)
MI.sdm = moran.test(res.sdm, listw=W_cont_el_mat, zero.policy=T)

## impact analysis
 trMatc <- trW(W_cont_el_mat, type="mult")
 trMC <- trW(W_cont_el_mat, type="MC")
 impacts( mod.sar_agg, listw=W_cont_el_mat)
 impacts( mod.sar_agg, tr=trMatc)
 impacts( mod.sar_agg, tr=trMC)


 # Space time Neighbourhood list creation Function (Author: Edzer
Pebesma)
 nbMult = function(nb, st, addT = TRUE, addST = FALSE) {
      stopifnot(is(st, "STF"))
      n = dim(st)[2]
      if (n <= 1)
            return(nb)
```

```
      L = length(nb)
      ret = list()
      FN = function(x,i,j,L) {
            ret = as.integer(x + i * L) # spatial-only, for time i+1
            if (addT) {
                  if (addST)
                        now = c(ret, j + i * L)
                  else
                        now = j + i * L
                  if (i > 0)
                        ret = c(ret, now - L) # time-previous: j-iL
                  if (i < (n-1))
                        ret = c(ret, now + L) # time-next: j+iL
            }
            sort(ret)
      }
      for (i in 0:(n-1)) {
            app = lapply(1:L, function(j) FN(nb[[j]], i, j, L))
            ret = append(ret, app)
      }
      attributes(ret) = attributes(nb)
      attr(ret, "region.id") = as.character(1:length(ret))
      ret
}


# Time Series preparation from layers of Spatial object
 yrs1 = 1992:2000
 time.xts = as.POSIXct(paste(yrs1, "-01-01", sep=""), tz = "GMT")

 def1 = STFDF(shp1, time.xts, data.frame( EVENTS_T, EVENTS_TB,
                  SSWI_T, SWI_T, SWI_TB,SWI_TB_1,SWI_TB_2,SWI_TB_3,
SSWI_TB,SSWI_TB_1,SSWI_TB_2,SSWI_TB_3 ,
                  SPI_T, SPI_TB,SPI_TB_1,SPI_TB_2,SPI_TB_3,POP_T,
POP_TB, POP_TB_1,POP_TB_2,POP_TB_3))

 formula = scale(EVENTS_T) ~ scale(EVENTS_TB) +
                  scale(SSWI_T) + scale(SSWI_TB) +
scale(SSWI_TB_1)+scale(SSWI_TB_2)+scale(SSWI_TB_3)+
                  scale(SWI_T) + scale(SWI_TB) + scale(SWI_TB_1) +
scale(SWI_TB_2) + scale(SWI_TB_3) +
                  scale(SPI_T) + scale(SPI_TB) + scale(SPI_TB_1) +
scale(SPI_TB_2) + scale(SPI_TB_3) +
                  scale(POP_T) + scale(POP_TB) + scale(POP_TB_1) +
scale(POP_TB_2) + scale(POP_TB_3)

 # temporal neighbourhood Preparation
 nlgal= poly2nb(shp1, queen=T)
 colw=nb2listw(nlgal)
 n=length(nlgal)
 tst = nbMult(nlgal, def1, addST = TRUE)
 colw = nb2listw(tst)

 # model (test) OLS with temporal neighbours where where
nbMult(nlgal, def_all, addT = TRUE)
 mod.lm <- lm(formula, as.data.frame(def))
 summary(mod.lm)
```

```
 # model 2 SAR with temporal neighbours where where nbMult(nlgal,
def_all, addT = TRUE)
 mod.sar <- spautolm(formula1, as.data.frame(def1), colw, family =
"SAR", method = "Matrix")
 summary(mod.sar, Nagelkerke=TRUE)

 # Spatiao-temporal neighbourhood Preparation
 nlgal= poly2nb(shp1, queen=T)
 colw=nb2listw(nlgal)
 n=length(nlgal)
 tst = nbMult(nlgal, def1, addT = TRUE, addST=TRUE)
 colw = nb2listw(tst)

 # model 3 SAR with spatio-temporal neighbours where where
nbMult(nlgal, def_all, addT = TRUE, addST=TRUE)
 mod.sar <- spautolm(formula1, as.data.frame(def1), colw, family =
"SAR", method = "Matrix")
 summary(mod.sar12, Nagelkerke=TRUE)

 # end of Lattice Approch
##################################################

# the end of analysis
#####################################################################
##############
```