Spatio-Temporal Analysis of Armed Conflict in Africa - Case Study: Political Conflict Events in Libya

Nicholas Johnston, MSc Data Science, City University London, April 2016

Abstract
This paper uses data provided by the Armed Conflict Location and Event Detection (ACLED) project to visualise instances of political violence across the African continent. A spatio-temporal exploration of Libya is conducted to gain understanding about the political stability of the country prior to, during and post the foreign military intervention that occurred in 2011. We demonstrate use of an effective spatio-temporal analysis framework that can be used as a model synthesis when applied to the ACLED dataset.

1 Introduction

The Armed Conflict Location and Event Data project (ACLED) provides data and analysis on political violence and protests across the African continent and ten countries in South and South-East Asia. The dataset was introduced by Raleigh and Hegre in 2009 [6]. The project serves a humanitarian purpose and through collection and analysis of the data hopes to reduce levels of political violence across these regions. This body of work falls under the umbrella of crisis mapping [2]; collection of real-time data that is analysed and visualised during a crisis. ACLED collect data from a wide variety of in-country sources and produce real-time analysis and visualisations of political unrest to enhance international understanding of these events.

1.1 Motivation

A variety of crisis mapping organisations exist around the world with the purpose of building knowledge of how a region can change in time relevant to external stimuli; climate, political unrest or disease.

This paper utilises the database produced by ACLED. African countries face numerous challenges in their pursuit of political stability. Violence in the name of a political ideology can halt progress. This paper builds on the excellent work delivered by the humanitarian project of ACLED to produce a case study of visual analytical research that showcases how effective use of a spatio-temporal framework can yield further insight from a large and complex dataset.

Our case study focuses on the country of Libya. In 2011, Libya was a prominent news headline due to the staging of a multi-state military intervention in the country. This came to a conclusion in October 2011 when the Libyan leader, Colonel Gaddafi was killed. We investigate the political events pre, during and post the military intervention to discover whether the loss of Gaddafi has increased or decreased political instability.

2 Data

ACLED data is made available to the public from their website [1]. The data contains over 100,000 events as of February 2016. The time interval for the data lies between 1997 and present day, providing the most complete information on political violence events within this region. Each recorded event contains a date and geographical co-ordinates of the specific location the event occurred in. Additional information includes the type of event, number of fatalities, responsible actors and changes in territorial control.

An event is considered for inclusion if it meets explicit criteria; the event was, either within or outside the context of a civil war, motivated by political ideology. Primary emphasis is placed on "…violence against civilians, militia interactions, communal conflict and rioting."

In order to contextualize the current instability of African countries and the scale of politically related conflicts, a high level overview of spatiotemporal analysis is provided on the entirety of the available data, 1997 – 2015.

2.1 Features

There were 23 references in the original dataset, I identified 12 as being important to my analysis. ACLED provide a 22-page codebook on how the data is captured and descriptions of both numeric and textual information. I describe the relevant data fields below:

Latitude – Numeric: Co-ordinate indicative of relative position to equator. Range [0, 90]
Longitude – Numeric: Co-ordinate indicative of relative position on Earth with relation to Greenwich Mean Time. Range [-180 , 180]
Year – Numeric, indicates year of event occurrence
Event-Date – Date time format: contains day, month and year of event
Inter1 – Numeric: Discrete value indicating the named actor involved in the event
Inter2 – Numeric: Discrete value indicating the named actor involved in the event
Interaction – Numeric: Discrete value indicating the interaction between the named actors.
Event Type – Text: Categorisation of the event, split into 8 sub-sections. Explanations provided below.
Location – Text: Name of city, town or village where the event occurred.
Notes – Text: Description of the event, granular detail often provided
Fatalities – Numeric: Count of fatalities caused by event occurrence.
Population – Numeric: Total number of citizens with permanent residence in the location.

2.2 Event Type

A major challenge in this field of work is gaining knowledge of and then understanding and interpreting the event that took place. ACLED are aided in this department by a vast network of collaborators who contribute information about the location they habitat within. This can then be corroborated with media networks and reliable sources. With such a vast quantity of information reported it is necessary to discretize the events to enable further analysis.

The ACLED team have identified 9 distinct event types that occur previous to, during or post a conflict. All reported events are therefore bucketed into 1 of the 9 categories when added to the dataset. This form of clustering is manual and relies on subject matter experts to pre determine the 9 categories and subsequently establish what event type a new event is bucketed within.

I will explore trends of event type over time and space, it is therefore necessary to understand how ACLED define each event type. I provide a description summary;

Battle-No change of territory – Control over contested territory does not change following violent dispute between two or more groups.
Battle-Non-state actor overtakes territory – Control of territory assumed by a non-governmental group following a violent dispute. The actor in this case is predominantly a rebel group.
Battle-Government regains territory – Control of territory is requisitioned by a government body or group acting in allegiance to a government.
Headquarters or base established – Sole actor event of non-violent type. Occurs when a permanent or semi-permanent base is established. Rare occurrence.
Strategic development – Non-violent event that occurs within the setting of violent dispute e.g. recruitment rally. Also included in this section are peace talks and high profile arrests.
Riots/Protests – Public demonstrations against political groups, government incumbents or rebel militias. Protests are non-violent, riots are violent.
Violence against civilians – Any act of violence committed against individual or groups of unarmed civilians.
Non-violent transfer of territory – Acquisition of territory with no violent conflict. Actors can be governments or rebel groups.
Remote Violence – Violence committed by a group that is spatially removed from the violence when it occurs e.g. bombings or IED attacks. Victims of remote violence can be groups in dispute or civilians.

2.3 Interacting Agents

The variables *inter1* and *inter2* define the principal and secondary agents involved in an interaction that became a conflict event. The dataset uses numeric values to signify an acting group e.g. 'Government / mutinous other' = 1. Eight categories were identified and coded as follows;

Code	Agent
1	Government or mutinous force
2	Rebel force
3	Political militia
4	Ethnic militia
5	Rioters
6	Protesters
7	Civilians
8	Outside/External force

Numeric representation of textual data facilitated the creation of the *Interaction* variable. *Inter1* and *Inter2* were combined to give a 2-digit representation of the interaction that occurred between agents e.g. 36 was an interaction of political militia and protesters.

3 Methodology

The exploration of data relating to human death caused by acts of violence can be emotive and as such can be a sensitive subject in which to participate. Discovery and interpretation of emotive data is necessary to advance progress towards political stability. The statistics gathered demonstrate what is happening, this paper aims to explore how we can visually interpret the 'what' to better understand the 'why' and 'how'. In the absence of emotion this is a Geographical time series problem, events linked to temporal and spatial attributes.

Scheiderman proposed a visualization workflow suitable for spatio-temporal analysis comprising of three simple steps; overview first, zoom and filter, details on demand [3]. Proposed in 1996 the mantra holds relevance to modern visual analytics. Visual analytics has emerged as a scientific discipline capable of combining modern computing power with representative visualizations in order to extract meaning and insight. To this extent the original mantra proposed by Schiederman has been extended to incorporate the algorithmic complexity of computational analysis.

Keim et al [4] introduced a new mantra as applied to visual analytics in 2008; 1) analyse first, 2) zoom, filter and analyse further, 3) details on demand. This modern approach to Schiederman's original mantra utilizes the speed and memory of modern technology coupled with human interaction to extract the most knowledge possible from data. Fig.1. drawn by Keim et al [5] provide a useful visual workflow to better demonstrate the process of visual analytics and its methods of yielding knowledge.

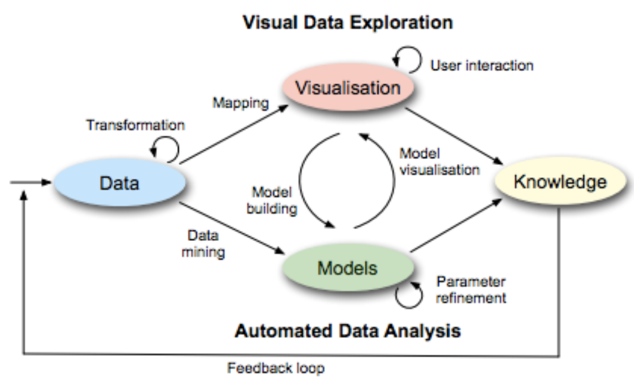

Fig.1 Visual Analytics Workflow

3.1 Related Work

Recently, various academic bodies have begun compiling large datasets in the pursuit of micro level analysis of in country violence. In addition to the work done by Raleigh and Hegre, the pioneers of ACLED [6], event datasets include the Political Instability Task Force (PITF) Worldwide Atrocities Dataset [7] and the Social Conflict in Africa Database (SCAD) [8]. We have chosen to analyse ACLED due to the inclusion of non-violent and violent data which we believe will be more suited to a spatio-temporal analysis as non-violent events may prove to be pre-cursors to violent events.

3.2 Spatio-Temporal Analysis

Spatial-temporal models are built when the dataset contains a combination of temporal and spatial data. The ACLED dataset contains conflict events measured in space and tracked along a time period. Spatio-temporal models are often the result of a large/complex dataset. The computer aggregates, clusters and filters the data in preparation for analyst interaction where the pattern recognition of the human brain takes over.

This paper follows the analytical framework first proposed by Andrienko et al [9] for spatio-temporal modelling of large/complex data.

1) Cartographic map display – We use the following methods to display spatio-temporal data.

Choropleths – Data plotted on choropleths was normalised prior to plotting, this is an essential step to avoid simply producing a density plot that reflects the population rather than the desired effect size. The higher the population of a country the higher the potential cost of collateral damage from conflict events. I chose to normalise by population density per km^2. We argue that this is more appropriate than the total population as the more land a country has the more opportunity exists for a political conflict to occur within its borders. Population densities were found at [10]. If normalisation has not occurred this is explicitly stated.

Heat maps – Similar in kind to choropleths but do not rely on overlaying data onto a world map and so can be expanded beyond countries. Different shapes can be used to represent attributes, the size and colour of these shapes represent features of the attributes. Excellent examples can be found in work by Rand [11].

Tree maps – Nested rectangles of varying size and colour are displayed in a fixed size structure. Each branch of the tree, one rectangle, represents one attribute of the data and is sub divided into smaller branches. Dimensions of the data vary with size and hue. This technique was first introduced by Scheiderman [12].

Charts – Plotting along an x and y axis, varying visuals can be created by plotting data in two dimensions; Area charts, stacked bar charts and horizontal bar graphs are used in this paper.

2) Time series display – Temporal data is binned into discrete intervals in preparation for trend analysis. Spatial data will be aggregated and combined with temporal data to form spatial time series.

3) Clustering – A common technique used for unsupervised learning; k-Means or k-Nearest Neighbour. Clustering was out of scope of this paper due to lack of continuous data. Only the fatalities dimension used numbers for their cardinal meaning. All other dimensions used text or numeric values to represent categorical variables e.g. nominal numbers used for latitude and longitude.

4) Interactive visual interfaces – Production of visualisations that include temporal, spatial or attribute filtering.

3.3 Statistical Analysis

Statistical Descriptors – Excel was used to calculate mean and standard deviation values at various stages of the analysis.

Data Transformation – Normalisation of data produced new values used later in analysis e.g. dividing sum of fatalities by population density per km^2.

Pearson Coefficient – The Pearson coefficient was used to calculate the correlation coefficient between various variables.

3.4 Computation

Spatial Aggregation – summation of events aggregated by spatial location e.g. country. Each example has been assigned a country based on its latitude and longitudinal co-ordinate. Co-ordinates are mapped to a country according to the database collated by fallingrain.com [13] a website dedicated to building a data warehouse for co-ordinates and estimated population of villages, towns, cities and airports.

Temporal Aggregation – summation of events according to their assigned time. A year in this case assumed the standard definition, beginning on 1st Jan and ending 31st Dec.

Reshaping – Transforming data between long and short form dependant on software used for analysis. Tableau requires data to be in long format for plotting, the structure of the original dataset was wide format. This was converted using pivot tables in Excel. If this process required automation Python and R are both capable of quick transformation.

3.5 Software

Tools used for analysis:

Python – Scipy used for Pearson coefficient calculations, Numpy and Pandas used for reshaping and data transformation.

Microsoft Excel – Used primarily for data transformation, use of pivot tables and vlookups made generating desirable data structure quick and easy.

Tableau – Used for generation of choropleths, heat maps, temporal trend analysis and additional statistical visualization. Excellent clustering and filtering functionality facilitates quick analysis.

4 Implementation & Analysis

This paper implemented a three stage analysis using the mantra proposed by Keim et al [4] accompanied at each of the three stages by the framework proposed by Andrienko et al [9].

4.1 Analyse First

The overview provides a contextualized understanding of how volume of conflict events has altered temporally and which African nations have been affected the most by political conflict.

Following the mantra of Keim, this involved harnessing both the computational process of spatial and temporal aggregation of the data examples, followed by human pattern recognition to pick out key trends. Jankowski et al [14] state that this first step is essential to give the analyst a chance to spot interesting patterns which would otherwise be lost in large and complex datasets.

Area charts were created which combined the quantitative total event counts and total fatality counts coupled with their cause. Displaying the data in an area chart allowed visualization of the total time period which enables trend analysis for each event type by time period. Hue was chosen as an agent to distinguish discrete variables [15]

The data was filtered by 'Event Type' and 'Fatalities' to produce two-time series distributions showing how the number of reported events and the sum of fatalities changed over time. The analytics software used to perform these processes was Tableau which automates the aggregation to allow the analyst additional time for visualization.

Figures 1 and 2 offer immediate knowledge, since 2010 the number of conflict events in Africa has risen dramatically. This rise has been driven by three events; Battle – no territory change, riots/protests and violence against civilians.

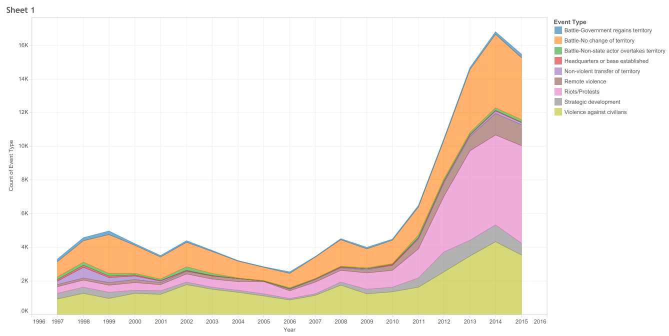

Figure.1 - Trends in event type count per event type per year.

Figure.2 shows the sum of fatalities has returned to numbers seen at the beginning of the 21st century. This steady increase is noticeable post the relatively stable period of 2005-2009. The significant contributors to this rise are; Battle – no territory change and violence against civilians.

The two figures appear visually to correlate with each other with the exception of 1998/99. Further exploration shows that Angola was responsible for 73% and 38% of these fatalities respectively. Combined, Eritrea and Ethiopia were responsible for an additional 53% in 1999. Ground truth knowledge of this time period understands that the huge spike in fatalities was a direct cause of the civil war in Angola which persisted between 1998 and 2002 and the Ethiopia/Eritrea war 1998-2000. Due to the impact of war, 1998 and 1999 show anomalous fatality numbers.

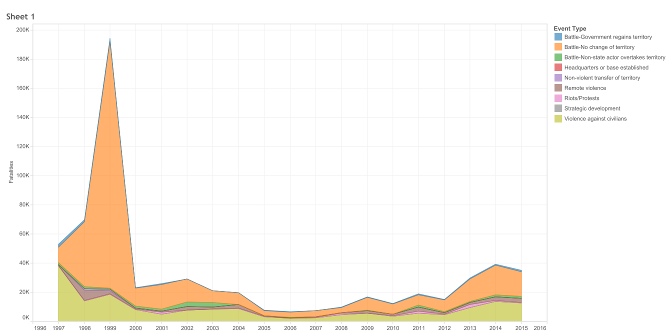

Figure.2 - Trends in fatality count per event type per year.

Using hue as a categorical distinguisher allows the analyst to easily discern that while Riots/Protests are a major contributor to the rise in conflict events, they are not a significant contributor to fatality counts.

If the spike in fatalities for 1998 and 1999 caused by the aforementioned wars are treated as anomalous events and subtracted from the data, the two visualizations appear to follow a similar trajectory. This was tested using a Pearson correlation coefficient which confirmed a moderate correlation of 0.45, the corresponding p-value was 0.054, marginally outside the standard threshold to be determined statistically significant. We can conclude from this statistical test that a relationship of some proportion exists between number of conflict events and number of fatalities.

The pink band is very prominent in figure.1 and as discussed makes a significant contribution to total event counts, it is however far less prominent in figure 2. This disparate contribution could reduce the correlation. To confirm this theory, the Pearson correlation coefficients were calculated for fatality sums against both 'Battles – no change of territory'[1] and 'violence against civilians'[2].

Table 1. Correlation of Sum of Fatalities v.s X

X	Pearson Coeff	p-value
Battles – no change of territory'[1]	0.68	0.001
Violence against Civilians[2]	0.60	0.007
Total Event Count	0.45	0.054

Table1. shows the strongest relationship to sum of fatalities are the events where Battles occur with no change in territory. Prior to investigation it would be expected that fatalities would increase with a larger number of battles or larger number of acts of violence. This result increases confidence in the integrity of the data.

We, the human analysts, identified key time periods from the area charts. We then created further visuals to understand the countries that were an influencing factor on the identified time periods. Heat maps were created, which allowed all features and attributes to be shown on the same visual; the size of a square was used to represent relative number of events while a divergent black-red colour scheme was used to represent number of fatalities. The timeline was filtered to show only the key time periods chosen by the analyst. Both size and colour used a min-max range, the effect of this is to show how extreme countries are in total event and fatality count.

Following creation, the minimum value for total events was adjusted to 100 events. The reduction in the number of countries represented in the graphic minimised irrelevant information. This process was repeated, once for absolute values and once for normalised values.

Figure.3 – Heat map – Absolute Values: Event count & fatality count by country and year

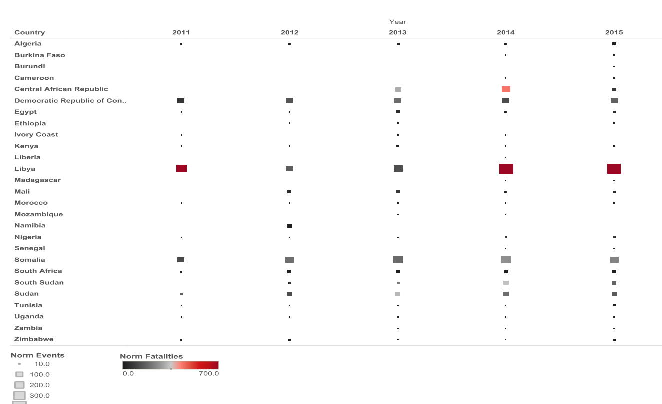

Figure.4 Heat map – Normalised Values: Event count & fatality count by country and year.

Data transformation was required for heat map visualisations; the sum of fatalities was normalised by each countries population density per km^2. To facilitate the data transformation, the data was manipulated using a pivot table, exported to a separate workbook and joined with population density values.

The heat maps are useful in understanding that although Nigeria exists at the maximum end of the range in number of absolute fatalities it is not the majority contributor to total conflict events. The Democratic Republic Congo, Egypt, Libya, Somalia, South Africa, South Sudan and Sudan all contribute a comparable or greater number. In contrast, when viewed in normalised terms Libya's population is greatest effected by number of political conflict events and number of fatalities.

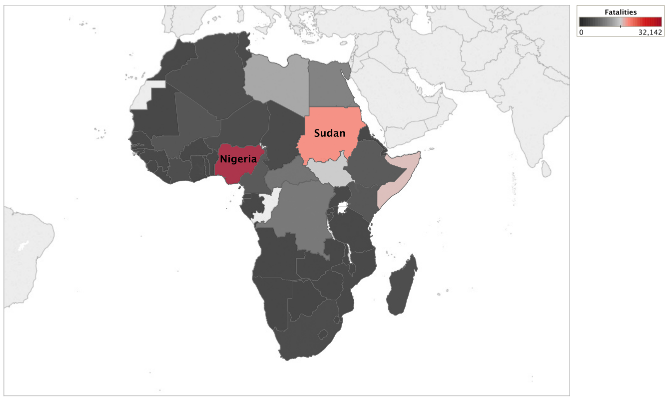

Figure.5 – Choropleth – Absolute Values: Fatality count by country, 2011-2015.

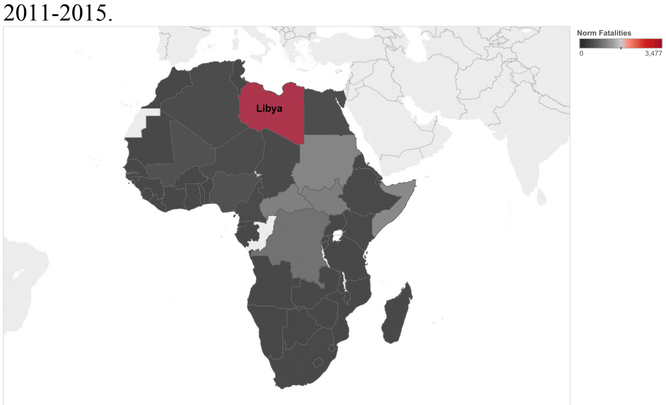

Figure.6 – Choropleth – Normalised Values: Fatality count by country, 2011-2015.

The choropleths utilised the same timeline filtering, colour scheme and range as the heat map. Event count was excluded as a feature.

The divergent colour scheme makes it easy to understand Nigeria has suffered an extreme number of fatalities in comparison to its continent neighbours in absolute terms. However, Figure.6 shows that Libya has in fact suffered the greatest number of fatalities in proportion to its population density. The choropleth shows that between 2011 and 2015 Libya has suffered double the number of proportional fatalities than the nearest effected countries.

4.2 Filter, Zoom & Analyse

The middle stage of the process produced country specific analysis. The spatial data was filtered and we zoom away from a continental view to focus our analysis on the country of Libya.

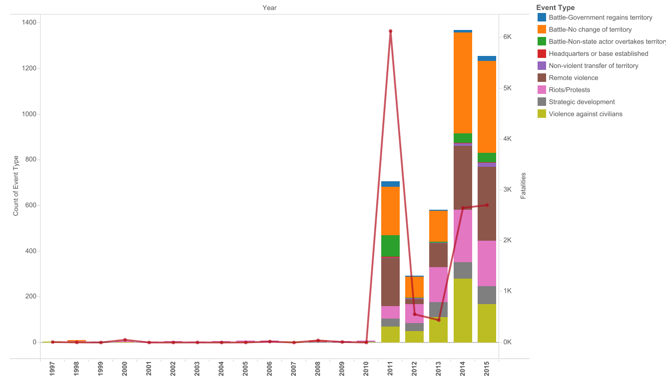

Figure.7 Stacked bar chart – Event count distinguished by hue, trend line shows sum of fatalities per year

Figure 7 is dramatic in understanding how few politically related conflict events, and thus fatalities caused by conflict events, occurred prior to 2011. It is not within the scope of this paper to investigate why the Libyan population was so sedate from 1997-2010 but we can speculate with some confidence that the leadership of Gadaffi was an influencing factor. We know that the Libyan intervention was a direct result of his leadership and therefore we expected a spike in event counts during 2011. Unexpectedly we see that the country has not recovered from the disposal of Gadaffi and although loss of life has not been as severe as 2011 the number of events remains significant year on year. The proportion of event types appears consistent across the last 5 years; Battle – no territory gained, remote violence, riots/protests and violence against civilians are all prominent post 2011.

Point maps were created to spatially visualise; population densities, event counts and fatality counts. Hue was used to distinguish event type and variance in sphere size was used to represent event count and /or fatality count.

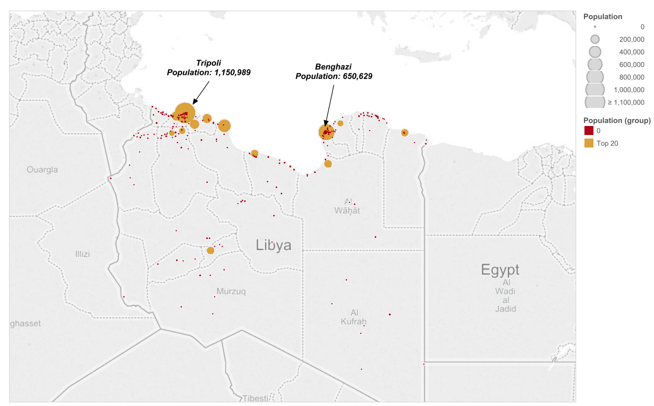

Figure.8 – Spatial location of at least 1 event; orange circles represent a top 20 most populous city, scaled by size proportional to population. Red circles represent towns/villages.

Figure.8 gives understanding of the spatial location of Libya's population. The top 20 most populous cities were located and plotted in orange, figures found at [16]. Six of the top 20 most populous cities were found not to have had any events within this time period and were not plotted. The remaining red circles are towns or villages with populations of under 60,000 but contained a minimum of 1 conflict event. We learn that the majority of the population reside along Libya's Northern coastline, this includes Libya's capital Tripoli and second most populous city, Benghazi. We expect the majority of events and fatalities to occur along the coastline with particular prominence in close proximity to the two principal cities. We are interested in the absolute numbers so this visual is essential in understanding the population spread of Libya.

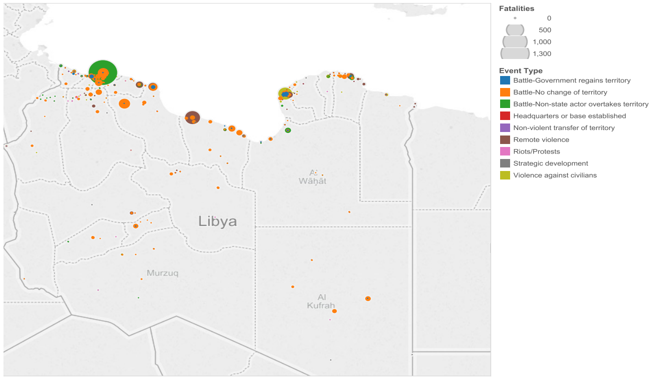

Figure.9 Events resulting in fatalities, size represents sum of Fatalities, hue represents Event Type.

Figure.9 is a static snapshot of the sum of fatalities connected with a spatial location. The time period shown in this image is 2011 – 2015. All years were additionally explored individually and pre 2011 was explored as an aggregated view. In connection with figure.7 the map was incredibly sparse for those years prior to 2011, whether individual or aggregated. Through visualisation of individual years, it was apparent that 2011 was dominated by a few key events which led to significant fatality counts. For example, one conflict event located in Tripoli resulted in 1,300 fatalities. This can be seen dominating figure.9, represented by the green sphere. In contrast 2014 and 2015 which both show fatality counts in excess of 2,500 had no singular event produce more than 100 fatalities.

Fatality counts and event counts were tabulated for the years 2011-2015. A hue was superimposed over the tabulated figures and the saturation scaled according to the proportional volumes. Fatality saturation range was set between 0 and the maximum sum of fatalities associated with 1 location; Tripoli, 2,677. Event count saturation range was between 0 and 200. Adding colour to the figures allows the analyst to identify key locations and years with little effort, in the absence of colour all values must be analysed, brute force computation, to extract required knowledge. Two tables were created, one showing counts for locations of populations greater than 60,000, making it a 'top 20' location by population and a second table for locations with populations lower than 60,000.

	Year					
Location	2011	2012	2013	2014	2015	Grand Total
Tripoli	2,524	18	34	58	33	2,667
Sirte	1,217	1	20	23	264	1,525
Benghazi	434	47	132	405	262	1,280
Misratah	728	1	7	2	13	751
Sabha	0	97	14	121	102	334
Ajdabiya	118	1	12	18	76	225
Az Zawiyah	155	0	15	16	10	196
Gharyan	0	94	0	51	1	146
Tagiura	54	1	1	1	36	93
Tobruk	10	0	1	6	23	40
Al Khums	0	3		0	20	23
Tarhuna	0	2	1	0	12	15
Yafran	7			0	0	7
Al Marj		0		1	0	1

Fatalities
0 2,667

Table.2 Tabulated fatality totals by location and year. All locations have > 60,000 citizens. Blue saturation represents proportional sum of fatalities

	Year					
Location	2011	2012	2013	2014	2015	Grand Total
Darna	12	3	14	79	226	334
Buatni			1	184	96	281
Bani Walid	226	23	1	0	0	250
Kiklah	12			238	0	250
Zlitan	236	12	0	0	1	249
Bannah				228	17	245
Al Layti		1	1	70	151	223
Al Aziziyah	0	3		198	20	221
Al Kufrah	0	109	20	11	63	203
Bin Jawad	55			30	98	183
Qasr Bin Ghashr				162	0	162
As Sabiri al Gharbi			4	36	94	134
Brega	128	0	0	1	0	129
Al Hawary		5	0	55	47	107
Az Zahra				100	0	100

Table.3 - Tabulated fatality totals by location and year. All locations have < 60,000 citizens. Blue saturation represents proportional sum of fatalities

The two tables display disparate fatality counts. This was expected, political events are more likely to occur in regions of importance such as large urban areas. Table.3 was filtered to only show a five-year fatality total > 100. Of interest was how the fatality counts in 2014 and 2015 were largest in areas outside of Tripoli. We can identify 11 locations over the two-year period that suffered fatality losses in excess of 100 people while Tripoli only suffered 91 across the two years. This could be indicative of increased stability in Tripoli. We decided to explore this further and identify whether a reduction in the number of events was the cause.

	Year					
Location	2011	2012	2013	2014	2015	Grand Total
Tripoli	148	59	163	187	111	668
Sirte	42	2	20	20	66	150
Benghazi	35	83	166	238	112	634
Misratah	93	5	12	18	43	171
Sabha	11	7	19	57	40	134
Ajdabiya	29	2	12	24	43	110
Az Zawiyah	34	3	10	29	37	113
Gharyan	12	5	1	24	24	66
Tagiura	12	1	3	3	10	29
Tobruk	8	1	6	15	14	44
Al Khums	6	3		3	5	17
Tarhuna	2	2	1	1	8	14
Yafran	8			1	1	10
Al Marj		1		3	3	7

Count of Event Type
1 200

Table.4 – Tabulated event counts for 'top 20' locations by year. Brown saturation chosen to represent proportional number of events.

We can immediately disprove the notion above; event count has remained consistent across 2014/15 with the count of 2011. We can therefore infer that the type of political event has instead altered. In the absence of foreign military intervention, fatalities caused by the foreign militaries have disappeared. The stability however has unfortunately not improved. Instead it appears that while the majority of fatalities and events were concentrated in Tripoli in 2011 this has slowly filtered to different locations. An identical table for non 'top 20' locations was produced which supported this finding.

4.3 Details

The purpose of the final stage of the analysis is to extract all knowledge possible in the pursuit of addressing the original motivation of the analysis. In addition, it is useful to now explore unexpected findings from the higher level analyses which may either, aid the original motivation or answer questions that arose during working.

We address our original motivation by focusing on the time period of 2011, this built a timeline of events that displayed how political instability began, why foreign military intervention occurred and how political instability has now become a permanent fixture of the Libyan landscape.

Filtering the time period to Jan 1st 2011 – Dec 31st 2011 yielded a detailed understanding of Libyan political events during the year. The chart shown in figure 10 has been displayed in descending order of total interactions. The first column displays the principal agent; the second column shows the secondary agent. The horizontal bars represent the count of interactions between the two agents.

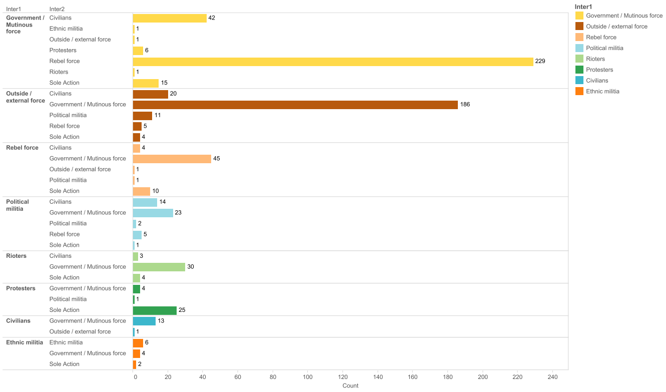

Figure.10 – Chart showing interaction counts as a result of interactions between agents.

The decision to assign principal and secondary agent was one taken by the curator of the dataset. We have been unable to identify the process behind this assignment and as a result have been forced to make the following assumptions; the assignment of principal and secondary agent are not done randomly and the order carries significance in resulting interactions. For instance, we would attribute 229 interactions to 'Government / mutinous other' when interacting with 'Rebel forces'. However, we would only assign 45 interactions to 'Rebel forces' when interacting with 'Government / mutinous other'.

The chart tells us that 'Government / mutinous other' were responsible for the most interactions as principal agent, they were also the highest cause of interactions when acting as secondary agent in all events except 'Ethnic militia'. We can conclude that the Libyan government was the most frequent actor during 2011. 'Outside / external force' was a minimal actor as a secondary agent but was responsible for over 200 interactions as the principal agent. This is understandable, foreign groups had the ability to interact through remote violence, for example; air strikes. It would have been very difficult for other groups to initiate an interaction with a group that had limited physical presence in the country. 'External/ other forces' includes NATO forces or the UN, their involvement arose from the foreign military intervention.

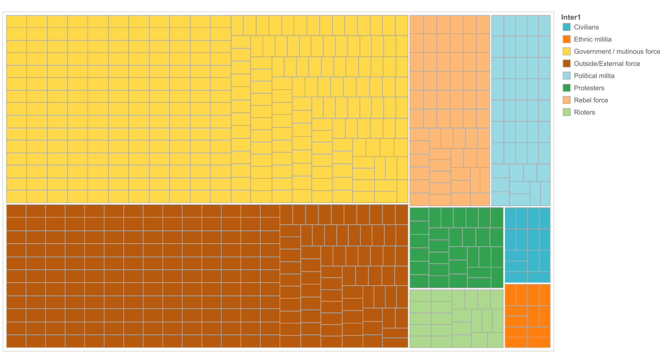

Figure.11 Count of events by principal agent during 2011

The creation of two tree maps facilitated our ability to extract knowledge quickly. Having identified the government as the key interacting agent, we wanted to create a view that would corroborate the horizontal bar chart but offer a simpler method for comparison. All blocks in Figure.11 are equal size and represent one event independent of fatalities. It is much easier to extract that as principal agent the Government and external forces account for nearly ¾ of all events.

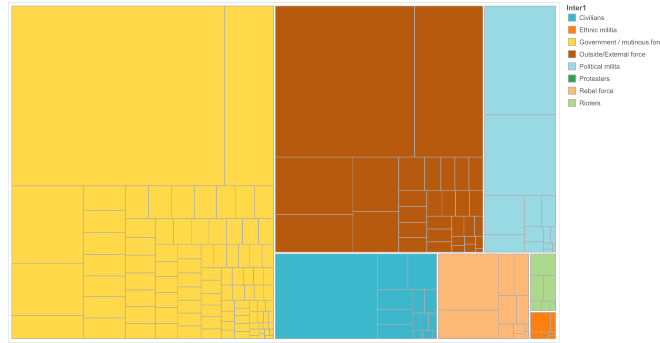

Figure.12 Number of events with at least 1 connected fatality by principal interaction agent. Block size proportional to number of fatalities.

Our second tree map included fatality counts. Figure.12 establishes that the two most frequent protagonists of interactions were also the most violent. The block sizes emphasize the volume of fatalities caused by one event, we see two stand out events; 1 brown, 1 yellow and a secondary level of events; yellow, brown, dark blue, light blue and beige. All are worth further investigation to understand why so many fatalities occurred. It is interesting to note that although protesters were responsible for a comparable number of events with rioters, political and ethnic militia they did not initiate events that led to any deaths.

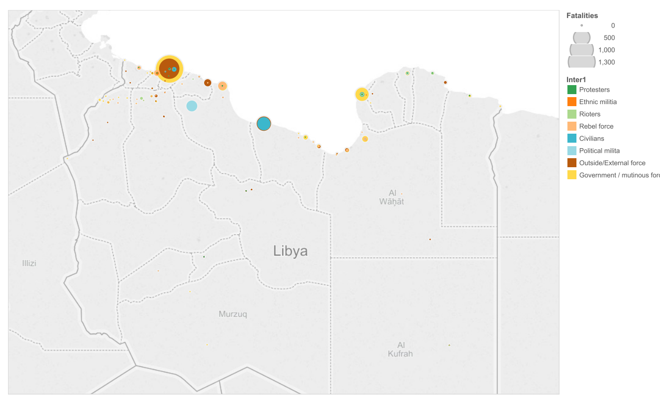

Figure.13 Spatial representation of fatalities by principal agent

We replicate the visual shown in Figure.12 but this time incorporate spatial data and plot the fatalities according to event location. The tree map was easily interpreted and therefore important in understanding fatality amounts. Figure.13 is much less interpretable in terms of quantifying fatality totals by principal agent. It does however give insight into how the fatalities were distributed across Libya. We immediately identify 3 key areas along the Northern coast; Tripoli, Sirte and Benghazi. In contrast to Figure.9 we can see that political unrest had not yet moved away from the Northern coast in 2011. Fatalities especially, were concentrated along the coast line as depicted by the larger spheres. Only one incident involving a political militia has resulted in significant loss of life inland. On closer inspection this location was Bani Walid, a known defensive stronghold of pro Gaddafi forces.

A temporal understanding of 2011 was gained by discretising the temporal data into months. Two temporal plots were created which when coupled together were useful in documenting the event types and interaction protagonists of those events that led to the highest fatality counts.

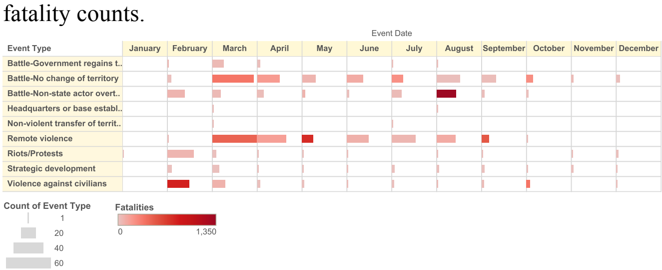

Figure.14 – Bar plot of event type by month. Saturation represents fatality count.

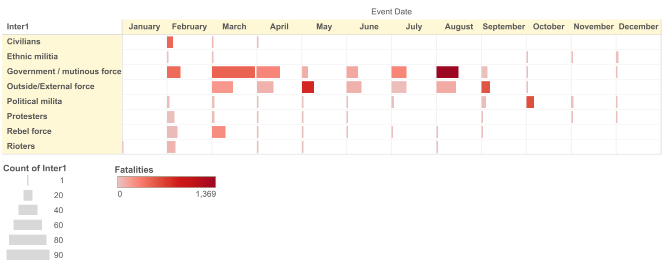

Figure.15 – Bar plot of principal agent interactions by month. Saturation represents fatality count.

Both plots highlight the relevance of the time period February – September. The reduction in events from September until the year end occurs directly after the month in which the most fatalities were suffered. The majority of fatalities suffered in august were due to a battle that resulted in a change of territory suggestive of a significant victory. The primary cause of fatalities prior to August are Battles with no change of territory and remote violence caused by Government and external forces. If August can be viewed as the beginning of the end, February appears to be the spark that ignited the dispute. There is a large volume of riots and protests in tandem with violence against civilians, leading to fatalities. All interacting agents except external forces appear to have been active in February.

Our final visuals are a combination of temporal data, spatial data, interaction agent, notes and source. The addition of notes and source provide specific event knowledge explaining how and why fatalities were caused. We have built sufficient knowledge from higher level analyses to filter for specific locations and time periods. Figure.16 is a snapshot of February 2011 of the Northern cities Benghazi and Sirte. Figure.17 is a snapshot of August 2011 and Tripoli. This method was repeated for all events of significance that were identified from the tree maps.

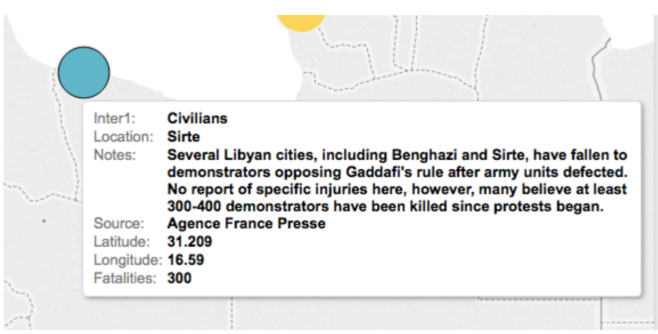

Figure.16 – Display of Notes connected with specific event of interest

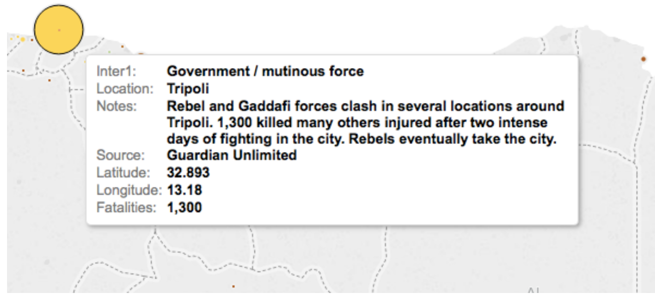

Figure 17– Display of Notes connected with specific event of interest

Figure.17 connects the visuals depicted in figures 9 and 13, the large fatality event centred in Tripoli was a fragmented battle that spread through the city with fighting lasting two days. Figure.9 showed that the fighting led to a change of territory occupation and figure.13 demonstrated the principal agent was the government. The accompanying notes confirm both of the previous findings. This demonstrates that even in the absence of notes, the analytical method used through this body of research was informative and effective at extracting knowledge. The addition of notes allows us, human agents, to identify areas of interest in time and space and then use the Notes for further information.

5 Results

To better address our original motivation, we have split our answer into three sections.

Libya Pre Foreign Intervention – The country specific analysis started in section 4.2 showed that Libya was, in political conflict terms, a stable and peaceful country from 1997 – 2011. Our visual shows minimal fatalities and very few politically motivated events. We make no attempt to address the reasons for this stability as this would be speculation but the evidence provided by this analysis is robust in ascertaining that Libya was relatively peaceful in comparison to the average African nation during this time period.

Libya During Foreign Intervention – In section 4.3 we showed that during the months of February – October 2011, Libya became a very politically active nation. Over the course of the full year 2011, over 700 conflict events occurred with each event producing an average of 8 fatalities. This is in stark contrast to the previously peaceful 14 years. We must conclude that in February 2011 the peace was broken and political opponents of Gaddafi were pivotal in the increase of political events and fatalities. This was further exacerbated by the intervention of foreign militaries. This meets expectations, Libya would not have become a prominent news feature during this time if peace and stability were in evidence.

Libya Post Foreign Intervention – Colonel Gaddafi was killed in October 2011, following his death foreign militaries exited the country to allow political mediators to establish a new regime. The evidence shows that Libya has been destabilised since the events of 2011. Political events have increased in 2014 and 2015 to numbers beyond those realised in 2011. Thankfully, fatality counts have declined on the final count of 2011 but are still far in excess of fatalities seen 1997-2010. We have shown the country is far less peaceful than the years preceding the fall of Gaddafi.

We can conclude that the country since 2011 has become one of the most dangerous and politically unstable countries across the African continent. We demonstrated that for this period, the country has suffered more loss of life in terms of population per km^2 than any other country in Africa. This, we expect, was not the intentions of the states who contributed to the foreign military intervention and should perhaps, in combination with evidence from other countries who have been subjected to similar interventions, be investigated thoroughly to learn future lessons.

6 Discussion

This research paper has successfully extracted knowledge regarding political events in Libya from 1997-2015. We believe the work demonstrated here can be replicated across spatial regions and/or time periods contained within the ACLED database. Our work can be viewed as a synthesis for future case studies.

We iterated through statistical analysis, computational techniques and visualisations at each stage of the spatio-temporal framework. All three analytical methods complemented our extraction of knowledge. As a case in point, the normalisation of fatality data, coupled with correlation between event counts and fatality counts informed the production of choropleths.

Each subsequent step of our 3 step framework was informed by the work preceding it. In section 4.2 we focussed on the time period post 2011 due in part to the visuals produced in section 4.1. We zoomed further in section 4.3, this time dictated by our findings in section 4.2.

At each step of analysis, we explored both expected and unexpected findings. Use of time and spatial filters allowed agents to quickly progress through numerous visuals in the pursuit of identifying the important information.

6.1 Limitations

The ACLED dataset contains only one continuous feature, the remainder are discrete. This reduces the computational techniques available for analysis, we discussed this in our analysis framework. If time had allowed, we could have found additional data that could be joined with the ACLED data in order to extract further knowledge of the actors involved in the conflict events.

Each event is given equal parity in the dataset. For instance, the violence that led to 1300 deaths in Tripoli is recorded in a similar manner to a non-violent conflict resulting in no fatalities. This reduces the insight we can learn about the provocations that led to the actor's behaviour and thus the nature of the conflict. We are restricted to temporal and spatial patterns that offer interesting historical insight but may prove ineffective at implementing future prevention.

6.2 Future Work

Work that builds upon the knowledge presented in this paper would require additional data from alternative sources. We believe we have successfully answered our motivating questions and future work would look to address the new questions that have arisen as a result of answering the original. This could be achieved using the techniques demonstrated in this paper.

Future work that involved the analysis of ACLED data would choose a different region to concentrate an analysis on. In 2014/15 the terrorist group Boko Haram has been very active in Nigeria. We showed that Nigeria has suffered the highest fatality counts in recent years in absolute terms and a study of this time period and geography following the framework laid out in this paper would yield further insight into where and how fatalities are increasing.

The 'Notes' dimension contained textual information, the potential to employ NLP techniques exists. Latent Dirichlet Allocation is a technique that could have provided an alternative to the pre-selected actors and event descriptions. The same visualisations could then be retained and identification of population groups or prominent political figures may come to the fore.

References

[1] Raleigh, Clionadh, and Caitriona Dowd. *ACLED Codebook*. 2016. Print.

[2] "Crisismappers". *Crisismappers.net*. N.p., 2016. Web. 22 Mar. 2016.

[3] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In IEEE Symposium on Visual Languages, pages 336–343, 1996.

[4] Keim, D. et al (2008): Visual Analytics: Definition, Process, and Challenges. In Kerren, A., Stasko, J.T., Fekete, J.-D., and North, C. (Eds.) Information Visualization – Human-Centered Issues and Perspectives. Volume 4950 of LNCS State-ofthe-Art Survey, Berlin: Springer, pp.154-175

[5] Keim, Daniel. *Mastering The Information Age*. Goslar: Eurographics Association, 2010. Page 23, Print.

[6] Raleigh, Clionadh, and Håvard Hegre. *Introducing ACLED: An Armed Conflict Location And Event Dataset*. Oslo: International Peace Research Institute, 2009. Web. 17 Mar. 2016. Centre For The Study Of Civil War.

[7] "Data: Political Instability Task Force Worldwide Atrocities Dataset". *Eventdata.parusanalytics.com*. N.p., 2016. Web. 17 Mar. 2016.

[8] Denver, University. "Social Conflict In Africa Database | University Of Denver". *Du.edu*. N.p., 2016. Web. 17 Mar. 2016.

[9] Andrienko, Natalia. *A Visual Analytics Framework For Spatiotemporal Analysis And Modelling*. Sankt Augustin: Fraunhofer Institute. Print.

[10] "African Countries By Population 2015 - Statisticstimes.Com". *Statisticstimes.com*. N.p., 2015. Web. 2 Apr. 2016.

[11] Rand, D.A. "Mapping Global Sensitivity Of Cellular Network Dynamics: Sensitivity Heat Maps And A Global Summation Law". *Journal of The Royal Society Interface* 5.Suppl_1 (2008): S59-S69. Web.

[12] Shneiderman, Ben. "Tree Visualization With Tree-Maps: 2-D Space-Filling Approach". *TOG* 11.1 (1992): 92-99. Web.

[13] "Directory Of Cities And Towns In World". *Fallingrain.com*. N.p., 2015. Web. 24 Mar. 2016.

[14] Jankowski, Piotr et al. "Discovering Landmark Preferences And Movement Patterns From Photo Postings". *Transactions in GIS* 14.6 (2010): 833-852. Web.

[15] MacEachren, A M, C A Brewer, and L W Pickle. "Visualizing Georeferenced Data: Representing Reliability Of Health Statistics". *Environ. Plann. A* 30.9 (1998): 1547-1561. Web.

[16] "Most Populated Cities In Libya". *Worldatlas.com*. N.p., 2016. Web. 2 Apr. 2016.