Exploratory Analysis: A look into conflict in Myanmar

Alexandra Plassaras (amp2261)
G4071 - Advanced GIS and Spatial Analysis
May 16th, 2016

Introduction and Background

The purpose of this paper for me was to conduct an exploratory analysis on geo-referenced data and test out my GIS and Spatial Analysis skillset. Exploratory Data Analysis (EDA) is an approach to conduct data analysis that employs a variety of techniques (mostly graphical) to maximize insight, uncover underlying structure, extract important variables and patterns, and develop outliers and anomalies as well as many other purposes. [1] Thus for this report I wanted to use the GIS and Spatial Analysis skills I have learned over the past year to conduct an analysis of a data set to see what I could find.

After deciding on the fact that I wanted my focus for this paper to be exploratory data analysis, I decided to use a dataset from the Uppsala Conflict Data Program (UCDP). The reason I chose this dataset is because it is one of the most accurate and well-used data-sources on global armed conflicts and its definition of armed conflict is widely accepted as a standard in how conflicts are systematically defined and studied.[2] UCDP defines conflict as: "a contested incompatibility that 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."[3] Since I found data that was geo-coded and very diverse and had a lot of information for decades I decided to use this as my data set. As you will see in the results section, after conducting my analysis I

---

1 "1.1.1. What Is EDA?" 1.1.1. What Is EDA? N.p., n.d. Web. 16 May 2016. <http://www.itl.nist.gov/div898/handbook/eda/section1/eda11.htm>.

2 "Program Overview." - Uppsala University, Sweden. N.p., n.d. Web. 16 May 2016. <http://www.pcr.uu.se/research/ucdp/program_overview/>.

3 Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand (2002) Armed Conflict 1946-2001: A New Dataset. Journal of Peace Research 39(5). http://www.pcr.uu.se/digitalAssets/124/124920_1codebook_ucdp_prio-armed-conflict-dataset-v4_2015.pdf

ended up focusing much of my report on Myanmar and the deaths that have occurred due to conflict there. Therefore in the conclusion/discussion section I will touch upon some of my findings and how they relate to the real world.

Methodology

The dataset used in this report was taken from the Uppsala Conflict Data Program (UCDP). The main information that was taken from this dataset was the location where deaths occurred, the year the deaths occurred and how many deaths were a result of conflict in each place. In regards to location, the UCDP database aimed at collecting the best spatial resolution possible however due to various limitations not all locations were coded with the same specificity. The following table outlines the various types of location specificities[4]:

Location Code	Meaning
1	exact location of the event known and coded
2	event occurred within at maximum a ca. 25 km radius around a known point. The coded point is the known point.
3	only the second order administrative division where an event happened is known. That administrative division is coded with a point representing it (typically the centroid).
4	only the first order administrative division where an event happened is known. That administrative division is coded with a point representing it (typically the centroid).
5	the only spatial reference for the event is neither a known point nor a known formal administrative division, but rather a linear feature (e.g. a long river, a border, a longer road or the line connecting two locations further afield than 25 km) or a fuzzy polygon without defined borders (informal regions, large radiuses etc.). A representation point is chosen for the feature and employed.
6	only the country where the event took place in is known.
7	event in international waters or airspace

Table 1 – UCDP's Geo-precision Classification

[4] Sundberg, Ralph, and Erik Melander, 2013, "Introducing the UCDP Georeferenced Event Dataset", Journal of Peace Research, vol.50, no.4, 523-532
Croicu, Mihai and Ralph Sundberg, 2015, "UCDP GED Codebook version 4.0", Department of Peace and Conflict Research, Uppsala University

For the purposes of this analysis, the data was filtered to include only events that were geo-referenced with at least three decimal places. All coordinates that did not meet this standard were removed from analysis. Additionally all coordinates were truncated to only have the first three decimal places for each location. By doing this, the results are able to pinpoint an event within 111 meters.[5]

The variable for the count of deaths caused by conflict was created using three variables from the UCDP dataset. These three variables for each event are lowest reliable estimate of total fatalities (*low_est*), highest reliable estimate of total fatalities (*high_est*) and best estimate (*best_est*). The problem with using any of these variables is that they vary. For example, in 2001 and in 1993 in Afghanistan a certain event's estimated death toll was:

Year	Best_est	High_est	Low_est	New_est
2001	441	4441	441	2882
1993	0	3717	0	929.25

For the purposes of this report the average of low and high estimates were first taken followed by an average of that number with the best estimate. This new number then became an average estimation of deaths per location and time (see *New_est* column above).

5 "How to Measure the Accuracy of Latitude and Longitude?" Lat Lon. N.p., n.d. Web. 16 May 2016. <http://gis.stackexchange.com/questions/8650/how-to-measure-the-accuracy-of-latitude-and-longitude>.

Furthermore, the given the size of the UCDP database, not all information was used. The Georeferenced Event Dataset, which is a subset of the UCDP data, was used as this dataset is an event-based and georeferenced dataset. Each event in this dataset (compared to the larger dataset) includes events that have coordinates attached for each event. Each unit of analysis is a single event of organized violence in which at least one person has been killed.[6] More specifically the UCDP defines an event as "An incident where armed force was taken by an organized actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date".[7]

For the purpose of this paper, deaths that resulted from a conflict as measured by UCDP was the unit of analysis focused on in this paper. As this paper is an exercise in conducting exploratory analysis, the whole dataset was first looked at and narrowed down as the analysis continued. First the Asian continent was chosen to look at deaths caused by conflict. Then the region surrounding Myanmar was selected given its relatively high death count and its numerous borders with other countries. Lastly, the state and administrative regions were looked at within Myanmar to conduct spatial analysis and confirm whether there were space-time clusters in existence in this dataset. The types of analysis conducted in this study looked at standard coordinate mapping onto shapefiles, aggregating deaths up to the country level, exploring spatio-temporal relationships and then finally running a space-time analysis using a space-time

---

[6] Croicu, Mihai Catalin, Ralph Sundberg, 2015, " UCDP Georeferenced Event Dataset Codebook Version 2.0", Department of Peace and Conflict Research, Uppsala University. And Croicu, Mihai Catalin, Ralph Sundberg, 2012, "UCDP GED Conflict Polygons Dataset Codebook version 1.1-2011", Department of Peace and Conflict Research, Uppsala University. And Ralph Sundberg & Erik Melander, 2013, "Introducing the UCDP Georeferenced Event Dataset", Journal of Peace Research, July 2013, Vol.50, No.4, 523-532 http://www.pcr.uu.se/research/ucdp/datasets/ucdp_ged/ and http://ucdp.uu.se/downloads/ged/ucdp-ged-40-codebook.pdf

[7] *Ibid.*

permutation probability model that looked for high rates of death in a given area over a span of time. Spatial regression using residual analysis was attempted but neither the queen contiguity weight with 1 order of contiguity nor the queen contiguity with 2 orders of contiguity were significant enough to continue the analysis. The first order had a Moran's I of 0.0441705 with a p-value of 0.168 while the second order had a Moran's I of -0.197184 with a p-value of 0.138. Since neither were statistically significant this analysis will not be covered in the report but the results can be seen in Appendix A.

Additionally the shapefiles that were used to map our georeferenced conflict deaths were taken from DIVA-GIS, a free computer program for mapping and geographic data analysis.[8] In part of the analysis the data was separated into two groups – 1) Myanmar and the surrounding administrative regions in India, China and Thailand and 2) Myanmar only. This was done to examine further whether there were high death counts along Myanamr's border or not. Then, as the analysis continued, the scope was finalized into only looking at deaths that had occurred in Myanmar. To see the maps used in both sets of analysis, see Appendix B.

Results

The first part of the analysis for this report was to map the coordinates where deaths had occurred on a map. Figure 1 shows this information. It is important to note that the coordinates represent any deaths that occurred as a result of conflict within about 110 miles of the actual location. Using this specificity there are 2862 unique locations where deaths occurred. It is important to note that this number does not represent the amount of deaths that occurred, merely the location.

8 "Download Data by Country." DIVA-GIS. N.p., n.d. Web. 16 May 2016.

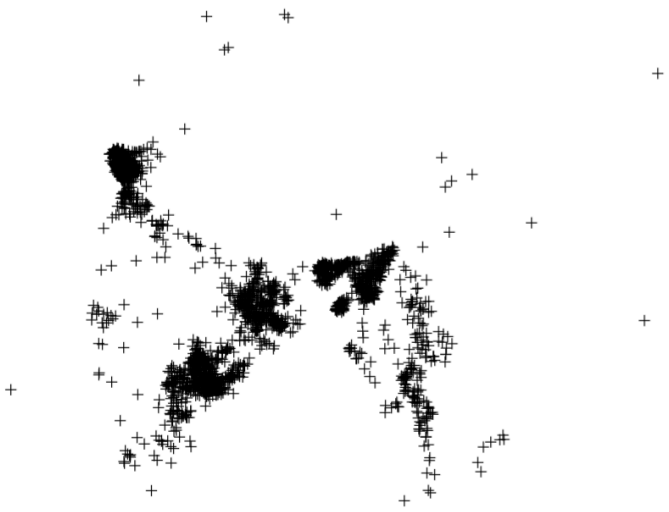


Figure 1 – Conflicts Caused by Death in Asia

From Figure 2 below we see that the highest incidents of death caused by conflict in Asia are found in Afghanistan, India, Pakistan, Myanmar, and Thailand. Given that Afghanistan has been rife with conflict where many countries have attempted to affect the outset of the conflicts – for example: the Soviet Union and the United States – and that both India and Pakistan who also had high amounts of conflict caused deaths have been in perpetual conflict with one another over various issues (i.e. Kashmir), I chose to look at a country that has a high conflict death count but not a lot of international participation in the country. Thus my two best options were Myanmar and the Philippines. In the end I chose Myanmar mainly because it shares borders with multiple countries yet has had little international influence on its conflicts, as they have mainly been internal ethnic conflicts.

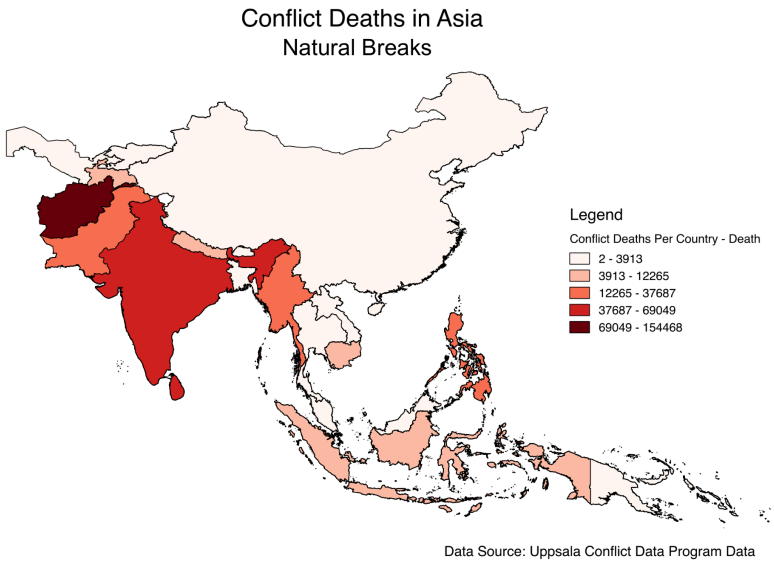


Figure 2 – Aggregate Conflicts Caused by Death in Asia by Country

Figure 3 shows the unique locations of deaths caused by conflict in Myanmar and the surrounding areas. These areas include Yunnan Province in China, Arunachal Pradesh, Nagaland, Mizoram and Manipur in India, and Chiang Rai, Mae Hong Son, Tak and many others in Thailand. The reason this analysis was done looking at bordering regions as well was because often conflicts, especially ethnic conflicts don't end right at the border. Spillover effect might have been an issue to further investigate. That is why for the next four visualizations this report will be comparing results from deaths in Myanmar compared to deaths in Myanmar and its surrounding areas. From the image below we can also see that the Indian bordering regions appear to have many more unique locations where deaths occurred compared to the other bordering countries and even Myanmar itself. There might be a few explanations for this – 1) Information about deaths related to conflict might not exist or be impossible to confirm in places like China where the government continues to censure what the world knows about internal affairs,

2) Some of the conflicts that plague Myanmar have been pushed out of the country towards India[9] or 3) perhaps there is another explanation that has yet been explored.

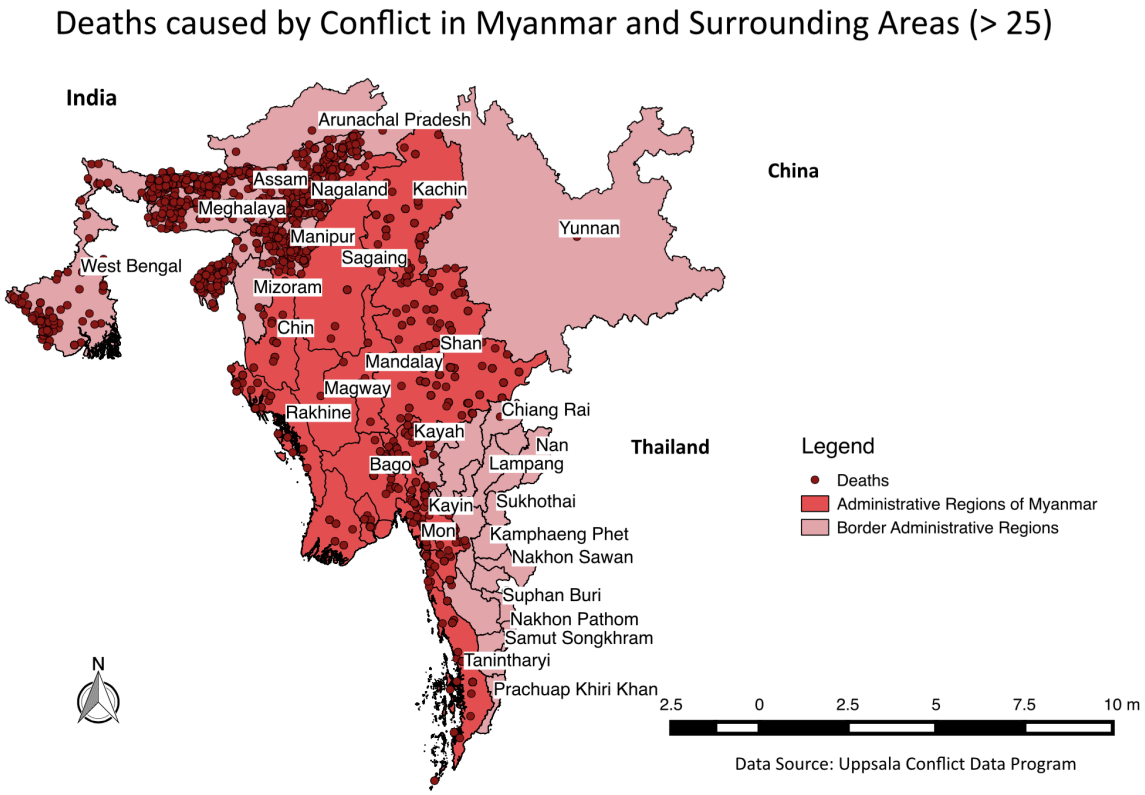


Figure 3 – Deaths caused by Conflict in Myanmar and Surrounding Areas

Next the analysis looked at potential evidence of spatio-temporal relationships looking across 2826 unique locations over a span of 26 years. After filtering for both Myanmar and bordering areas and Myanmar, 659 unique locations were found to be within the surrounding borders and Myanmar and 175 unique locations were found to be in Myanmar. For this analysis I looked at the death observations in 4-year periods. Figure 4 below shows the results of squared deaths over time (which maximize the variation

9 *"The artificial delineation of the borders was one of the major factors that led to the newly independent states of India and Burma reaping a harvest of insurgencies since 1947-48. There is still a lack of stability in the location of tribal populations, particularly on the Thai, Chinese and Indian borders due to refugee movement and migration."* "The Invisible Country Ethnicity & Conflict Management in Myanmar Sushil K. Pillai*." The Invisible Country Ethnicity & Conflict Management in Myanmar. N.p., n.d. Web. 16 May 2016.

within the data) in the two regions. On the left hand side we see that the bordering regions are affecting the results heavily since there are many unique locations clustered around the Indian borders (dark blue points). We also see that there are some slightly dense points that show up along the Thai border in the visualization on the left (blue dots) that do not show up when looking only at Myanmar. However we do see dense points on both maps that show high volumes of deaths in Myanmar (the yellow dots). Overall we see that the trend for both sets of visualizations is that increase in deaths from 1989 to 1997, followed by relatively similar levels of deaths up until 2009 ending in a decrease in deaths by 2013.

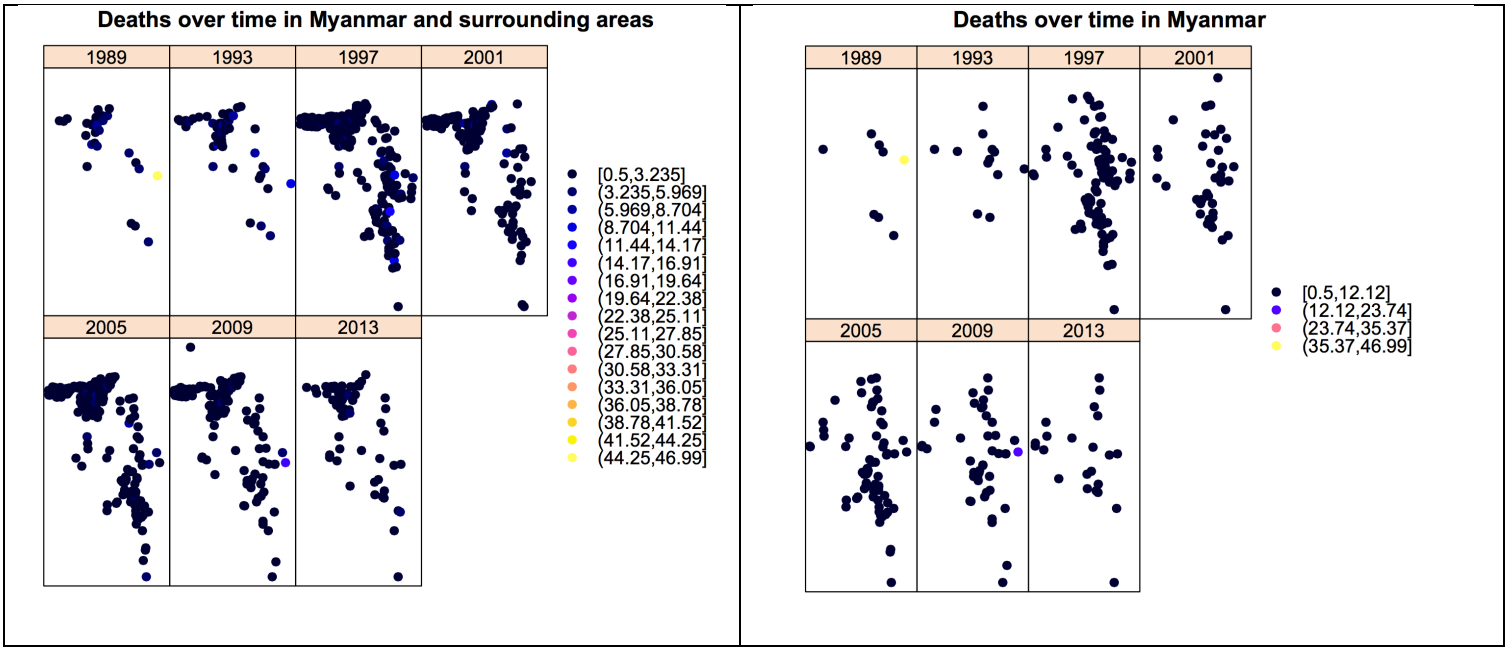


Figure 4 – Deaths over time in both Myanmar and surrounding areas (left) and in Myanmar (right)

Another approach to exploring space-time data is to examine the weighted mean time count, which measures abundance and early presence of deaths occurring in these areas. Using a Hovmoller Diagram we can review the weighted mean time count for both

areas. In the left-hand visualization we see an increase in deaths over the years however, given that the data used for this visualization contained 659 unique locations it is difficult to ascertain which locations followed a particular pattern. When we examine only the 175 unqiue locations found in Myanmar we see a slightly different picture. First we see that there was a high level of death in the early 1990s in a particular location that after this time period didn't see many deaths related to conflicts. Then again around 2010 we see another spike in deaths in a particular region. Thus, this graph tells us that deaths were not necessarily increasing over time. This result led to the decision to look at space-time clusters in Myanmar (see Figure 7 below).

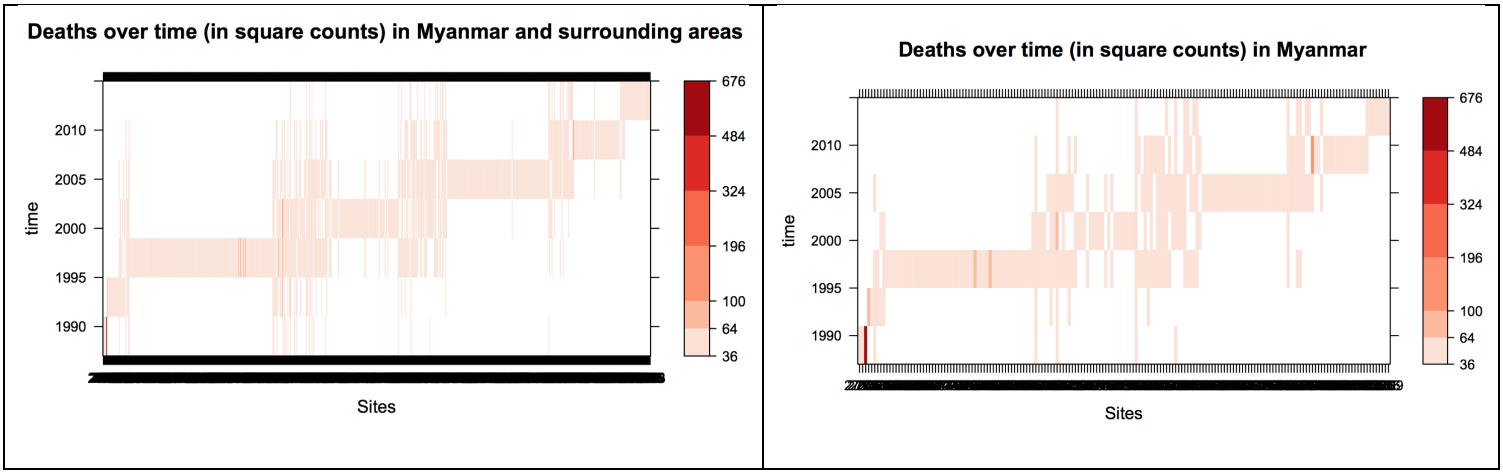


Figure 5 – Deaths over time in both Myanmar and surrounding areas (left) and in Myanmar (right)

Before analyzing the data for space-time clusters, the death count was mapped within Myanmar only as seen in Figure 6. Here we see high counts of deaths in the Ayeyarwady Region, Shan State, Sagaing Region, Kachin State, and Kayin State.

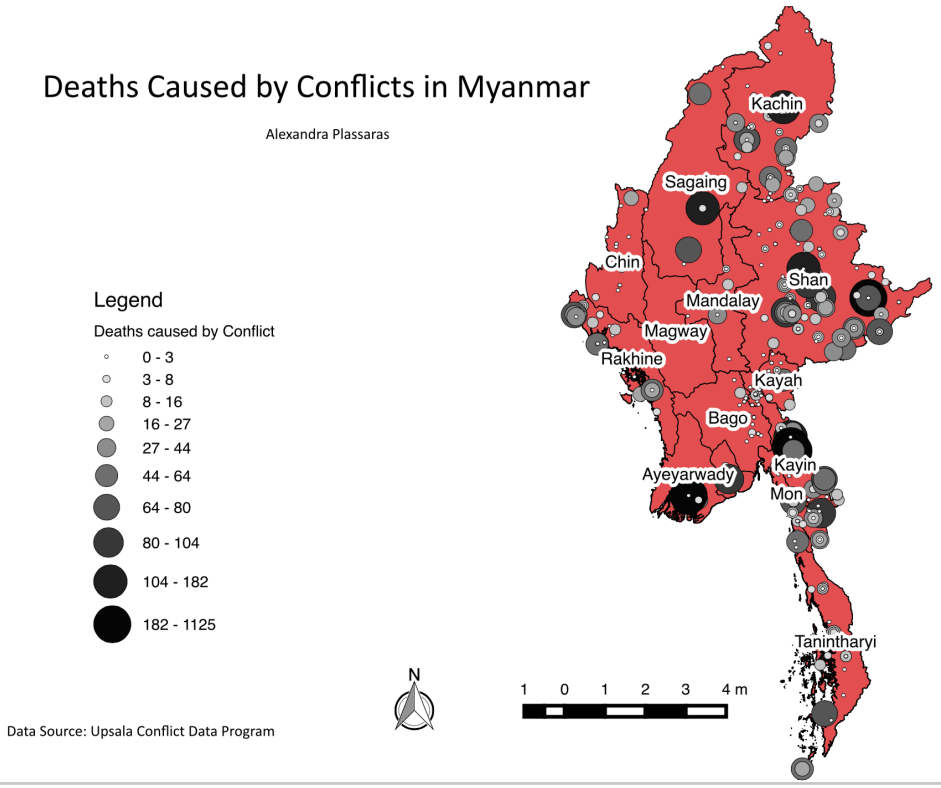


Figure 6 – Distribution of Deaths Caused by Conflict in Myanmar

State/Region	Year	p-value
Kachin	2011/1/1 to 2014/12/31	< 0.00000000000000001
Ayeyarwady	1991/1/1 to 1991/12/31	< 0.00000000000000001
Tanintharyi, Mon	1996/1/1 to 2000/12/31	< 0.00000000000000001
Bago, Yangon	2006/1/1 to 2007/12/31	< 0.00000000000000001
Mandalay	2003/1/1 to 2003/12/31	< 0.00000000000000001
Kayah	1996/1/1 to 1996/12/31	< 0.00000000000000001

Table 2: Space-Time Cluster Results

One would expect that given these large amounts of deaths focused on a singular location would equate to space-time clusters. In order to detect significant clusters simultaneously in both space and time I used space-time data exploration to see what results the data had for us. The results shown in Table 2 show the area and the given time period for which significant space-time clusters were found. Figure 7 displays the results visually.

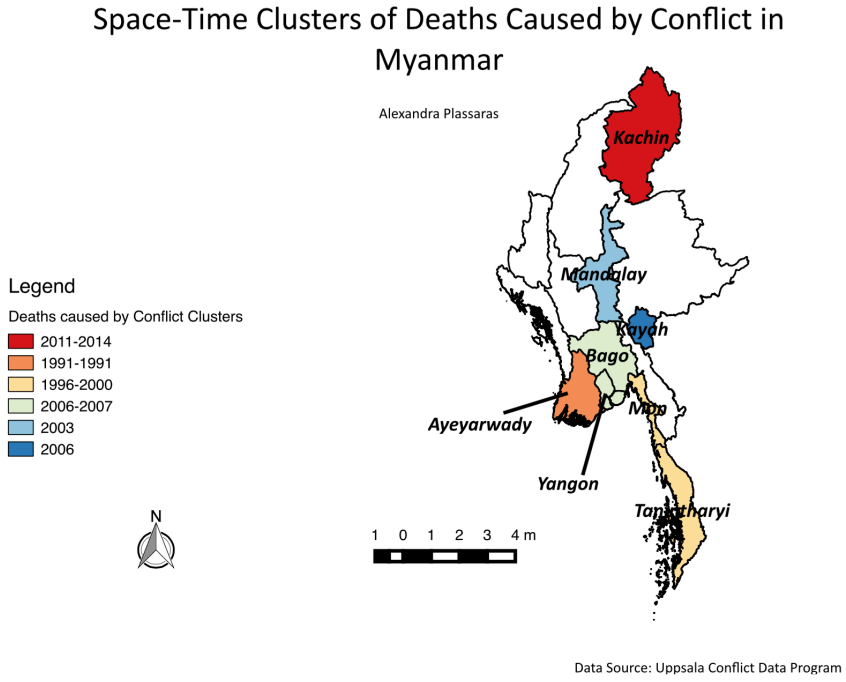


Figure 7 – Space-Time Clusters Caused by Conflict in Myanmar

From Figure 6's depiction of high death counts in the Shan State and Sagaing Region it is puzzling why these regions were not also found to have significant space-

time clusters. Additionally the Tanintharyi and Mon areas as well as the Bago and Yangon areas did not appear to have high distributions of deaths over the time period but our analysis proved differently.

Conclusion

After systematically analyzing all conflict deaths starting in Asia to narrowing it down to one singular country, we now have mathematical proof that the space-time clusters found in this report are statistically significant and did not happen by chance. Indeed we can see that the regions highlighted in Figure 7 have historically been placed where ethnic conflict has occurred over the years in Myanmar. Over the years Myanmar has been ruled by a military junta, which is notorious for treating non-Burma ethnic groups harshly which has led to many minorities taking up arms. Some examples of this are insurgencies in the Mon, Kachin, Karenni and Shan-dominated regions. In addition, as the government offers ceasefires to insurgents, the conditions that come with this agreement often include provisions for the group in question to fight against any remaining non-ceasefire groups.[10] Examples of this include conflicts between the SSA-S (Shan State Army - South Command) against UWSA (United Wa State Army) and DKBA (Democratic Kayin Buddhist Army) against KNU (Karen National Union).

Some interesting further study might include separating the conflict related deaths by conflict (for example looking at deaths as a result of the SSA-S vs. the KNU) to have a better understanding of how each individual conflict has changed over the years. While my analysis has been a useful exercise and interesting, it probably does not serve a tool

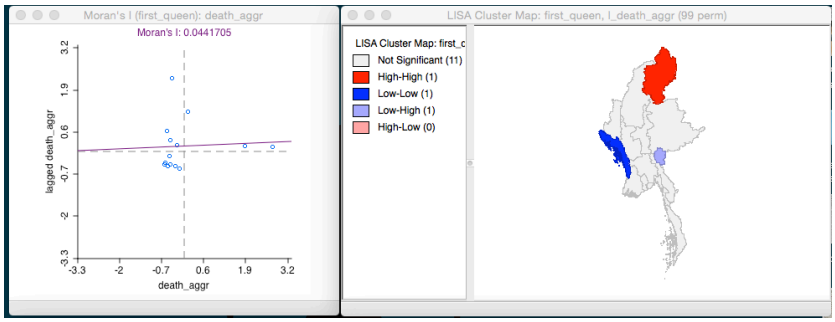
10 "UCDP - Uppsala Conflict Data Program." UCDP - Uppsala Conflict Data Program. N.p., n.d. Web. 14 May 2016.

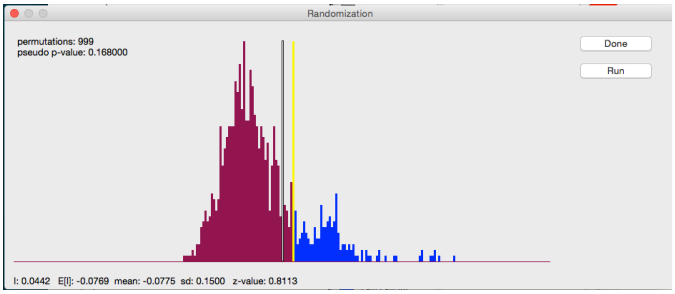
for policy makers or peacemakers/builders to use given that this analysis is done on a country-wide level. Perhaps something more helpful, especially to institutions and organizations that are looking to support conflict-affected areas would be a county level analysis to see exactly where are the places that might need to most help in rebuilding their communities after decades of conflict and death.

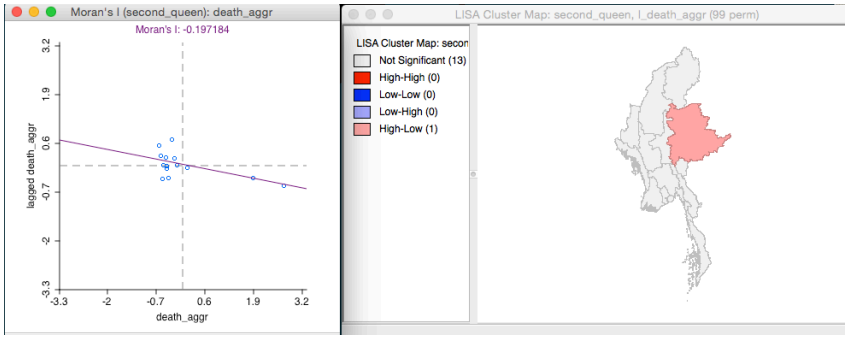
Additionally, comparing space-time clusters along the borders of Myanmar to see how much the violence has changed would also be something that could prove useful. Lastly, it would also be quite interesting to see if given this information and some sort of well-established index or indices if I could create a prediction model that would look at trying to predict where the next conflict related deaths would happen in the country. If that were possible, there would be many ethical questions that would need to be raised and answered on what, if anything, to do with that information.

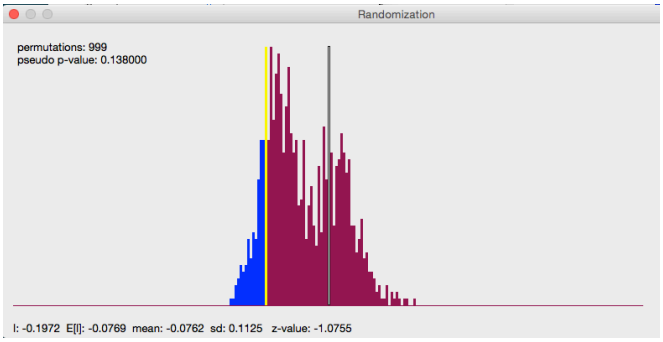
Appendix

Appendix A – First and Second Order Queen Contiguity Results









Appendix B – Maps of Myanmar and Surrounding Border Territories (Left) and Myanmar (Right)