A Regional Study of the Relationship Between Rainfall and Violent Conflict in Ethiopia

Acknowledgements

I would like to thank Phelan Yu and Noah Golowich for sitting through discussions about this project and providing helpful feedback. I would like to thank Nathan Mueller for all his help along the way. I would like to thank Marena Lin for the many crucial e-mails and meetings in the final days; and for taking the time to advise me in both the writing and computation processes. Finally, I would like to thank my advisor, Professor Peter Huybers, for the invaluable guidance and encouragement in both starting and following through with this project.

Contents

List of Figures

Chapter 1

Introduction

In the past few years there has been a growing interest in the connection between climate and conflict. There has been a large increase in the number of papers written on the subject. These papers vary widely in terms of geographic scope, time scale, conflict type, and the aspects of climate and weather they deal with. Perhaps due to this variety in methods and scope, there is not consensus in the literature. Hsiang et al. (2011) are of the opinion that climate is indeed linked to conflict. They found that in countries where the weather system is strongly linked to the El Niño Southern Oscillation (ENSO), there was a significant increase in conflict during El Niño years. Hsiang et al. looked at all ENSO teleconnected countries as a whole, noting that, in general, these countries are dryer and hotter in El Niño years. However, there is reason to believe that ENSO may have a more nuanced influence on precipitation and temperature, and may affect various regions within a single country in opposite ways (Beltrando and Camberlin, 1993). This paper will focus on just one country: Ethiopia. The influence of ENSO on different regions within Ethiopia will be studied. Ethiopia has a particularly interesting climate and topography that comprises arid deserts and cool lush highland forests at elevations higher than 2400 m. With this varied topography come equally varied weather patterns. Thus, it may not be the case that all parts of Ethiopia are hotter and dryer in El Niño years. In fact, parts of Ethiopia seem to have above normal rainfall in El Niño years (Beltrando and Camberlin, 1993). This paper will deal with Ethiopian precipitation and conflict on a regional level, taking into account the both the spatial patterns of rainfall and the geographic distribution of ethnic groups.

This paper comes at an especially relevant, though unhappy, time for Ethiopia. Ethiopia is currently experiencing what threatens to be its most severe drought in 50 years (Famine Early Warning Systems Network, 2015). This drought, along with political and economic factors, is causing a famine that is predicted to be worse than the famine of 1984, during which 1 million people died (IRIN, 2015). This past year was also the strongest El Niño event in about 30 years - which supports the hypothesis that

El Niño is linked to dry years in the tropics. This drought is most severe in the north and central eastern part of Ethiopia. This timing and location agree with Beltrando and Camberlin's (1993) findings, as well as my own findings, that El Niño years are correlated with below normal precipitation in north and central eastern Ethiopia. In addition to this drought, in November of last year protests began in the Oromo region of Ethiopia (Human Rights Watch, 2016). Government crackdown on the protests has become bloody, with an estimated 200 people killed so far. As both a severe drought and violent conflict are happening currently in Ethiopia, the need to understand both misfortunes, and their relationship, seems particularly important.

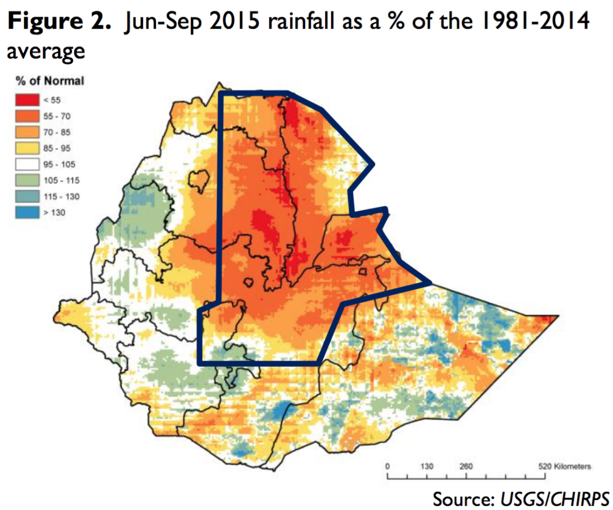


Figure 1.1: Map of the current drought severity from the Famine Early Warning System (2015)

This paper will focus specifically on intra-country conflict between ethnic groups, and between the government and various ethnic groups. After doing a review of the literature on climate and conflict, Gleditsch (2012) believes that these types of conflict are the most likely to be linked to climate. The hypothesis in this paper is that conflict is related to precipitation anomalies. In Ethiopia specifically, I will show that precipitation patterns in different parts of the country are oppositely correlated with ENSO, so some years' precipitation anomalies will be felt unevenly across the country. This makes the relationship between precipitation and conflict more complex, because above normal and below normal precipitation, and their various effects, could occur in close proximity. Perhaps such a precipitation disparity over a relatively small geographic area could exacerbate conflict between regions, moreso than a widespread but evenly felt anomaly.

Ethiopia is particularly interesting to look at because ethnic groups and precipitation, as I will show later, correspond tightly to administrative region boundaries. There is regional homogeneity but country-wide heterogeneity in both precipitation and ethnic groups. Additionally, there have been frequent intra-state conflicts in the recent past (Salehyan et al., 2012). Most of the population is reliant on rain-fed agriculture, so precipitation has an immediate and direct impact on people's lives. Unfortunately, precipitation in Ethiopia has become more erratic and less predictable in the past few decades, and it is predicted to become even more erratic in the future (Jury and Funk, 2013). This makes it increasingly important to be able to predict precipitation, and to understand its effects.

In the remainder of the introduction I will provide background on Ethiopia's ethnic groups, climate, and agriculture. This background is important for understanding the direct and immediate impacts that precipitation anomalies have on Ethiopians, and understanding the possible mechanisms that could cause precipitation anomalies to increase the likelihood of intra-state conflict. In the remaining chapters I will discuss the close correlation between rainfall patterns and the geographic distribution of ethnic groups; the relationship I observed between ENSO and rainfall in different regions; and finally the possible connections between precipitation anomalies and conflict.

1.1 Ethiopia's Ethnic Makeup and Recent History

Almost 95 million people live in Ethiopia, and they belong to more than 14 distinct ethnic groups (C.I.A., 2016). The Oromo, 34.4% of the population, and Amhara, 27%, are by far the largest, followed by Somali at 6.2%, Tigray at 6.1%, Sidama at 4%. The remaining 22% of the population is made up of other small ethnic groups. Very few Ethiopians, less than 20%, live in urban areas. The vast majority of people are rural farmers or pastoralists whose livelihoods are closely tied to the land. Despite being the second largest ethnic group, the Amhara have the most political power and hold almost all federal government positions. This has lead to frequent conflict between the government and the Oromo people, the largest ethnic group, who often feel that they are left out of important political and economic decisions (Human Rights Watch, 2016). In fact, the current conflict in Ethiopia involves Oromo people protesting the expansion of the capital city, which is dominated by the Amhara people.

Ethiopia has had a turbulent recent political history. In 1991 the previous militaristic government, the Derg, was overthrown. In the same year a civil war erupted which resulted in the region of Eritrea claiming independence from Ethiopia. All of Ethiopia's coastline was in the Eritrea region. Its departure made Ethiopia the most populous landocked country in the world. Shortly after this, in 1994, Ethiopia completely redrew its administrative region boundaries as part of the country's new constitution (BBC, 2016). The 1994 administrative regions were based largely on a language survey conducted in the 1980s, and thus they map quite closely to the areas inhabited by the various ethnic groups. In most administrative regions, 90% or more of the population belongs to the ethnic group after which the region is named. The cooler wetter central and northern part of the country is much more densely populated than the dryer and hotter southern part. Ninety percent of the population lives north of the Great Rift Valley which cuts the country roughly in half (Federal Democratic Republic of Ethiopia Population Census Commission, 2008).

The Amhara region is located northwest of the Great Rift Valley in the northern highlands. Almost all the residents of Amhara are Amhara people. This region is characterized by highlands, cool temperatures, and lush forests. Most Amhara are Ethiopian Orthodox by religion and, as in the rest of the country, most Amhara rely on agriculture for their sustenance and livelihood (Federal Democratic Republic of Ethiopia Population Census Commission, 2008). The capital of Ethiopia, Addis Ababa, is located just south of Amhara in the Oromo region. One in two of the capital city's residents are Amhara, whereas only one in five of its residents are Oromo. Though Amhara is the second largest ethnic group in Ethiopia, it dominates the urban areas (Federal Democratic Republic of Ethiopia Population Census Commission, 2008). Federal government positions are almost all held by Amhara people.

In line with the national pattern, most of the residents of Oromia are Oromo people (Federal Democratic Republic of Ethiopia Population Census Commission, 2008). Oromia is the largest region in Ethiopia, spanning both the northern highlands and part of the southern lowlands and eastern desert. As with the Amhara people, agriculture is the livelihood for 90% of the Oromo people (Government of Ethiopia, 2016).

In the Somali and Afar regions to the east, and in southern Oromia and the parts of the Southern Nations Nationalities and Peoples (SNNP) region, most people are pastoralists or agro-pastoralists (Government of Ethiopia, 2016). Most Afar and Somali people, the ethnic groups in the respective regions, are also Muslim. These regions are hotter, dryer and at a lower elevation than the north. In contrast to other parts of the country, there is not much permanent agriculture in the Afar and Somali regions. As a result, many Afar and Somali people are nomadic. They travel between grazing grounds for their animals and places where they can trade. Some people in these regions practice shifting cultivation, meaning that they grow crops in one place for a few years, then move on to find fertile ground once the soil is spent.

There are a few more relevant details about Ethiopia's recent turbulent political history. There have been numerous wars and uprisings. The government was overthrown in 1974 (BBC, 2016). President Mengistu Haile and his military regime, the Derg, took power. During their reign they had thousands of people from the opposition party killed. A few years later, in 1977 Somalia invaded Ethiopia. In 1991 the government was overthrown again and the Derg was thrown out. As previously mentioned, there was a war where Eritrea claimed independence. Then eight years later, in 1999, there was another war with Eritrea over a border dispute. The political scene has also been fraught with disputes. From 1992-1995, after Mengistu was overthrown, there was an interim government. In 1995 Ethiopia had democratic elections and Negasso Gidada became president. That election, along with every national election since, has been accompanied by violent protests and has been disputed by at least one ethnic group (BBC, 2016). Through all of this, Ethiopia has also suffered severe drought roughly every 5 to 8 years, with particularly devastating drought-related famines 1972-1974 and in the mid 1980s (USAID, 2015; Beltrando and Camberlin, 1993). With Ethiopia's chaotic recent political history, diverse ethnic groups, and food shortages, it is no wonder that there are often conflicts between groups in the country.

This next section will provide background on the climates of Ethiopia's various regions, which vary just as much as their ethnic makeup.

1.2 Ethiopian Climate and Weather

Ethiopia is divided into three climatic zones that correspond to elevation: *dega*, above 2400m; *weina dega*, between 1500m and 2400m; and *kolla*, below 1500m (Cheung et al., 2008). The *dega* and *weina dega* zones are often called the highlands, and they will be referred to in the rest of this paper as such. They are located in the northwest of the country and along the Great Rift Valley which cuts across the country from the southwest to the central east. Most of the highlands are in the Amhara and Oromia regions. Mean temperatures in the highlands are around 15°C whereas in the *kolla*, the lowlands, it is much warmer, with mean temperatures around 35°C. The highlands also receive much more rainfall than the dryer lowlands. The lowlands surround the highlands, and are primarily in the Somali, Afar, and southern Oromia regions.

Ethiopia has two rainy seasons and one dry season. The first rainy season, *belg*, lasts from March through May. The second rainy season, *kiremt*, lasts from June through September and is the major rainy season for most of the country, including all of the highlands (Cheung et al., 2008). The *kiremt* rainy season is prominent in the northern highlands, and the *belg* rainy season is prominent in the southeastern lowlands. In fact, in some parts of the highlands, the *belg* rainy season is barely noticeable and there is really only one rainy season - the June through September *kiremt*. Conversely, though the *belg* rains are generally the shorter and smaller of the two, in the south and southeastern lowlands they are actually the primary rainy season. This north-south division between dominant rainy seasons aligns somewhat, though not exactly, with the pastoral versus agricultural regions. Most agriculture is in the northwest, and is more reliant on the greater *kiremt* rains, and most pastoralists are in the southeast, and are more reliant on the lesser *belg* rains. Even though the Afar and Somali regions have similar climates with their high temperatures and low annual rainfall, the *kiremt* rains are the primary rains for Afar and the *belg* rains are the primary rains for Somali. In the northernmost part of the country, around the Tigray region, the *kiremt* rains are shorter and lighter, lasting only from July-September. The dry season is called *bega* and typically lasts from October to January.

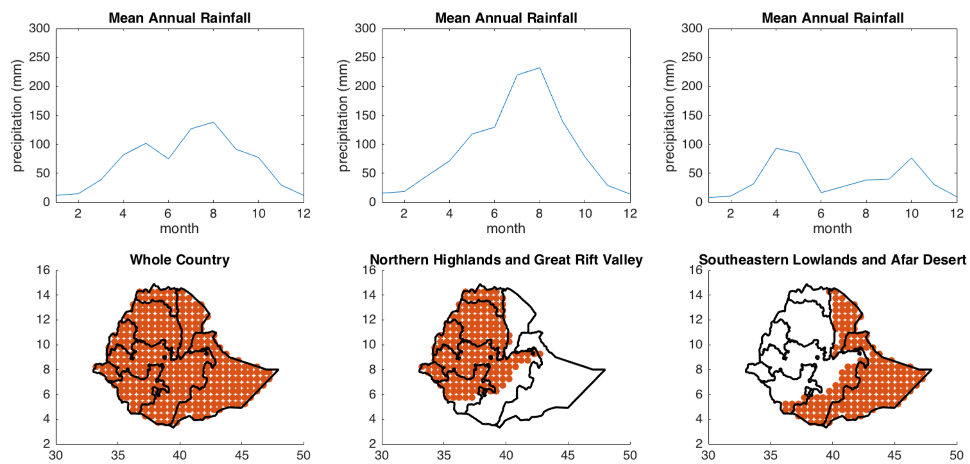


Figure 1.2: Mean annual rainfall for the entire country, the northern highlands, and the southeastern lowlands and Afar Desert. The *belg* and *kiremt* seasons are clearly distinguishable for the country as a whole and for the southeastern lowlands and Afar desert.

Ethiopian precipitation has been found to be related to the El Niño Southern Oscillation (ENSO) (Beltrando and Camberlin, 1993). El Niño years seem to coincide with below-average rainfall in much of Ethiopia, especially north of and around the Great Rift Valley. The next chapter will include my investigation into the relationship between ENSO and precipitation anomalies across the country.

With the predicted increase in global temperatures, we will likely see an increase in extreme weather events. In agreement with this trend, there already has been an increase in the variability of rainfall in Ethiopia (USAID, 2015). As temperatures continue to increase, it is predicted that this trend will continue and rainfall will become even more unpredictable. Both the number and severity of extreme precipitation events such as droughts and floods are likely to become more common. Droughts, floods, and unpredictable rain can all be detrimental to crops. In the next section I will give a brief overview of agriculture and pastoral lifestyles in Ethiopia, and how they are especially vulnerable to precipitation anomalies and unreliability.

1.3 Ethiopian Agriculture and Pastoralists

Ethiopia is extremely reliant on smallholder agriculture for both the wellbeing of its people and its economy. Ethiopian farmers grow 90% of the food that Ethiopians eat (IRIN, 2015). Roughly 1/3 of the land in Ethiopia is used for agriculture, and almost all of this is smallholder agriculture (C.I.A., 2016). The agriculture land is split almost evenly between pasture lands and croplands, with pasture lands taking up a slightly larger proportion. As smallholder agriculture is the prevailing type, the average farm size is quite small, typically less than a hectare (Cheung et al., 2008). These farms are usually run by a single family, and the family will often consume much of what they produce. Almost 90% of working people in Ethiopia are employed in smallholder agriculture, and the industry makes up over half of Ethiopia's GDP. Irrigation is very rare in Ethiopia, especially on smallholder farms. Which means that these farms are reliant on rainfall; thus the people are reliant on rainfall for their food and livelihoods.

The main crops grown in Ethiopia, by area of land planted, are cereal grains such as teff, wheat, barley, maize, and sorghum (Central Statistical Agency of Ethiopia, 2015). Pulses such as chickpeas and fava beans are also important crops. Pulses are often grown interspersed with other crops in the highlands to retain soil nutrients and to provide necessary protein for livestock. Oilseeds, such as sesame, neug, and linseed, are important cash crops. The particular crops grown vary by region. Most of the permanent agriculture in the country is in the highlands where there is greater water avaiability (Abbink, 1993). In the southeast and Afar desert permanent agriculture is much less common. Most people there are pastoralists who rely on livestock more than agriculture. Agriculture in these regions is primarily done by shifting cultivation, meaning that crops are grown on one plot of land for a few years, until the land becomes depleted and the crops are replanted elsewhere.

In the highlands, mostly in the Afar region and northern Oromia, it is cool enough throughout the year that long-cycle crops such as maize and sorghum can be grown (Cheung et al., 2008). These crops are very difficult to cultivate in the lowlands. In some parts of highlands there is only one harvest season per year. However, in areas where both the *belg* (March-May) and *kiremt* (July-September) rains occur (typically in the southern part of the highlands) there are oftten two harvests per year: one in the *belg* season and a second in the *kiremt* season. For this bi-annual harvest, short-cycle crops such as wheat, barley, teff, and some pulses are grown.

Amhara is the largest producer of teff, a grain used to make many staple Ethiopian foods (Government of Ethiopia, 2016). Major cash crops grown in amhara include sunflower, cotton, sesame, and sugarcane. These are typically grown in the lower-elevation parts of the highlands. Other cereals, pulses, and seeds are also grown here. Oromia produces half of the country's agriculture (Government of Ethiopia, 2016). Maize, barley, wheat, teff, pulses and seeds are common, and coffee is the major cash crop. Oromia also holds just less than half of the country's livestock. The southeast of Oromia is dryer and has little agriculture; people there are primarily pastoralists. Most people in the Afar and Somali regions are also pastoralists who earn their living through livestock and trade. Sheep, goats, and cattle are the most common livestock. Camels are also crucial, they provide milk and meat, and they transport people and goods between grazing grounds (Guliye et al., 2007). Some maize, sorgum, and beans are grown in these regions, moreso in the Afar region than Somali, but agriculture is a less important economic activity here relative to other regions.

Many Somali and Afar pastoralists are nomadic. They travel between good grazing land for their livestock and places where they can trade (Guliye et al., 2007). Others, who are agro-pastoralists, will sometimes stay in one place for a few years where they can reliably produce crops. Both pastoralists and agro-pastoralists are highly dependent on rainfall for their livelihoods. In times of drought, the areas they typically inhabit can become barren, and they will have to travel outside their normal region to find grazing grounds. In the past when this has happened, they have sometimes ended up encroaching on other ethnic groups lands, and this has caused violent conflict (Abbink, 1993). The USAID found that the Afar and Somali regions, as compared to other regions in Ethiopia, are especially vulnerable in times of drought or irregular rainfall (2015). The Afar region is one of the the worst affected by the current drought and famine.

Besides forcing pastoralists to migrate, there are other ways in which especially high, low, or erratic rainfall can negatively impact Ethiopians. Below average rainfall in the growing season can cause crop yields to decrease, or, in extreme cases, cause crops to fail completely. Breaks in rainfall during the growing season, even if rainfall over the whole season is not below average, can also hurt crop yields. This is especially true if the break or breaks come near the end of the season, as this can cause the crops to fail (Cheung et al., 2008). The timing of the beginning and end of the rains are also important. A negative effect of unpredictable rainfall is that farmers sometimes employ planting strategies that are risk-averse, and therefore likely to survive a wide variety of rainfall scenarios, but typically result in lower crop yields even if the rains are good. Too much rainfall can also damage crops. In the highlands, many crops are planted on terraced fields or slopes. Heavy rainfall in these areas can cause erosion of the soil or, in extreme cases, landslides that ruin entire fields. Roads in the highlands can erode, become rutted, or washed away in heavy rainfall. This impedes people's ability to travel to markets to trade for things they need, and sell what they produce. In general, floods are a bigger concern in the northwest highlands, and droughts or erratic rainfall ar bigger concerns in the southeast (USAID, 2015). USAID determined that the southeast part of the country and Afar region are most likely to be seriously negatively impacted by drought.

This study will study precipitation in terms of anomalies averaged over the three-month growing season. I suspect that precipication anomalies during the growing season may affect the likelihood of conflict. Though I do not deal directly with the causal mechanisms, I propose a few plausible ones here: low crop yeilds, crop failure, and drought in pasturelands. These are possible links between precipitation to conflict. Some specific mechanisms could be, first, as mentioned earlier, below-average rainfall in pastoral regions could force pastoralists to migrate and encroach on other ethnic groups lands. This could cause tension as the pastoralists compete with the encroached-upon group for land and resources. Second, imagine there are two regions that are home to separate and oft-disputing ethnic groups; there are many examples of this in Ethiopia. If one region receives above average rainfall and the other receives below average rainfall, this could lead one to have an especially prosperous year, and the other to have a year of hardship. In this case the prosperous group might take advantage of the vulnerability of the other group and choose this time to initiate violent conflict. Alternatively, perhaps the group experiencing hardship would, because they were desperate, try to capture land or resources from the other group, and stimulate conflict in that way.

To review, Ethiopia consists of disparate and largely geographically separate ethnic groups, who have a history of inter-group conflict. This motivates my looking at inter-group conflict and explains why conflicts may occur between regions within the country. Ethiopia also produces almost all of its own food. Most Ethiopians are small-lot farmers and pastoralists who are heavily reliant on rainfall. This motivates my looking at precipitation, as it directly impacts people's lives. Further, precipitation patterns in Ethiopia vary by geographic region. This geographic diversity motivates the study of precipitation and conflict by region, rather than over the country as a whole. The next section will explain the data and methods that I used to explore the relationships between precipitation and ENSO, and between precipitation and conflict.

Chapter 2

Data and Methods

The motivation for the data analysis was to explore the relationship between climate and violent intra-state conflict in Ethiopa. This was done in two steps. First, I investigated the relationship between ENSO and precipitation anomalies in different parts of Ethiopia. Second, I investigated the relationship between precipitation anomalies and the onset of conflict. The thought being that, if there was a relationship between ENSO and precipitation, and a relationship between precipitation and conflict, then perhaps understanding these relationships would help us be better able to predict conflicts in Ethiopia. The better we can predict conflict, the better we can prepare for and try to prevent it. The reasons conflicts start are extremely complex, and precipitation is most likely only a small part of the equation. As such, the hope is only that this analysis sheds some light on one piece of the conflict equation.

In the first section of this chapter I will introduce the data that was used. Next, I will explain how Ethiopia was divided into distinct geographic clusters based on precipitation patterns. After that I will discuss the analysis of the relationship between ENSO and precipitation anomalies in each of the geographic clusters. Last, I will go over what was done in investigating the relationship between precipitation anomalies and violent intra-state conflict. For the remainder of this paper, unless stated otherwise, conflict refers to violent intra-state conflict.

2.1 Data

For this study I used precipitation data put together by the Climatic Research Unit at the University of East Anglia (2012). I used the most recent version of their dataset, namely CRU TS v3.23. The temporal resolution is monthly, and the data covers the years 1901-2014. I used only the years 1982-2014 when investigating the relationship between ENSO and precipitation because the ENSO data covered only 1982 forward.

I used precipitation data for the years 1970-2014 when investigating the relationship between precipitation and conflict, because of the years spanned by the conflict data. I used only datapoints within Ethiopia, which ended up being 378 geographic points.

The conflict data used is from the Social Conflict Analysis Database, put together by the Robert A. Strauss Center at the University of Texas at Austin (Salehyan et al., 2012). Only conflicts that occured within Ethiopia were used[1]. Only anti-government violence and extra-government[2] violence were considered. Though other types of violence may also be influenced by precipitation anomalies, my hypothesis is that precipitation anomalies may increase the likelihood of conflict between regions. These two conflict types seemed most appropriate for the hypotheses. The geographic location of the conflict[3] and the number of deaths were included in the dataset.

To quantify ENSO, the ENSO 3 index was used. This was from the the National Oceanic and Atmospheric Administration website (NOAA/ National Weather Service, 2016). It is provided monthly for the years 1982 to the present. Only data up to 2014 was used because of the years covered by the precipitation data. The ENSO 3 index is the sea surface temperature anomaly measured in the western equatorial Pacific Ocean. This index was used because El Niño and La Niña events move eastwards across the Pacific, so events will be detected in the west earlier than in the east.

2.2 Clustering

As discussed above, rainfall patterns vary widely over Ethiopia. Some parts of the country have two distinct rainy seasons, others have only one; some of the highlands get over 100 cm of precipitation per year, whereas much of the Somali region gets just more than 10 cm of precipitation per year. I wanted to study the relationship between ENSO and each of these rainfall patterns individually, so I divided the country into distinct geographic clusters. Another reason to study precipitation by cluster, is that, as mentioned earlier, ethnic groups in the country are largely divided into distinct geographic regions; so if precipitation anomalies vary across geographic clusters, they will

[1] Conflicts that ocurred in Eritrea before its independence were not included.

[2] Anti-government violence is defined in the SCAD dataset as 'Distinct violent event waged primarily by a non-state group against government authorities or symbols of government authorities (e.g., transportation or other infrastructures). As distinguished from riots, the anti-government actor must have a semi-permanent or permanent militant wing or organization.' Extra-government violence is defined as 'Distinct violent event waged primarily by a non-state group targeting individual, or "collective individual," members of an alleged oppositional group or movement. As distinguished from riots, at least one actor must have a semi-permanent or permanent militant wing or organization. Government authorities are not listed as actors or targets.' (Salehyan et al., 2012)

[3] If the exact geographic location was not known, it was estimated by the creators of the dataset (Salehyan et al., 2012)

inherently vary across regions occupied by different ethnic groups.

The goal of clustering is to create distinct groups, clusters, of objects that are more similar to each other than they are to objects in other groups. In this case, each geographic point in Ethiopia is an object, and the goal is to sort them into clusters with similar rainfall patterns. I used the k-means clustering algorithm with k-means++ initial values to create geographic clusters based on rainfall patterns. This algorithm uses an iterative approach to find data points that are close to each other in euclidean space, and create clusters of those points that are close to each other. The k-means++ algorithm is used to choose initial points to seed the k-means clustering algorithm. K-means++ increases the likelihood of finding good clusters, in other words, clusters where the within-cluster variance is small and the between-cluster variance is large. The hope is that, if there is some natural grouping of the data in euclidean space, the k-means algorithm will discover it.

There are a couple of things to be wary of when using k-means. Namely, choosing features and dimensionality of the features, and choosing k,the number of clusters. For choosing features, it is helpful to have domain knowledge to know what characteristics[4] of the data points will be useful in distinguishing their simmilarity and dissimilarity. Fortunately, I used only precipitation data, which is relatively simple to understand, so specialized domain knowledge was not necessary. For choosing dimensionality, it is important to remember that k-means uses euclidean distance, so it works best in low dimensions, and less well in high dimensions. In very high dimensions, euclidean distance is not useful, because there is little notable difference in distances between pairs of points. Keeping this in mind, I used 13 features[5]. The features I used were mean annual rainfall, and mean rainfall for each month. These 13 features were calculated for each of the 378 geographic points in the precipitation data set.

There are various methods to choose k, the number of clusters to use in k-means. These methods can be based on domain knowledge, or quantifiable measures of goodness[6]. I was most interested in whether or not precipitation clusters were correlated with ethnic boundaries, so I used domain knowledge about Ethiopia to choose k. Choosing k was based on administrative regions, which, as mentioned earlier, correspond well to the location of ethnic groups. There are 11 administrative regions in Ethiopia, but two are chartered cities: Addis Ababa and Dire Dawa. The remaining 9 are based on ethnic territoriality. Of these 9, one is as small as the chartered cities, Harari, and three others are also quite small: Gambela, Benishangul-Gumaz, and Tigray. However,

[4]It should be noted that categorical and binary features can also be used in k-means, but only numerical features were used in this analysis.

[5]I considered decreasing the number of features, however, because the clustering was successful I left all 13 features.

[6]Goodness of clusters depends largely on the specific problem and data, so these quantifiable measures should be seen as useful information, not the ultimate way to determine k

from looking at the precipitation data, it was clear that Tigray had a distinct rainfall pattern from the surrounding regions, so I decided to include it in choosing k. This leaves 6 medium to large sized administrative regions. The largest region, Oromia, spans both the highlands and the lowlands. I noticed that it had at least two distinct rainfall patterns, so I added one more cluster, and completed the analysis with k = 7 clusters.

Though I chose k based on domain knowledge, I wanted to find the quantifiably 'best' number of clusters to make sure 7 was a reasonable choice. I used two methods to quantify the goodness of various numbers of clusters. The first method was the 'Elbow Method'. This method tracks the within-cluster sum of squares error (SSE) for various k. The idea is that increasing the number of clusters in the beginning will reduce the SSE substantially, and as k increases further reductions in SSE will be less drastic. In looking at the plot of SSE vs. cluster number, an 'elbow', a more acute angle, should be apparent. This is the optimal number of clusters, where an additional cluster reduces the SSE less substantially. However, as mentioned before, this method to measure the goodness of k should be taken as useful information, not as the definitive answer for choosing k. Using the Elbow Method, it is not obvious exactly which number of clusters is best, but it looks to be between 3 and 5 clusters.

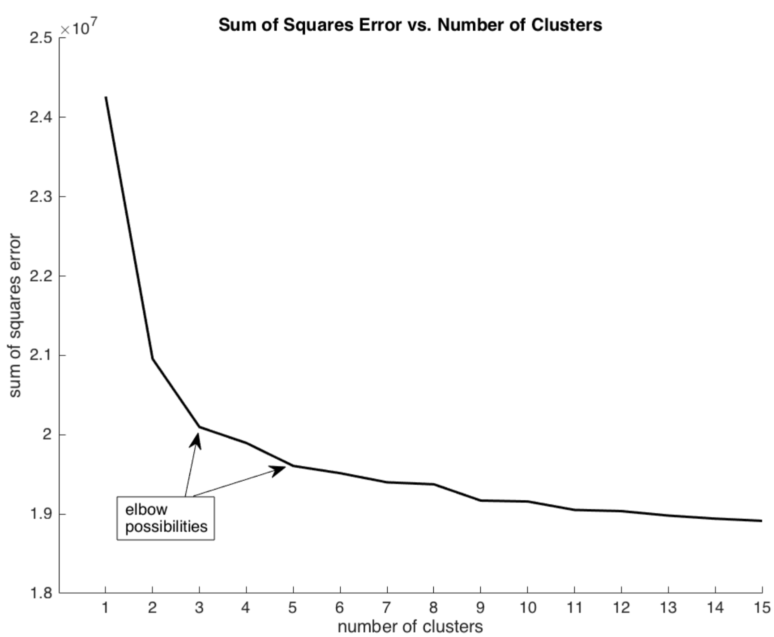


Figure 2.1: Elbow Method for determining the optimal number of clusters.

This is a shortcoming of the Elbow Method. Since the optimal number is determined visually[7], it is not always obvious what the optimal number of clusters is. Because of this, other quantitative methods of determining k have been developed. One such method is the Gap Statistic developed by Tibshirani, Walther and Hastie (2001). The Gap Statistic is similar to the Elbow Method in that it uses SSE as a measure of the variance explained by the clusters. However, the Gap Statistic compares the variance within the clusters of the data set to the variance within the clusters of a random uniformly distributed data set over the same region. This means that you compare the goodness of the clusters found in your data to the goodness of another set of clusters formed in data that should not contain any actual clusters. See figure for an illustration.

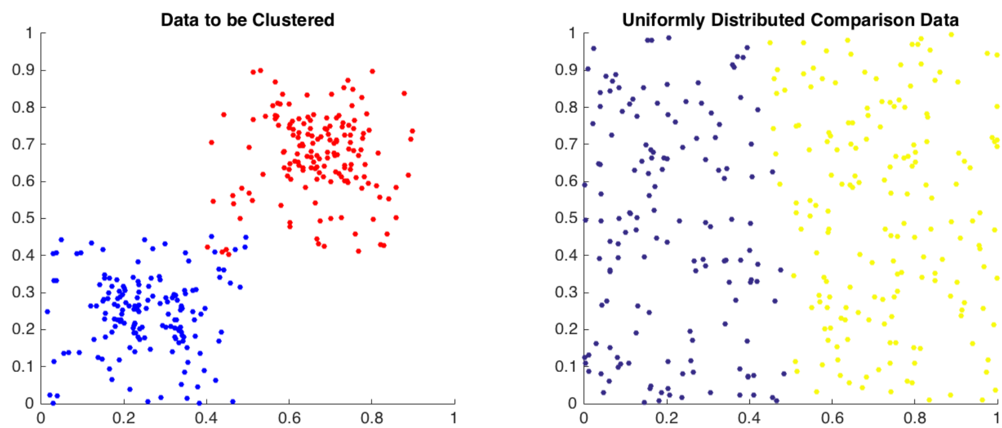


Figure 2.2: Gap Statistic. Example of a dataset to be clustered (left) and a uniformly distributed data set over the same region for comparison (right).

The Gap Statistic is defined below

$$Gap_n(k) = E_n\{logW_k^*\} - logW_k$$

Where W_k is the sum over all clusters of the normalized within-cluster SSE in the data set that you want to cluster, and W_k^* is the same quantity for the uniformly distributed comparison data set. n is the number of comparison data sets that are generated. See Appendix A for expansion of W_k.

The optimal k should be the k for which the Gap Statistic is largest. More concretely, the best k should be the smallest k such that $Gap(k) - (Gap(k+1) - s_{k+1}) \geq 0$. Although sometimes 1 will be the best k by this definition, and in that case the next smallest value for which the preceeding is true should be chosen. Where s_k is the combination of the standard deviation of the $log(W_K^*$ and the simulation error, which

[7] One could mathematically calculate the angle in the curve at each k instead of finding the elbow visually

becomes smaller the more comparison data sets are generated. With guidance from The Data Science Lab's blog post 'Finding the K in K-Means Clustering', I computed the Gap Statistic for the precipitation data. I tested k = 1 through 10, and generated 20 comparison datasets for each k. Using this method I found that 4 was the optimal number of clusters.

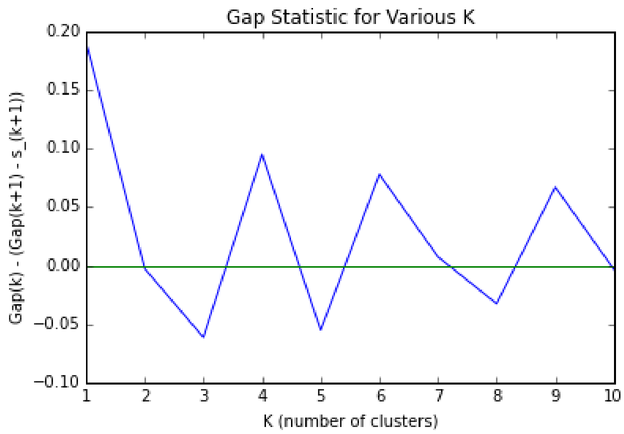


Figure 2.3: Gap Statistic. The best k should be the smallest k for which $Gap(k) - (Gap(k+1) - s_{k+1}) >= 0$

These quantitative measures selected 4 as the optimal number of clusters. As 7 is not much larger than 4, this verified that 7 was a quantifiably reasonable number of clusters for this data set. The k-means algorithm is non-deterministic. Though it will always converge, it will often find a local minima, rather than the global minima. Therefore, to find the tightest clusters at the global minima, many iterations of the algorithm must be run. I did 1000 iterations of k-means with 7 clusters to find the set of clusters with the lowest SSE. However, I noticed that this set of clusters resulted in one cluster that was not geographically cohesive. In the set of clusters with the next lowest SSE, all the clusters were geographically cohesive. Because the motivation for this analysis is largely based on the geographic cohesiveness of the ethnic groups, I chose to use this slightly less-tight group of clusters for the remainder of the analysis. See Appendix B for the non-cohesive cluster and the SSE difference between the two sets of clusters.

As mentioned above, I hoped that there would be some correspondence between the clusters and the locations of Ethiopia's ethnic groups. After clustering the data, it turned out that there is indeed such a correspondence. Many of the clusters map quite well onto the administrative regions, which means they map well onto the locations of ethnic groups.

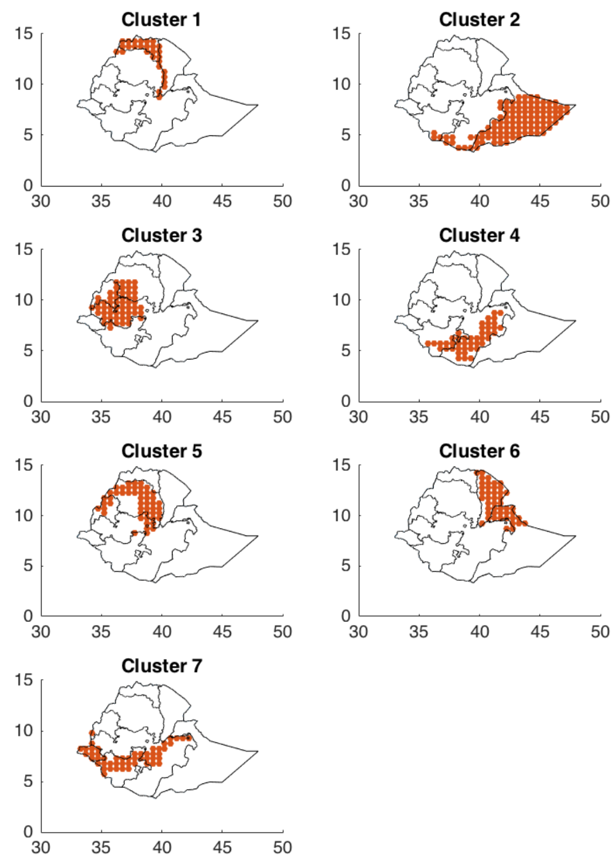


Figure 2.4: Clusters based on precipitation patterns. Administrative regions outlined in black.

Cluster 1 is mostly in the Tigray region, cluster 2 in Somali, cluster 5 in Amhara, and cluster 6 in Afar. Clusters 4 is mostly in southeastern Oromia, and clusters 3 and 7 do not correspond as tightly to administrative regions.

To find out how likely it was that this correspondence between clusters and administrative regions occured purely by chance, I created a simulation. In this null hypothesis simulation I chose a rectangle of points randomly in Ethiopia. The height and width of the rectangle were chosen from a uniform distribution, the bounds of which were chosen so that the mean rectangle area would be the same as the mean cluster area. I also made sure that the rectangle was at least as large as the smallest cluster. I then calculated what proportion of the rectangle was contained inside each administrative region, and took the maximum proportion. I then compared this to the proportion of the cluster that was most contained within that same region[8] I ran this simulation 10,000 times and only 15 rectangles, .0015, were as contained within administrative regions as the clusters used for the analysis. Though this is not a perfect test of the null hypothesis, it seems extremely unlikely that this level of correspondence between clusters and administrative regions would occur purely by chance. Also, it is worth noting that many of the clusters have a few points which are technically outside the respective administrative region, but are quite close to the boundary and still map impressively well to the shape of the administrative regions. See Appendix C for example rectangles and further explanation.

It is somewhat remarkable that the clusters based purely on precipitation patterns correspond so well to the administrative regions. It does make sense that the clusters are, in general, geographically cohesive, because locations near eachother will have the same or similar rainfall. However, this level of similarity to the administrative regions was not expected. However, this strong correspondence enables this analysis to pertain directly to the different ethnic groups who reside largely within various clusters. Above average rainfall in cluster 6 means above average rainfall for the Afar people. In the next section, the relationship between ENSO, and the precipitation anomalies in each cluster will be explored.

2.3 Precipitation and ENSO

It is generally accepted that it is dryer and hotter in the tropics during El Niño years (Hsiang et al., 2011). However, it has also been noted that the correlation between ENSO and precipitation in some parts of the tropics is the opposite, making it wetter than normal in El Niño years (Beltrando and Camberlin, 1993). Now that Ethiopia has been divided into distinct clusters, the relationship between precipitation anomalies in each cluster and ENSO will be explored. Only data in the *belg* and *kiremt* rain seasons will be analyzed. The reason being, that precipitation anomalies during the growing season are expected to have the biggest impact on crop yeilds and, in pastoral regions, the persistance of good grazing land. I will look at both rain seasons for each cluster. However, for the purposes of interpreting the results I first figure out which rain season is dominant for each cluster. Below, the primary rainy season of each cluster is shown.

Kiremt is the primary rainy season for clusters 1, 3, 5, 6, and 7, and *belg* is the primary rainy season for clusters 2 and 4. This agrees with the division USAID determined between *belg* dominant and *kiremt* dominant parts of the country (2015).

[8] I also controlled somewhat for size per administrative region. For example, since Somali is much larger than Tigray, for a rectangle to tested for Somali it had to contain at least 70 points, whereas for a rectangle to be tested for Tigray it needed to contain only 36 points. For reference, the cluster that is 93% within Somali contains 91 points.

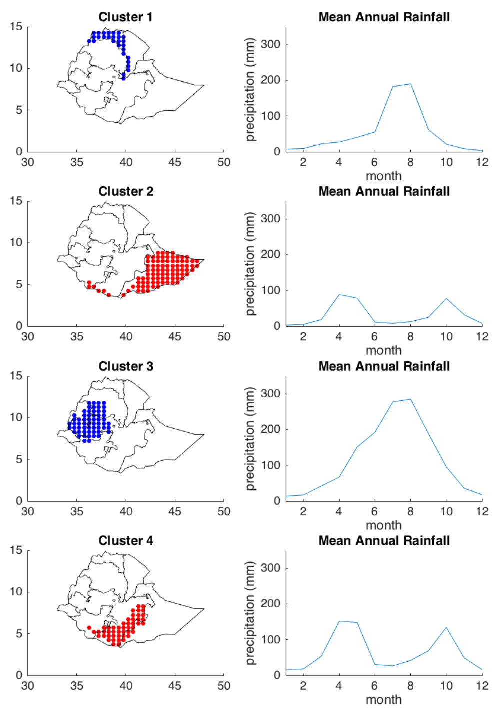


Figure 2.5: Clusters 1 through 4. *Belg* dominant clusters in blue, *kiremt* dominant clusters in red.

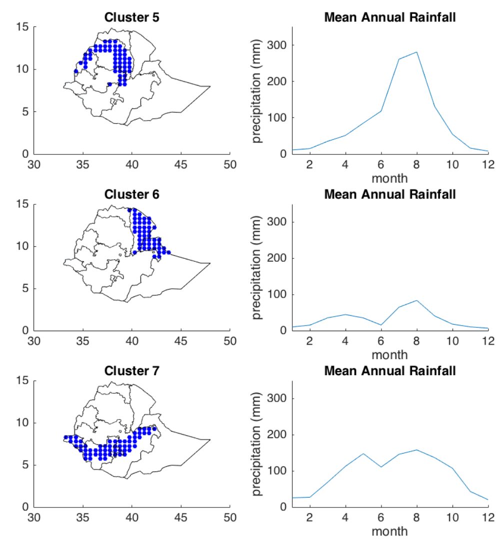


Figure 2.6: Clusters 5 through 7, all *kiremt* dominant.

Next, single linear regressions of the precipitation anomalies on the ENSO 3 index were computed. Regressions for both *belg* (March-May) and *kiremt* (July-September) 3-month periods were computed for each cluster. The regressions were done on the monthly data for ENSO 3 and precipitation, and then a second set of regressions were done using the mean value over the three-month period of each year. For the regressions on *kiremt* anomalies, there was a perfect split between *kiremt* dominant and *belg* dominant clusters: all *kiremt* dominant clusters were negatively correlated, and all *belg* clusters were positively correlated with ENSO 3. Meaning that during the *kiremt* season in an El Niño year, the *kiremt*-dominant northwestern half of the country should expect below normal precipitation, and the *belg*-dominant southeastern half of the country should expect above normal precipitation. The coefficients of these regressions are visualized below. The plot on the right shows which clusters were positively and negatively correlated.

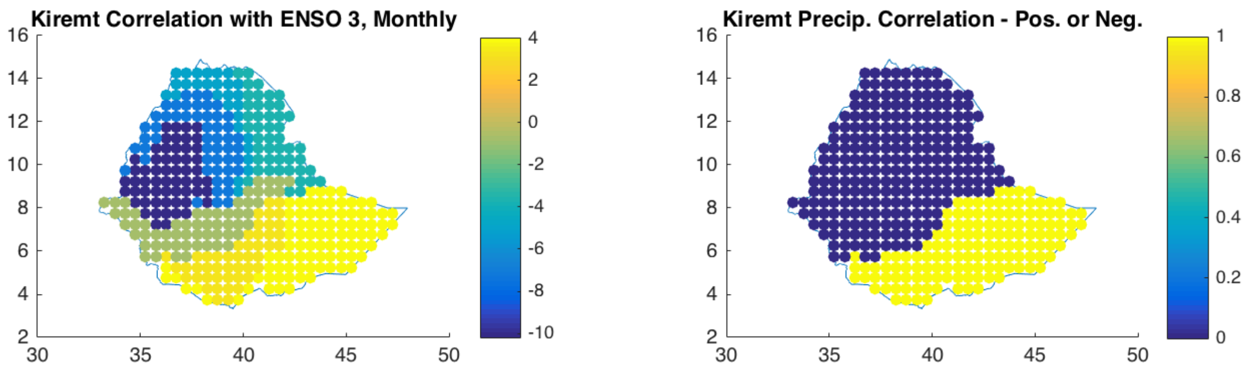


Figure 2.7: Coefficients precipitation anomalies and ENSO 3 during *Kiremt* for each cluster - monthly regression.

Only the results for cluster 1 were significant at a 90% level. However, clusters 2, 3, 4, and 5 are significant at an 80% level. Below is a table of the intercepts, coefficients, p-values, and r-squared from the monthly regressions of *kiremt* precipitation anomalies on ENSO 3. In the regression for the average precipitation anomaly and ENSO 3 values over the 3 *kiremt* months, the signs of the coefficients were the same. However, the results were less significant. See Appendix D for the map and table of results.

cluster	intercept	coefficient	p-value	r-squared
1	-0.0044291	-5.7993*	0.070424	0.10173
2	-0.39469	2.7336	0.36474	0.02657
3	1.2564	-9.8377	0.12752	0.073303
4	-2.1663	1.7742	0.69608	0.0049896
5	-5.3848	-8.0837	0.12643	0.073709
6	0.92847	-4.6125	0.12244	0.075224
7	0.52276	-2.3356	0.69103	0.0051655

*Indicates significance at 90%, **indicates significance at 95%.

The regressions on the *belg* season produced slightly different results. Also, the signs of the coefficients changed depending on whether the regression was done on each month during *belg* or on the average over the three months. In both, the two *belg* dominant clusters in the southeast were positively correlated with ENSO 3. *Kiremt* dominant clusters 1, 6, and 7 were also positively correlated with ENSO in the regressions done both monthly and averaged over the three months. *Kiremt* dominant cluster 5 was positively correlated in the monthly regression, but negatively correlated in the three-month averaged regression. Neither results for cluster 5 were close to being significant. The coefficients of the monthly and averaged regressions are visualized below.

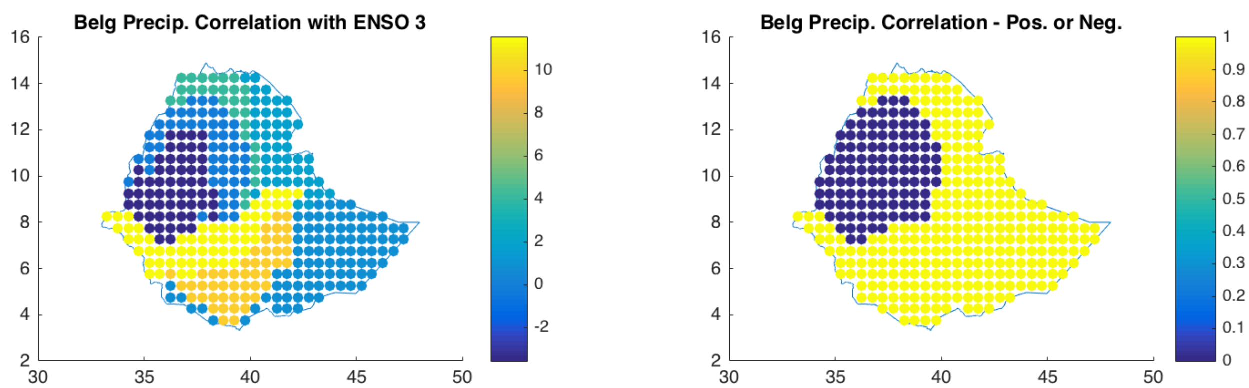


Figure 2.8: Coefficients precipitation anomalies and ENSO 3 during *belg* for each cluster - averaged regression.

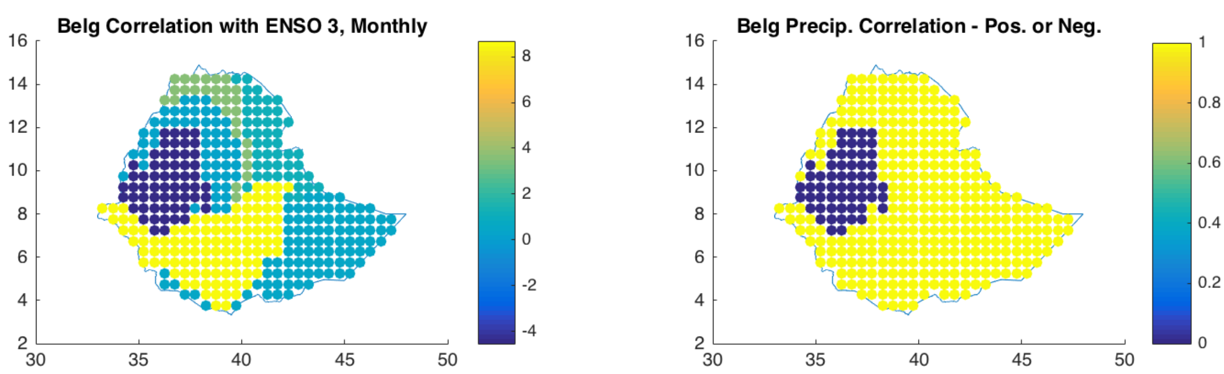


Figure 2.9: Coefficients precipitation anomalies and ENSO 3 during *kiremt* for each cluster - monthly regression.

Only the results for cluster 7 were significant. They were significant at a 90% level for the monthly regression, and a 95% level for the averaged regression. See table with the intercepts, coefficients, p-values, and r-squared from the monthly and averaged

cluster	intercept	coefficient	p-value	r-squared
1	-1.1371	4.059	0.19488	0.053592
2	-1.313	0.72783	0.86004	0.0010187
3	-0.59742	-3.5546	0.61882	0.0080818
4	-3.851	9.733	0.15832	0.063157
5	-1.0444	-0.13689	0.98119	1.82E-05
6	0.73888	1.7983	0.5758	0.010211
7	0.35131	11.535**	0.04956	0.11873

cluster	intercept	coefficient	p-value	r-squared
1	-1.1295	3.6979	0.13616	0.02275
2	-1.3073	0.45645	0.91163	0.00012762
3	-0.57663	-4.5347	0.47991	0.0051583
4	-3.828	8.6472	0.1412	0.022184
5	-1.0541	0.31963	0.94381	5.15E-05
6	0.75409	1.0814	0.71709	0.0013597
7	0.4158	8.4949*	0.088804	0.029563

regressions below.

The averaged regression results (left), and the monthly regression results (right). *Indicates significance at 90%, **indicates significance at 95%.

Better understanding of the relationship between ENSO and regional precipitation anomalies would better equip the Ethiopian government, and organizations such as the Famine Early Warning System to anticipate droughts and floods better. In the next section, the relationship between precipitation and conflict, as analyzed in this project, will be covered. An understanding of this second relationship exands the applications of the ENSO and precipitation relationship; because understanding both relationships would also improve our ability to predict the rise of potential conflicts.

2.4 Precipitation and Conflict

As mentioned previously, conflict data was gathered from the Social Conflict and Analysis Database (Salehyan et al., 2012). Only conflicts between groups in Ethiopia, and between the government and various groups that resulted in 5 or more deaths were included in this analysis. Nine such conflicts started[9] between 1970 and 2014. Based on the location at which the conflicts occured, they were sorted into each of the 7 clusters. This was done so that each conflict could be studied in relationship to the precipitation anomalies in the region in which it took place. There was one conflict which could have been sorted into either cluster 6 or 7. The analysis was performed with the cluster in each region. The goal in this part of the analysis was to discover whether or not there seems to be a relationship between precipitation anomalies and the conflict on a regional scale.

First, the average precipitation anomaly over the three month rain seasons of *kiremt* and *belg* and the start dates of conflict were plotted for each cluster.

[9]The start date of each conflict was used for the analysis.

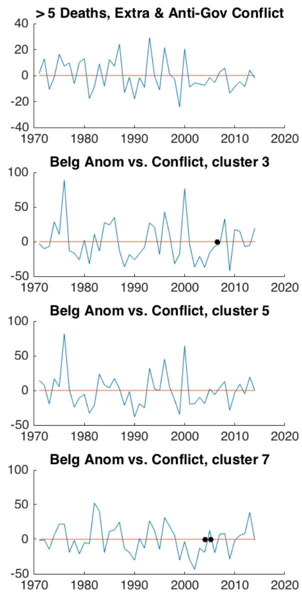


Figure 2.10: *Belg* precipitation anomalies (blue) and conflict start dates (black) for each cluster.

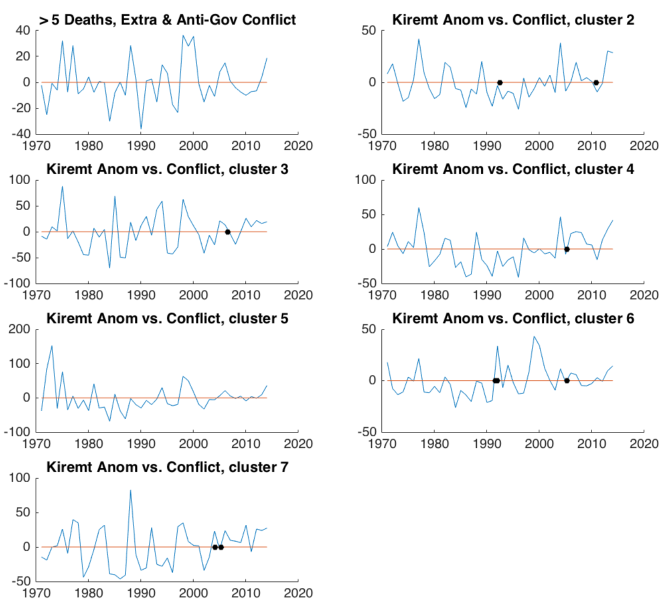


Figure 2.11: *Kiremt* precipitation anomalies (blue) and conflict start dates (black) for each cluster.

From a quick look at the plot of the *belg* anomalies, it seems that there is often below normal precipitation in the season preceding conflict. Next, *belg* season anomalies for the five years preceding each conflict were composited, so that the pre-conflict trend in each cluster could be more easily observed. Thin lines in each plot represent the 5-year *belg* anomalies preceding individual conflicts. Thick lines in each plot represent the mean 5-year pre-conflict *belg* anomalies for all of the conflicts for the cluster. There were no conflicts in some clusters. In the bottom right subplot, the mean 5-year pre-conflict *belg* anomalies for all conflicts in all clusters is plotted in blue. The same was done for kiremt[10].

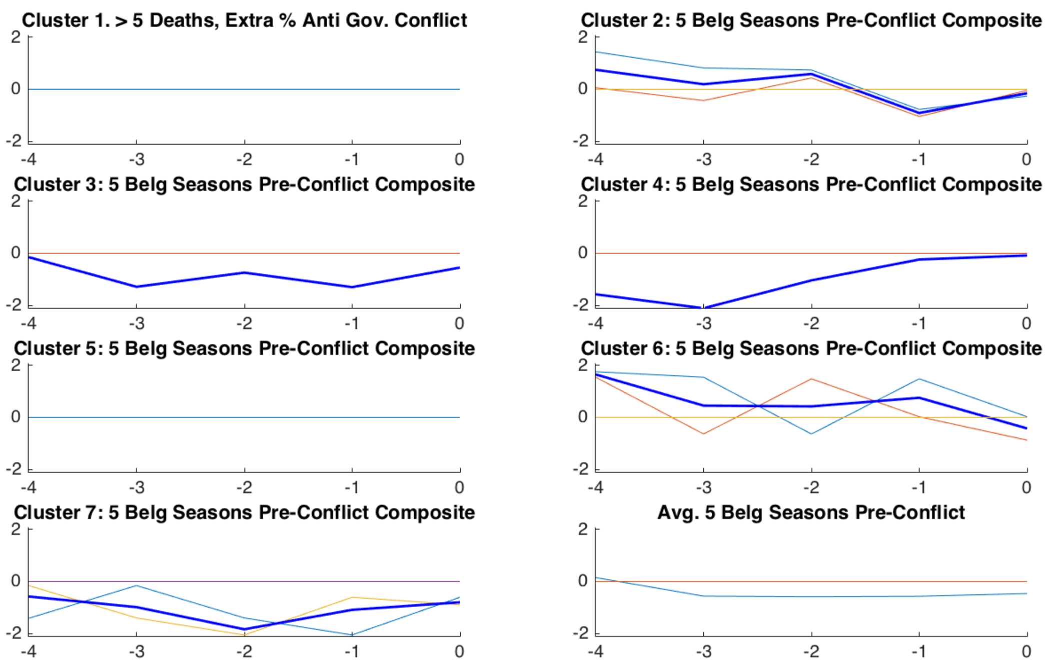


Figure 2.12: Composite 5 *belg* seasons preceding conflict for each cluster.

[10]These plots placed the conflict that could have belonged to cluster 6 or 7 in cluster 7. Placing the conflict in cluster 6 did not change the results drastically. The plots where the conflict was placed in cluster 6 can be found in Appendix E

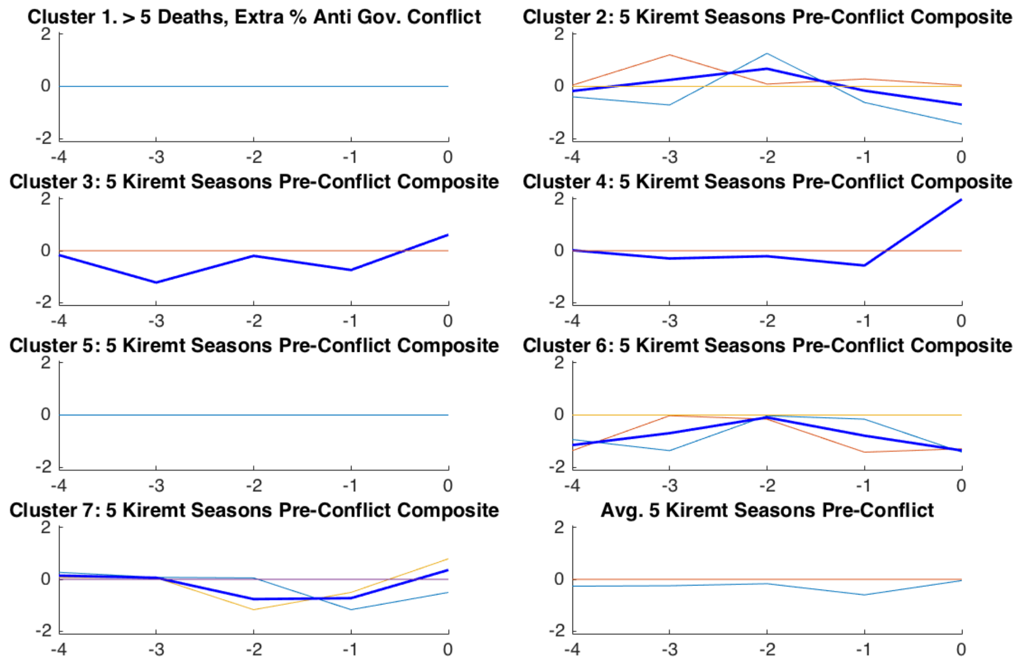


Figure 2.13: Composite 5 *kiremt* seasons preceding conflict for each cluster.

It seems that there is a trend of below normal precipitation in both *belg* and *kiremt* seasons preceding conflict. For *belg*, there is a four year trend of low precipitation preceding conflict. For *kiremt*, the season immediately preceding the conflict seems to have normal rainfall, but the four seasons preceding that have low precipitation, with a more acute defecit two seasons pre-conflict. This evidence supports the hypothesis that conflict is more likely to happen after years of below-normal precipitation during the growing season.

It is also worth looking separately at the *belg* dominant and *kiremt* dominant clusters above. Clsuters 2 and 4 are *belg* dominant, and the remaining clusters are *kiremt* dominant. *Belg* dominant cluster 4 has below-normal precipitation during both rainy seasons preceding the one conflict in the region. However, the defecit is more acute in the *belg* season. Though this is only one data point, it may indicate that precipitation deficits in the dominant rain season increase the likelihood of future conflict moreso than deficits in the non-dominant rain season. A distinction between *belg* and *kiremt* anomalies for cluster 2 is less clear. There is a deficit in both rain seasons for the two years preceding conflicts in the region. For the *kiremt* dominant clusters, cluster 5 has defecits in both rainy seasons before the one conflict in the region, with a larger deficit in the *belg* season. The same trend appears for cluster 7. However, the trend in cluster 6 is a pre-conflict deficit in the *kiremt* season and a pre-conflict surplus (excluding the season immediatly before the conflict) in the *belg* season.

Chapter 3

Results and Discussion

The success of the clustering on precipitation data was strong support for the regional approach of this analysis. Precipitation patterns mapped impressively closely to ethnic territories. This has interesting implications when combined with the relationships observed between regional precipitation and ENSO. ENSO is positively correlated with half the country, and negatively correlated with the other half during the *kiremt* season. In the belg season, all of the country besides the east-central highlands (clusters 3 and perhaps 5) is positively correlated with ENSO. This means that in La Niña and El Niño years, some regions, and therefore some ethnic groups, will likely have anomalously high rainfall, and other ethnic groups will experience anomalously low rainfall. As discussed in the introduction, both above normal and and below normal rainfall can be detrimental to agriculturalists and pastoralists in Ethiopia. Almost all Ethiopians fall into one of these two categories, so rainfall anomalies have the potential to make people extremely vulnerable. This by itself may increase the likelihood of conflict, because vulnerable people are desperate and may resort to violent conflict. Or, because non-vulnerable groups may choose an opportune time to instigate conflict when the other group is vulnerable. In addition to this, the predicted disparity in rainfall anomalies between regions might create or exacerbate the difference in well-being between groups. Investigating the relationship between precipitation disparity in regions and conflict between regions would be an interesting avenue to explore further. I looked into this relationship briefly, but, I suspect as a result of some faulty assumptions and methodology I was using at the time, I did not observe anything of note.

The existence of a relationship between precipitation and conflict has been observed before by Hsiang et al. (2011) and Theisen et al. (2013) among others. However, I have not come across such a small-scale study on precipitation and conflict. The analysis above showed a trend of regional below-normal rainfall in both rainy seasons in the few years preceding conflict in that same region. Only nine conflicts were involved in this investigation. It would be interesting to do a study on a wider array of conflicts within Ethiopia, or perhaps on the same type of conflict over a broader spatial scale, to see how the results may or may not vary. Additionally, it would be interesting to take a more in depth look at each of the conflicts, determine the parties involved, and look at the precipitation anomalies where both parties reside. This result is promising. However, the sample size, geographic scale, and time scale are small. In addition to this, Ethiopia's fascinating precipitation and ethnic territory correspondence, as well as ENSO's influence on regional precipitation, likely do not exist in many other countries. As such, the results of this project should not be extrapolated to other countries.

I regret that I did not include the present drought and conflict in my analysis. The precipitation data from CRU does not cover 2014 to the present. However, there are other precipitation data sets, and it would be interesting to see how well the current drought and conflict fit the trends observed here. The current drought in the northeast of the country certainly supports the results of the precipitation on ENSO regressions. These regressions predicted below average rainfall in the northeast *kiremt* season during El Niño events, and this is precisely the current situation. From the results of the regression, we would also expect to see below normal precipitation in the northwest of the country this year. However, the drought is less severe in the northwest. This may have to do with differences in climate between the northwest and the northeast. The northeast has a much dryer climate, and small rainfall deficits here are more likely to have serious negative effects. The northwest, by contrast, has a wetter climate, and small rainfall deficits are likely not as detrimental.

One more avenue that I think may be worth exploring, is the relationship between climate and the founding of rebel groups. In my reading about recent extra-government and anti-government conflicts, it seemed that often the violence was initiated by a rebel group formed in the previous three to five years. It seems that the formation of rebel groups may be a signal of discontent, increasing dissident organization, and therefore of possible conflict. Thus it would be interesting to study if either weather or climate might be related to the inception of rebel groups.

3.1 Conclusion

The initial motivation for this project was to investigate the relationship between Ethiopian conflict and weather and climate indicators such as temperature, precipitation, evapotranspiration, sea surface temperature, etc. Perhaps the most surprising, and informative, part of this project was the strong spatial correspondence between precipitation patterns and ethnic territories. This correspondence was discovered early in the process. Because of this surprising result, and because a large majority of Ethiopia's population comprises farmers and pastoralists who are dependent on rainfall, precipitation was chosen as the weather indicator to focus on. In this project, precipitation was the mode by which climate could be linked to conflict. If there is a relationship between climate, via ENSO, and precipitation, and a relationship between precipitation and conflict, then there is a relationship between climate and conflict. A relationship between ENSO and Ethiopian precipitation was observed. The *kiremt* rains are negatively correlated with ENSO in the northwest half of the country, and positively correlated with ENSO in the southeast. The *belg* rains are negatively correlated with ENSO only in the east-central highlands, and are positively correlated with ENSO in the rest of the country. Additionally, it seems there may also be a connection between regional precipitation anomalies and conflict: conflict is often preceded by a few years of below normal rainfall in both rainy seasons. The strength of the relationship between both ENSO and regional precipitation, and between precipitation and conflict could be tested more thoroughly. However, I hope that this analysis has demonstrated a novel, regionally-based approach to the topic of climate and conflict. The current famine and violent conflict in Ethiopia is proof that we still have much to do in terms of our ability to anticipate precipitation anomalies and conflict. As there seems to be no more debate on the existence of climate change, understanding the effects climate change, including effects on precipitation and conflict is critical.

Appendix A

Gap Statistic

The Gap Statistic.

D_k is the within-cluster sum of squares error (SSE) for cluster k:

$$D_k = 2n_k \sum_{x_i \in C_k} ||x_i = \mu_k||^2$$

Where C_k is cluster k, x_i is a data point assigned to cluster k, and μ_k is the cluster center.

W_k is the sum over all clusters of the normalized within-cluster SSE. It is a measure of variance of all the clusters.

$$W_k = \sum_{k=1}^{K} \frac{1}{2n_k} D_k$$

Where K is the number of clusters and n_k is the number of data points assigned to cluster k. The Gap Statistic is defined as:

$$Gap_n(k) = E_n\{logW_k^*\} - logW_k$$

Where W_k is the sum over all clusters of the normalized within-cluster SSE for the data set that you want to cluster, and W_k^* is the same quantity for the uniformly distributed comparison data set. n is the number of comparison data sets that are generated. The best k should be the smallest k such that $Gap(k) - (Gap(k+1) - s_{k+1}) \geq 0$. Where s_k is the combination of the standard deviation of the $log(W_K^*$ and the simulation error.

Let $sd(k)$ be the standard deviation for the uniformly distributed comparison dataset. Then $s_k = \sqrt{1 + 1/n} sd(k)$ is a measure of the standard deviation and the simulation error for simulating n comparison distributions.

Appendix B

Choosing Tight vs. Geographically Cohesive Clusters

I chose to use a set of clusters with a slightly higher SSE, because the set of clusters with the lowest SSE had a geographically non-cohesive cluster.

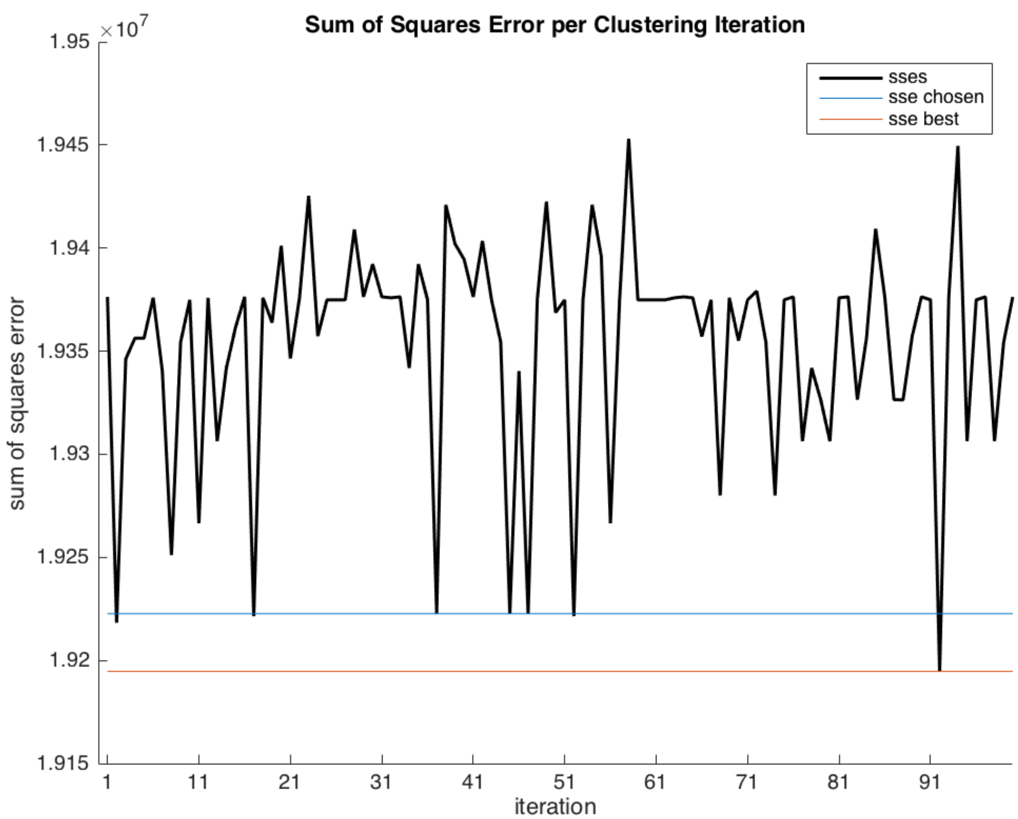


Figure B.1: SSE for 100 iterations of k-means clustering.

Here is the set of clusters with the lowest SSE. Cluster 4 is geographically non-cohesive.

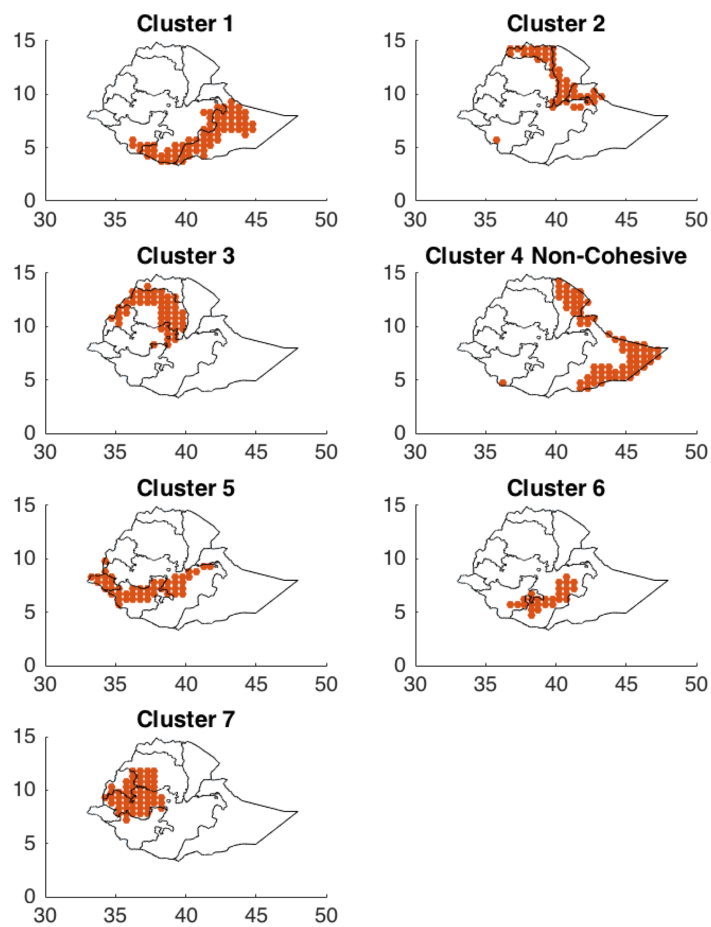


Figure B.2: Clusters with lowest SSE. Though I did 1000 iterations of the k-means algorith, this plot is of only 100 iterations to make it easier to see visualize.

Here is the set of clusters used in the analysis. The clusters are more geographically cohesive, and the SSE was only slightly larger than the best set of clusters.

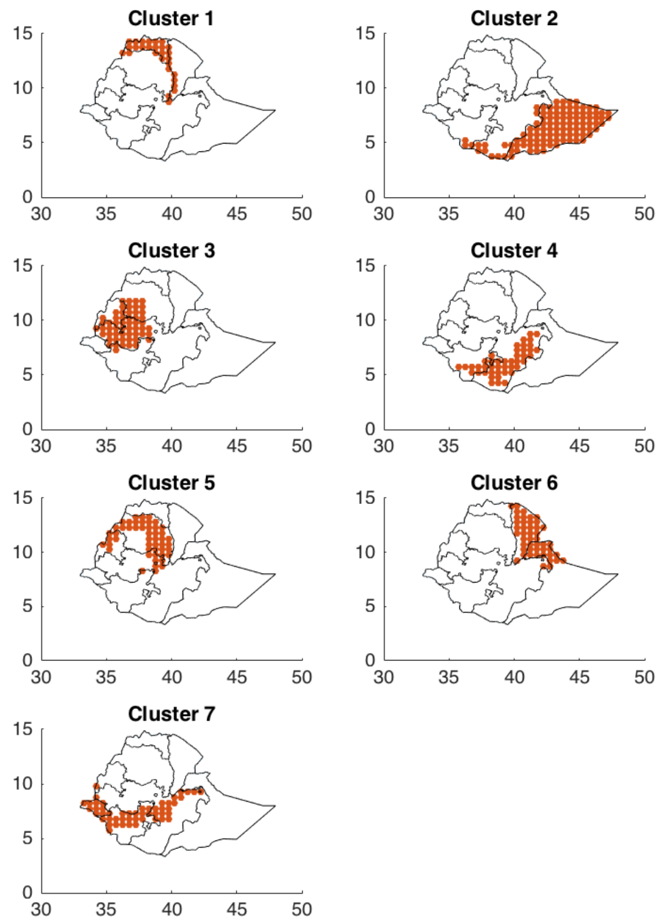
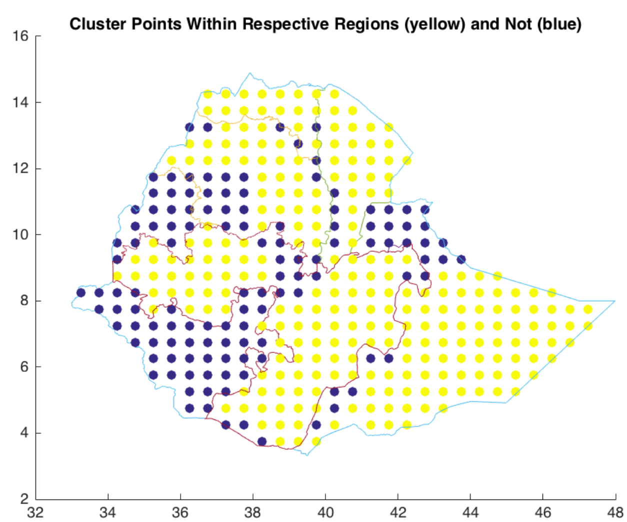


Figure B.3: Clusters used for analysis.

Appendix C

Testing the Null Hypothesis of Cluster to Administrative Region Correspondence

The points in the clusters used for analysis that were inside their respective administrative region (in yellow) and those that were outside their respective region. It is worth noting that many points outside the administrative regions are quite close to the administrative regions.



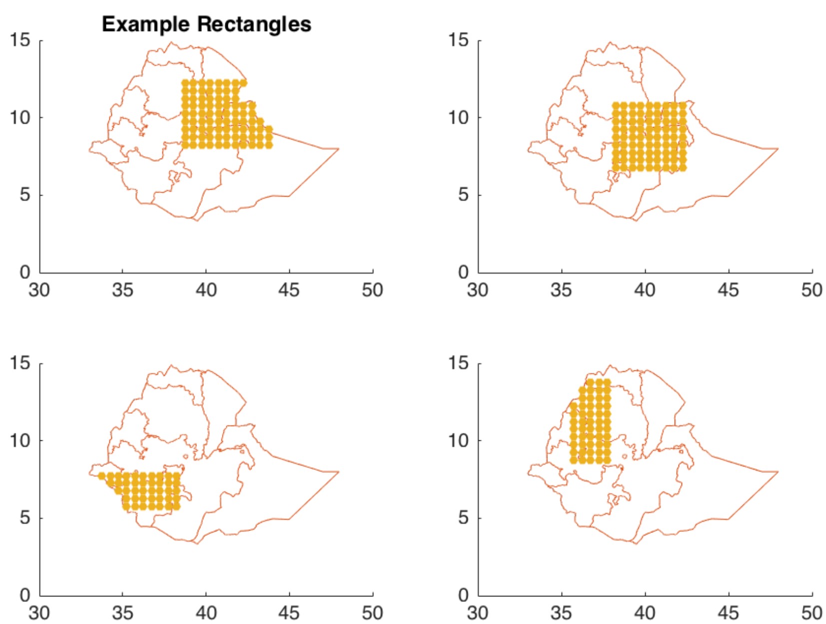
The percentage of points within the respective administrative region for each cluster:

```
for region 1: 0.53571 of the cluster points are inside the respective region
for region 2: 0.93333 of the cluster points are inside the respective region
for region 3: 0.55172 of the cluster points are inside the respective region
for region 4: 0.78723 of the cluster points are inside the respective region
for region 5: 0.64706 of the cluster points are inside the respective region
for region 6: 0.6087 of the cluster points are inside the respective region
for region 7: 0.43103 of the cluster points are inside the respective region
percentage of points in respective regions = 0.67196
```

The percentage of points within the respective administrative region for each cluster excluding 7, which did not correspond as well to a single administrative region:

```
for region 1: 0.53571 of the cluster points are inside the respective region
for region 2: 0.93333 of the cluster points are inside the respective region
for region 3: 0.55172 of the cluster points are inside the respective region
for region 4: 0.78723 of the cluster points are inside the respective region
for region 5: 0.64706 of the cluster points are inside the respective region
for region 6: 0.6087 of the cluster points are inside the respective region
percentage of points in respective regions = 0.71562
```

Examples of the rectangles generated to test the null hypothesis that clusters corresponded to administrative regions purely by chance:



Appendix D

Map and Table of Results from Precipitation ENSO 3 Regression

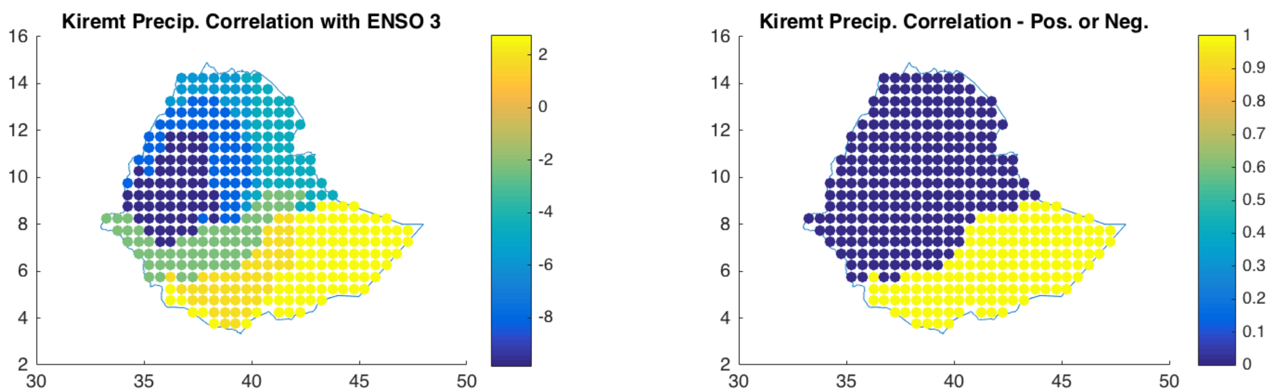


Figure D.1: Coefficients precipitation anomalies and ENSO 3 during *Kiremt* for each cluster - averaged regressions.

cluster	intercept	coefficient	p-value	r-squared
1	-0.01103	-5.0305 *	0.089581	0.029422
2	-0.40582	4.0295	0.17438	0.018936
3	1.2594	-10.187	0.10445	0.026948
4	-2.1794	3.3046	0.41045	0.0069959
5	-5.3919	-7.2558	0.10114	0.027464
6	0.92093	-3.7344	0.11779	0.025033
7	0.50884	-0.71519	0.87843	0.0002424

Appendix E

Composited Precipication Anomalies with Ambiguous Conflict Placed in Cluster 7

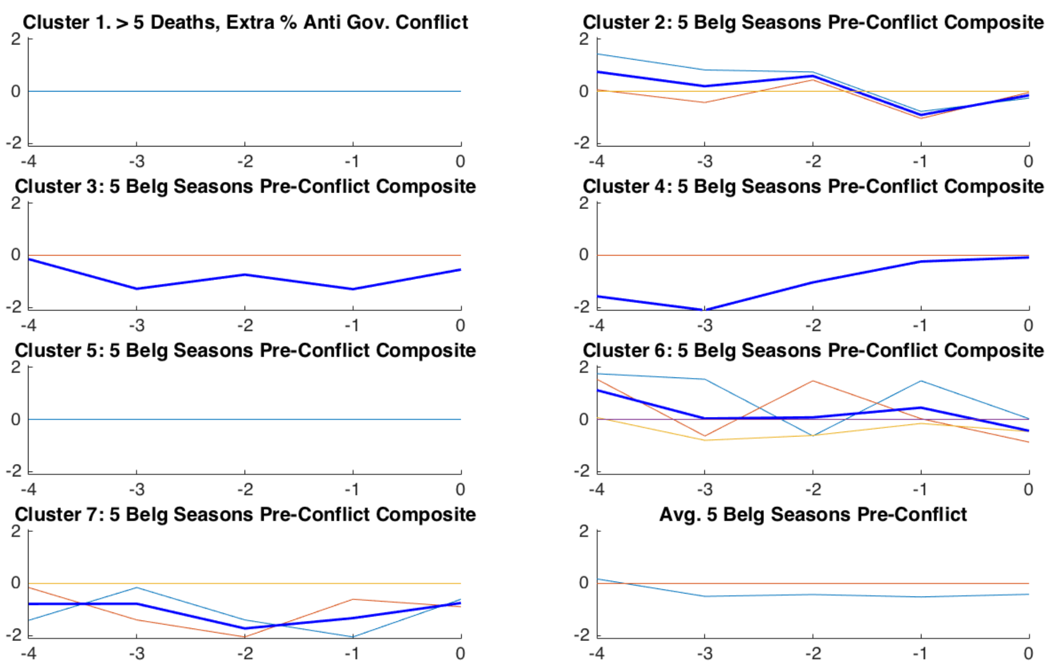


Figure E.1: Composite 5 *belg* seasons preceding conflict for each cluster.

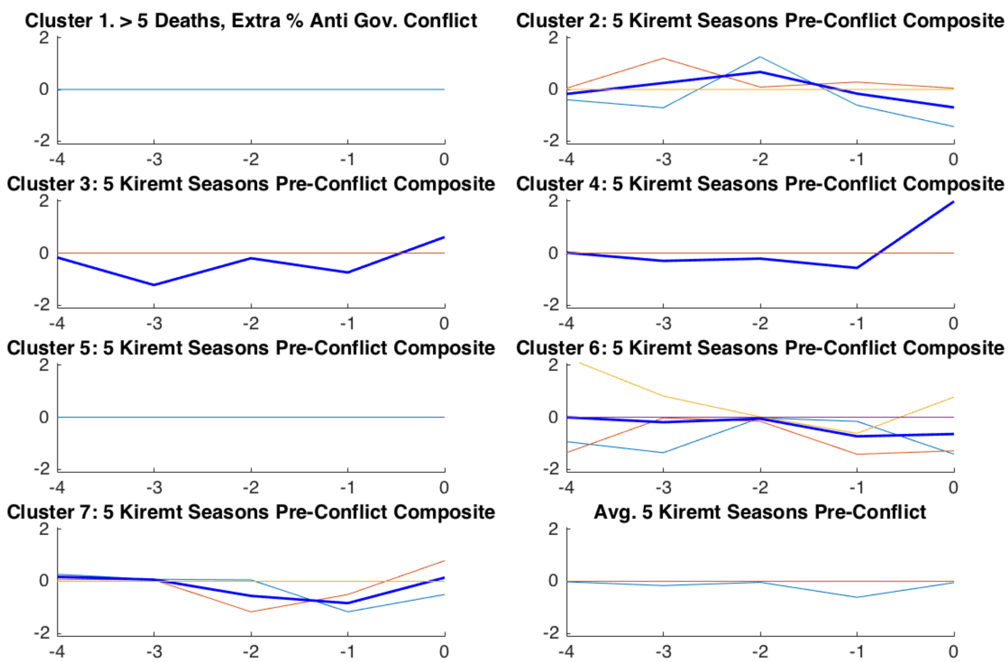


Figure E.2: Composite 5 *kiremt* seasons preceding conflict for each cluster.

Bibliography

Abbink, J. (1993). Famine, Gold and Guns: The Suri of Southwestern Ethiopia, 1985-91. *Disasters*, 17(3):218–225.

BBC (2016). Ethiopia profile - Timeline. http://www.bbc.com/news/world-africa-13351397. Accessed: 2016-03-28.

Beltrando, G. and Camberlin, P. (1993). Interannual Variability of Rainfall in the Eastern Horn of Africa and Indicators of Atmospheric Circulation. *International Journal of Climatology*, 13:533–546.

Central Statistical Agency of Ethiopia (2015). Table 2: Area, Production, and Yeild of Crops For Private Peasant Holdings for Meher Season 2014/2015.

Cheung, W. H., Senay, G. B., and Singh, A. (2008). Trends and spatial distribution of annual and seasonal rainfall in Ethiopia. *International Journal of Climatology*, 28:1723–1734.

C.I.A. (2016). The World Factbook: Ethiopia. https://www.cia.gov/library/publications/the-world-factbook/geos/et.html. Accessed: 2016-03-28.

Famine Early Warning Systems Network (2015). ETHIOPIA Food Security Alert.

Federal Democratic Republic of Ethiopia Population Census Commission (2008). Summary and Statistical Report of the 2007 Population and Housing Census. Technical report, Addis Ababa, Ethiopia.

Gleditsch, N. P. (2012). Whither the weather? Climate change and conflict. *Journal of Peace Research*, 41:3–9.

Government of Ethiopia (2016). Ethiopian Government Portal: Regional States. https://web.archive.org/web/20140802031634/http://www.ethiopia.gov.et/web/pages/regional-states. Accessed: 2016-03-28.

Guliye, A., Noor, I., Bebe, B., and Kosgey, I. (2007). Role of camels (Camelus dromedarius) in the traditional lifestyle of Somali pastoralists in northern Kenya. *Outlook on Agriculture*, 31(1):29–34.

Hsiang, S. M., Meng, K. C., and Cane, M. A. (2011). Civil conflicts are associated with the global climate. *Nature*, 476:438–441.

Human Rights Watch (2016). Ethiopia: No Let Up in Crackdown on Protests. https://www.hrw.org/news/2016/02/21/ethiopia-no-let-crackdown-protests. Accessed: 2016-03-29.

IRIN (2015). How Bad is the Drought in Ethiopia? http://newirin.irinnews.org/dataviz/2015/11/19/how-bad-is-the-drought-in-ethiopia. Accessed: 2016-03-29.

Jury, M. R. and Funk, C. (2013). Climatic trends over Ethiopia: regional signals and drivers. *International Journal of Climatology*, 33:1924–1935.

NOAA/ National Weather Service (2016). Monthly Atmospheric and SST Indices. http://www.cpc.ncep.noaa.gov/data/indices/.

Salehyan, Idean, Hendrix, C. S., Hamner, J., Case, C., Linebarger, C., Stull, E., and Williams, J. (2012). Social conflict in Africa: A new database. *International Interactions*, 38(4):503–511.

Theisen, O. M., Gleditsch, N. P., and Buhaug, H. (2013). Is climate change a driver of armed conflict? *Climate Change*, 117:613–625.

USAID (2015). Climate Variability and Change in Ethiopia: Summary of Findings. Technical report, Washington, DC.