UiO University of Oslo

Heating Up the Dry Debate

A sensitivity analysis of drought and communal conflict in Sub-Saharan Africa (1989-2014)

Fredrik Methi

Master's Thesis in
Peace and Conflict Studies

Department of Political Science
University of Oslo

Spring 2020

Word Count: 29,703

Abstract

Climate is getting hotter — both in fact and as a research topic. Yet there is little scientific consensus on the conflict potential carried by changes in climate and weather patterns. Previous studies investigating the links between drought and conflict have relied on precipitation-based measures of drought. However, in the same way that peace is not just the absence of war, drought is not just the absence of precipitation. What is needed is a greater focus on the relationship between the *impacts* of drought and conflict potential. In this thesis, I identify three contrasting drought indicators that are assumed to capture different theoretical concepts of drought: a precipitation-based measure (SPEI); a vegetation-based measure (NDVI); and a socio-economic measure (EM-DAT). These three measures are not only assumed to capture different theoretical concepts of drought, but also different stages in the drought cycle. Lack of precipitation may lead to less vegetation, loss of crops and deterioration of pasture, which in turn may spark a socio-economic disaster. By using these three measures, I answer the research question: *Do different conceptualisations of drought affect the likelihood of communal conflict?*

Drawing on novel high-resolution data on communal conflict events and droughts in Sub-Saharan Africa from 1989 to 2014, this thesis evaluates the relationship between drought and communal conflict on the local level. Results from mixed-effects multilevel logistic regression show that all three drought measures are associated with a higher risk of communal conflict, but the effect differs across the various measures. The higher the measurable physical impact of drought, or the closer we get to measuring the socio-economic impacts of the drought, the higher the risk of experiencing communal conflict.[1]

[1]Dataset and do-file used in this thesis can be downloaded from the following Dropbox folder: https://www.dropbox.com/sh/6wqi2xdooqo7nkz/AADpOzER0DZ5SZ0QL9QwaFc6a?dl=0

Acknowledgements

“Some people have a way with words, and other people ... not have a way”
- Steve Martin

Writing a thesis in a foreign language I certainly feel attached to the latter category. Yet here I am, handing in a paper containing 30,000 words in this foreign language. This whole project would not have been possible without the excellent teaching, guidance and advise of my supervisor, Andreas Forø Tollefsen. Thank you for always leaving the door and inbox open.

Next, I want to thank the rest of PRIO for letting me write my thesis in a highly inspiring environment. Special thanks to the CVP department and CLIMSEC/CROP for allowing me to present my work and giving thorough feedback. Also thanks to the other MA students at the Learners Loft for moral support.

Furthermore, I would like to thank all friends and family for helping me compile these hundred pages into something (hopefully) readable. Thanks to Sindre and Kristian for thorough feedback and keeping me entertained with numerous quizzes throughout these two years. Also thanks to Mikael for feedback and good discussions along the way.

Special thanks to mum and dad for letting me stay at home during the pandemic, feeding me with ideas and (more importantly) free food.

Most of all, I would like to thank Guro for her love and support, as well as for reading thoroughly through the paper and giving comments.

Fredrik Methi

Oslo, 22 June 2020

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Climate is getting hotter — both in fact and as a research topic. In the preceding years the relationship between climate and conflict has received broad attention. In 2007, Al Gore and the Intergovernmental Panel on Climate Change (IPCC) were awarded the Nobel Peace Prize on the basis that climate change may lead to "increased danger of violent conflicts and wars, within and between states" (Mjøs, 2007). Two years later, in his award speech for receiving the Nobel Peace Prize, Barack Obama stated that "[t]here is little scientific dispute that if we do nothing, we will face more drought, more famine, more mass displacement — all of which will fuel more conflict for decades" (Obama, 2009). Yet neither of these statements were built on solid scientific ground. For the last 15 years a number of researchers has tried to unveil the true relationship between climate variability and conflict, but results are diverging (Mach et al., 2019). In fact, results are so diverging that other scholars have felt the need to call for peace among climate-conflict researchers (Solow, 2013).

In order to better understand the causal mechanisms at play, I shed light on what I argue to be a problematic aspect in the existing literature: how to measure drought. Most research on the climate-conflict nexus focuses on the implications of drought as drought is the natural hazard commonly assumed to carry the largest conflict potential (von Uexkull, Croicu, Fjelde & Buhaug, 2016, p. 2). However, with few exceptions, former researchers use precipitation-based measures as proxies for drought (see e.g. Buhaug, 2010; Miguel, Satyanath and Sergenti, 2004; or von Uexkull, 2014). This is problematic as a drought is much more than the lack of rainfall. In the same way that peace is not just the absence of war, drought is not just the absence of precipitation. Hence, I argue that precipitation-based measures of drought lack operational validity as they are not always coherent with the researchers own theoretical claims. In order to overcome this challenge, I use three different measures of drought in order to answer the research question:

Do different conceptualisations of drought affect the likelihood of communal conflict?

These three different measures are The Standardized Precipitation Evapotranspiration Index (SPEI), the most commonly used drought index which is mainly based on precipitation anomalies; The Normalized Difference Vegetation Index (NDVI), a remotely-sensed measure taking into account the level and quality of vegetation; and The Emergency Events Database (EM-DAT), a database compiled of disaster data from various crisis reports. My main reason for using these three indicators of drought is twofold. First, these indicators may reflect three different theoretical definitions of drought. SPEI can be perceived as a measure of meteorological drought, NDVI as a measure of agricultural drought and EM-DAT as a measure of socio-economic drought. Second, these three measures may represent a sequence of drought, from lack of precipitation to lower vegetation quality to a socio-economic disaster, all playing a key role in the hypothesised link between drought and armed conflict.

To investigate the conflict potential carried by these drought indicators I use communal conflict incidences as the dependent variable. Communal conflicts are small-scaled, non-state conflicts and are commonly assumed to be particularly prone to climate variability (Butler & Gates, 2012; Fjelde & von Uexkull, 2012; Nordkvelle, Rustad & Salmivalli, 2017). These conflicts are particularly prominent in Sub-Saharan Africa (Brosché & Elfversson, 2012). Moreover, a large number of communal conflicts in Sub-Saharan Africa is between farmers and herders[2], which has been a prominent conflict type for decades. Since both farmers' and herders' livelihood rely on stability from nature, farmers rely on water for crops and herders on healthy pasture, these are expected to be particularly vulnerable to climatic shocks. For these reasons I limit the scope of this thesis to focus on Sub-Saharan

[2]In this thesis I use the term "farmer-herder conflict" as an umbrella term for all conflicts consisting of either farmers, pastoralists or both.

Africa. Due to the availability of data, I focus on the time period between 1989 and 2014.

In order to construct a convincing answer to the research question, I run a spatially disaggregated multilevel logistic regression analysis. To capture the local dynamics, I use grid cells as units to allow for variation in drought and conflict within countries and sub-national units. Since communal conflicts tend to be small-scaled conflicts often taking place close to the source of disagreement (Fjelde & Østby, 2014), using grid cells seems analytically favourable. Moreover, since data are nested in a panel-data structure, I apply multilevel logistic regression as this estimator is well suited to handle nested observations. In other words, contrary to regular logistic regression or ordinary least squares (OLS) regression, the multilevel model does not assume independence of all observations. This is favourable as it allows measuring both whether some areas are more prone to experience drought and conflict, and whether the droughts and conflicts also tend to happen within the same temporal unit. Thus, I do not only consider the variation between cells, but also within cells over time.

1.1 Relevance and Contribution

To the author's knowledge this is the first large-N study investigating the effect of drought on violent conflict measuring different theoretical definitions of drought. Existing studies have examined the impact of various drought indicators, but all indicators have captured the same theoretical version of drought, namely meteorological drought (see e.g. Theisen, Holtermann & Buhaug, 2011). Meteorological drought is commonly defined as a measure based on the lack of rainfall. However, it is not the lack of precipitation per se that carries conflict potential, it is rather the *impact* of the lack of precipitation. Therefore, this thesis attempts to get closer to the tipping point where a drought may spark a conflict, by looking at more concrete operationalisations of drought.

Based on Wilhite and Glantz (1985), I argue that SPEI can be used as a proxy for meteorological drought, NDVI as a proxy for agricultural drought and EM-DAT as a way of measuring socio-economic drought. As briefly touched upon, these three measures do not necessarily reflect three different concepts of drought, but can also be perceived as three different components in the causal chain of a drought. However, regardless of whether these indicators reflect three different concepts or three different components of the same concept, this would still mean that both NDVI and EM-DAT are closer to measuring the social impacts of a drought than SPEI. Ultimately, it is the social impacts of drought we expect to carry conflict potential.

Based on existing theory and former findings, I create two testable hypotheses. The first hypothesis suggests that there should be a positive relationship between all three drought indicators and the likelihood of communal conflict. Since all droughts are assumed to deteriorate living conditions of farmers and herders, the likelihood of violent conflict over scarce resources are assumed to be higher during and after droughts (Homer-Dixon, 1999). The second hypothesis states that the greater physical impact of the drought, the higher likelihood of experiencing conflict. Here I assume that measuring meteorological drought can be perceived as a measure containing a low degree of physical impact. Measuring the vegetation quality can be perceived as measuring the physical impact of a drought to a higher extent than solely relying on precipitation-based measures. And measuring the socio-economic consequences of a drought can be perceived as the measure containing the highest level of physical impact. The hypothesis therefore suggests that the closer we get to the tipping point where the impacts of drought are expected to contribute to conflict, the higher the likelihood of conflict. Hence, I expect SPEI to a have a low, albeit positive, effect on the likelihood of communal conflict; NDVI to have a medium effect; and EM-DAT to have a strong effect.

This study is an important contribution to the research frontier as it raises questions on elements largely neglected in previous research: what is a drought and how should we measure it? Thus, the aim of the study is to examine other operationalisations of drought, which in turn may be used to measure drought in future research. Moreover, the relationship between the different operationalisations of drought and conflict also reflect different implications.

1.2 Findings and Implications

My thesis offers several findings and implications concerning the relationship between drought and conflict. First, all three drought indicators are positive in the main model where I consider the direct, indirect and total effect of drought. Thus, they all point in the same direction that droughts are associated with a higher likelihood of seeing communal conflict across Sub-Saharan Africa. This supports the first hypothesis that all drought measures are associated with a higher likelihood of communal conflict.

Second, although the three drought indicators point in the same direction, the size of the coefficients vary in line with my second hypothesis. The SPEI variable is only significant when considering the lack of precipitation in neighbouring areas. When all neighbouring cells experience a SPEI drought, the odds of seeing communal conflict increase by 47% controlled for confounding factors. NDVI, on the other hand, is significant in all models and suggests that an NDVI drought increases the odds of experiencing conflict by 39-92% depending on whether the drought occurs in the main cell, neighbouring cells or both. EM-DAT obtains by far the strongest positive and significant coefficients, suggesting that a socio-economic drought doubles the odds of experiencing violent conflict. This supports the second hypothesis that the larger physical impact of the drought, the higher probability of experiencing a communal conflict. Moreover, this could also be interpreted as the closer we get to measuring the social impacts of a drought, the higher the conflict potential.

The third major finding in this thesis is that comparing vegetation quality (NDVI) across land types is problematic. Some areas consisting of cropland, shrubland and grassland have a higher variation in vegetation than areas covered by trees and rainforest. This raises a fundamental question: Do all land types have the same probability of experiencing drought? The answer to this question affects the computation of the NDVI cutoff. I argue that some areas are more prone to experiencing droughts, hence I calculate a percentage-based cutoff from the cell mean, instead of basing the cutoff on the standard deviations of each cell. This makes some areas more likely to experience NDVI droughts than others. However, this imposes a problematic aspect concerning the causal claims as only some areas are able to experience drought and these tend to be the same areas experiencing communal conflict.

To further investigate this, I run a subset regression analysis of the Sudano-Sahelian Zone and results show that two of the three NDVI estimates are negative, albeit not significant. This means the correlation between NDVI drought and communal conflict is based on spatial correlation (taking place in the same areas) rather than temporal correlation (taking place in the same years). Thus, the generalisability of the effect of NDVI drought on conflict should be carefully considered.

Nonetheless, the findings in this thesis are important. First they have scientific implications. Since there are differences between SPEI, NDVI and EM-DAT regarding the relationship between drought and communal conflict, they capture different aspects of a drought. I am not to say that NDVI or EM-DAT measure drought *better* than SPEI, but they arguably capture some aspects of drought that SPEI are not able to capture. Hence, further research should take this into consideration when measuring drought. Moreover, the findings also have political implications. If socio-economic drought carries a larger conflict potential than meteorological drought; how can we prevent meteorological droughts becoming socio-economic droughts? Drought resilience, such as irrigation and general improvements in living standards, makes people less vulnerable to meteorological drought. Hence, results from this thesis also suggest that these improvements could potentially reduce the risk of experiencing communal conflict.

1.3 Structure of the Thesis

I have started this thesis with a summary of the project and a short introduction of the knowledge gap this thesis is trying to fill. The remaining structure of the thesis is as follows:

Chapter 2: I continue the thesis with an outline of the ongoing research debate through a literature review. First, I summarise the research conducted on climate variability and conflict in general, with a particular focus on the relationship between drought and conflict. Moreover, I present a discussion of the literature on communal conflict. Through the literature review I highlight the knowledge gap which this thesis seeks to fill.

Chapter 3: In chapter three I present my theoretical framework and argument. In this chapter I elaborate on the exact definitions of drought and conflict applied in this thesis. Here I also present my theoretical argument and argue why this thesis is scientifically interesting. I end the chapter by formulating the hypotheses I will be testing in the analysis.

Chapter 4: In the fourth chapter I present the data used in the analysis. Since one of my main contributions to the research frontier is the use of new and unconventional ways of measuring drought, I devote a relatively large space to this chapter. I carefully discuss the strengths and limitations with the different drought indicators and examine the correlation between them. Finally, I discuss the use of control variables, which has been source of debate in this research field, and outline the control variables used in the analysis.

Chapter 5: Chapter five is devoted to the research design and statistical methods applied in order to test the various hypotheses. I begin this chapter with a brief discussion of the role of causality in social science, before I elaborate on the value of the multilevel logistic regression model. I end the chapter by examining the most important assumptions in order to be able to draw causal inferences from the results estimated by the multilevel model.

Chapter 6: In the sixth chapter I present and analyse the results from the regression models. In order to make sure these results are robust; I run several robustness tests and discuss whether using a sample consisting of the whole Sub-Saharan Africa is problematic when examining the relationship between drought and conflict. Furthermore, I briefly discuss whether the three drought indicators should be considered three different components of a drought, rather than three different definitions of drought.

Chapter 7: In the final chapter I summarise the thesis, discuss its strengths and limitations, as well as its political and scientific implications. I end the chapter, and thesis, by proposing several interesting avenues for future studies on the drought-conflict nexus.

1.4 Technical Note

This paper has been written in the Computer Modern font made by Donald E. Knuth using LaTeX. All data processing and statistical analyses have been conducted in Stata/SE 15.1. All graphs have been made in R using the ggplot2-package and maps have been made using QGIS version 3.8.3-Zanzibar. Spatial weights have been constructed using GeoDa 1.14.

Chapter 2

Literature Review

This chapter is devoted to the existing literature on the climate-conflict nexus. I divide the chapter in two. First, I begin by outlining the literature on climate variability and conflict. This is a relatively new research field emerging mostly during the last two decades. Second, I outline the literature on communal conflicts in particular, with the main theoretical focus on farmer-herder conflicts.

2.1 Climate Variability and Conflict

The idea that weather and climatological factors carry conflict potential has existed for centuries. For instance, in Shakespeare's *Romeo and Juliet*, Benvolio tells Mercutio that the hot weather makes it more likely that a fight will break out. Instead of retreating inside, they stay outside, and violent clash breaks out. Similarly, in *L'Étranger* by Camus, the protagonist Mersault suffers a heat stroke on the beach and eventually ends up shooting a man. And in *Roman Blood* by Saylor, Gordianus explains to Tiro that the number of stabbings usually increase during warmer weather in Rome.

Although all these three examples are from fictional literature, they all portray the same scenarios where humans are more violent during harsh weather conditions. Are the relationships the authors illustrate solely fictional or are they rather based on real connections?

The research field on climate and conflict has emerged over the last two decades. When researchers first started exploring the field, the focus was mainly put on the direct linkages between climate variability and conflict such as the effect of temperature or precipitation on conflict. In Africa, Burke, Miguel, Satyanath, Dykema and Lobell (2009) find a positive relationship between higher temperatures and more conflict, suggesting that more conflicts take place during warmer years. Similarly, Miguel et al. (2004) find a positive relationship between less precipitation and more conflicts in Sub-Saharan Africa. However, both these studies have been challenged in later years. Buhaug (2010) argues that Burke et al.'s (2009) findings are only valid for one specific operationalisation of conflict, and the significant relationship diminishes with further robustness tests. Similarly, Ciccone (2011) finds that Miguel et al.'s (2004) findings are only valid when using data between 1979-1999, not when the data are extended to 2009. In a response, Miguel and Satyanath (2011) suggest that this may be due to Africa's recent economic growth in non-agrarian sectors making people less vulnerable to climatic shocks.

In following years, most research has focused on the relationship between drought and conflict, as drought is the climatic hazard commonly assumed to carry the largest conflict potential (von Uexkull et al., 2016, p. 2). When measuring drought, researchers have tended to use proxies such as either raw precipitation data or standardised measurements such as The Standardised Precipitation Index (SPI). Counter-intuitive to most hypotheses, Theisen (2012) and Witsenburg and Adano (2009) find that wetter years are associated with both more conflicts and more intense conflicts in Kenya. Meier, Bond and Bond (2007) find the same association in the Karamoja Cluster – the border region between Kenya, Uganda, South-Sudan and Ethiopia. A suggested explanation for this relationship is that more rainfall leads to higher vegetation and more camouflage, which in turn makes it easier to track and ambush cattle without getting caught (Meier et al., 2007, p. 731). Similarly, wetter years lead to higher economic value of raiding pastoralists' cattle than during dryer years and with increased rainfall, violent actors may find it easier to expropriate wealth from the population (Theisen, 2012, p. 84; Salehyan & Hendrix, 2014, p. 241). On the other hand, other scholars find that warmer years see more conflicts in East-Africa and Sub-Saharan Africa in general, whereas precipitation does not matter (O'Loughlin et al., 2012; O'Loughlin, Linke & Witmer, 2014).

In recent years, researchers have evolved beyond the use of sole precipitation data to more complex drought-measures. Salehyan and Hendrix (2014) use Palmer's Drought Severity Index (PDSI)[3] and find that across Africa, more precipitation, not droughts, are associated with more conflicts. Von Uexkull et al. (2016), Harari and Ferrara (2018) and Döring (2020) all use the Standardized Precipitation Evapotranspiration Index (SPEI) which builds on SPI, but also takes into account potential evapotranspiration such as temperature. Evapotranspiration is defined as the sum of evaporation, which comprises water vaporisation from non-living objects such as soil, water bodies and wet surfaces, and transpiration, which includes water vaporisation from plants (Berner Jr, 2009). Although both PDSI and SPEI arguably measures drought better than solely relying on precipitation, they primarily measure meteorological drought.[4]

Additionally, researchers have focused on different types of conflict. Hendrix and Salehyan (2012) use a broad definition of conflict by including disruptive activities such as demonstrations, riots, communal conflict and anti-government violence. Across Africa, they find that wetter years are associated with more *violent* conflict. Whereas extreme deviations, such as particularly dry and wet years, are associated with *all types* of conflict – not just violent conflict. Similarly, in East-Africa, Raleigh and Kniveton (2012) find differences between conflict types. Particularly dry years are associated with more *rebel* conflicts, whereas wet years are associated with more *communal* conflicts.

A large focus has been put on communal conflicts as these conflicts are expected to be particularly prone to weather events. By focusing solely on communal conflicts, Fjelde and von Uexkull (2012) find that the likelihood of experiencing communal conflict is higher following a drought in Sub-Saharan Africa. Nordkvelle et al. (2017) find that droughts increase the likelihood of experiencing communal conflicts in Nigeria, Kenya, Uganda, Sudan and India. Similarly, Döring (2020) finds that the lack of water availability, not just rainfall, but also the lack of ground water, is associated with more communal violence.

Due to a lot of contradictory findings and a lack of consensus, researchers have argued that there does not seem to be any clear direct linkages between climate

[3]The PDSI measures meteorological drought through a combination of precipitation, temperature and soil condition (Salehyan & Hendrix, 2014).

[4]I elaborate more on various drought definitions in Chapter 3.

variability and conflict. Instead researchers should better specify the contextual factors in which climate variability is assumed to affect conflict (von Uexkull et al., 2016). Some of these factors have been agricultural production and agricultural dependence. In Sub-Saharan Africa, von Uexkull (2014) finds that areas with rainfed croplands see an increased risk of civil conflict violence following droughts. Similarly, Harari and Ferrara (2018) find that droughts during growing season are both associated with more conflict events and conflict onsets in Africa in general. On the other hand, when examining ethnic groups dependent on agriculture, von Uexkull et al. (2016) find no clear relationship between drought and conflict onset, but a relationship between drought and conflict events, suggesting that droughts may lead to longer conflicts, but not necessarily more conflicts.

Focusing on ethnic groups, Theisen et al. (2011) investigate the relationship between drought, ethnic marginalisation and conflict. They do not find any evidence of a relationship between drought and civil conflict, although they find a strong link between ethnic marginalisation and conflict. Suggesting that politics and structural factors cause conflicts, not the environment.

When focusing on drought and communal conflict, researchers often assume the conflict potential induced by drought is mediated through resource scarcity. A second way drought is assumed to affect conflict potential is through the loss of food production and higher food prices. Buhaug, Benjaminsen, Sjaastad and Theisen (2015) find a strong connection between weather patterns and changes in food prices, but no relationship between food prices and violent conflict. On the other hand, Koren (2018) finds that in years where wheat and maize yields are higher, countries see more conflicts. Fjelde (2015) investigates individuals' propensity to take up weapons, and finds that income shocks in the agricultural sector substantially increase the risk of violent events, suggesting that loss of income makes farmers and pastoralists more inclined to join rebel groups. Similarly, in Africa, both Vestby (2019) and von Uexkull, d'Errico and Jackson (2020) find that participation in violence would have been more likely if an individual experienced a deterioration of living conditions due to drought. On the other hand, Linke, O'Loughlin, McCabe, Tir and Witmer (2015) find little support that droughts make people more likely to support the use of violence in three specific regions in Kenya.

From this literature review so far, it is evident that research on the climate-conflict nexus is prone to different contextual specifications and operationalisations. Even two meta-studies trying to sum up the debate reach different conclusions (Hsiang, Burke & Miguel, 2013; Theisen, Gleditsch & Buhaug, 2013). Moreover, Busby (2018) argues that a reason for this conundrum is due to the inclusion or exclusion of control variables. According to Busby, studies not including control variables tend to find a strong relationship between climate variability and conflict, whereas studies including these not tend to find any clear relationship. I elaborate more on the importance of, and reason for, including control variables in Chapter 4 and 5.

In order to make a robust recap of the ongoing debate on climate and conflict, Mach et al. (2019) conduct expert interviews with eleven climate and conflict experts.[5] They find that all experts agree that climate variability have affected armed conflict within countries in some way over the past century. However, they also agree that other factors carry a larger conflict potential than climate variability, such as low socioeconomic development, low capabilities of the state, inter-group inequality and recent history of conflict (Mach et al., 2019, p. 194). Across these experts, there is low confidence in the exact mechanisms through which climate variability affects conflict risks. In particular, economic shocks and dependency on natural resources are judged likely to be one possible mechanism of the climate-conflict relationship. Climatic hazards can hinder agricultural productivity or affect food prices, while also have direct effects such as floods, droughts, heat waves or cyclones (Mach et al., 2019, p. 195). However, on the other hand, dependency on natural resources can also stimulate cooperation and thus decrease conflict risk if conditions are unfavourable for sustaining an armed group (Mach et al., 2019, p. 195). One example of this is Witsenburg and Adano (2009) finding that people in northern Kenya do not necessarily engage in violent conflict over access to water resources. Instead of causing conflicts, droughts and times of resource scarcity foster cooperation and warring communities are seen to reconcile in order to use water and pasture together (Witsenburg & Adano, 2009, p. 520).

[5] These eleven experts are Halvard Buhaug, Marshall Burke, James D. Fearon, Christopher B. Field, Cullen S. Hendrix, Jean-Francois Maystadt, John O'Loughlin, Philip Roessler, Jürgen Scheffran, Kenneth A. Schultz and Nina von Uexkull.

To sum up the literature on climate and conflict, it seems as if findings diverge based on the inclusion or exclusion of contextual variables and the causal pathways examined. Yet one of the few things researchers have in common is the use of precipitation-based drought measures as their explanatory variable. Hence, there is a scientific knowledge gap regarding whether other operationalisations of drought yield different estimates on the conflict risk.

2.2 Communal Conflict

The literature on communal conflicts tends to be different from the literature on climate and conflict relationships. Whereas most research on the climate-conflict nexus is quantitative, research on communal conflicts tends to be qualitative with in-depth case studies and ethnographic research.

A communal conflict is often defined as a conflict between informal groups (Brosché & Elfversson, 2012). In short, this means that neither conflict part is a state nor a permanent rebel group. In Sub-Saharan Africa, a large part of communal conflicts is between farmers and herders (von Uexkull & Pettersson, 2018). According to Turner, Ayantunde, Patterson and Patterson (2011) the literature on farmer-herder conflicts has mainly clustered around two strands. The first strand portrays these conflicts as often induced by resource scarcity and environmental security. The second strand argues that these conflicts reflect cultural animosities between farmers and herders, and should not be reduced to a simple case of environmental security.

The first strand of literature on this type of conflict relates to the resource-scarcity hypothesis suggested by Homer-Dixon (1999). Homer-Dixon (1999) propose that scarcity of renewable resources can contribute to civil conflicts – such as insurgencies or ethnic clashes. Scarcity of renewable resources can include water resources, grazing land and arable land. Although Homer-Dixon (1999) is careful in stating the exact linkages between scarcity and conflict and denotes that "[e]nvironmental scarcity is never a sole or sufficient cause of large migrations, poverty, or violence[...]" (Homer-Dixon, 1999, p. 16), he has often been depicted as one of the forerunners of the environmental scarcity strand. One reason for this is that he argues that with the coming climate changes, we are likely to see more conflicts with an environmental connection. This has led many to believe that there exists a link between resource scarcity and conflict.

In the literature, there exist several examples on how drought and scarcity may affect conflict. For example, Moorehead (1989, cited in Shettima & Tar, 2008) finds that conflict between farmers and pastoralists in the Niger River delta erupted when the delta became drier. Since the delta became drier, farmers started cultivating new parts of the delta, overtaking pastoralists' grazing areas. Eventually, this led to violent conflicts between the two groups. Similarly, Hundie (2010, p. 141) argues that droughts in the mid-1990s in Ethiopia resulted in scarcity of rangeland. This resulted in new routes for pastoralists, which eventually led to more frequent violent confrontations between the Karrayyus group and the Afar group. Hundie (2010, p. 142) also found these conflicts to be more intense during droughts.

In South Sudan, the two most populous groups, the Dinka and the Nuer, have been fighting each other for almost a century over agricultural land (Wig & Kromrey, 2018, p. 415). And across the border, in the state of Southern Kordofan in Sudan, droughts have affected both the timing and the migratory routes pastoralist groups use when moving north in the rainy season and south during the dry season (Bronkhorst, 2011, p. 15). Conflicts have frequently erupted between pastoralists and farmers because of these shifting routes (Bronkhorst, 2011). In the Asante Akim North District in Ghana, droughts in neighbouring countries have led to a large immigration of Fulani herders (Amankwaa, 2019). This is a typical example of rural-to-rural migration which is often assumed to spark tensions between the indigenous "sons-of-the-soil" and new migrants (Fearon & Laitin, 2003).

Even though there are many examples of conflicts where climate variability seems to have played an important role, the scarcity-strand has been heavily criticised for creating overly simplistic, reductionistic portrayals of conflicts which in fact are socially produced (Benjaminsen, Maganga & Abdallah, 2009; Turner et al., 2011). Similarly, Turner (2004) have argued that farmer-herder conflicts are complex conflicts that environmental security analysts have used to extend the notion of scarcity induced conflicts to more modern resource wars.

In Sub-Saharan Africa, occupations of farmers and herders are often tied to ethnic and caste identities (Turner, 2004). Often these relationships are even more complex where they are composed by social, political and economic strands such as patron and client; landlord and tenant; sellers and buyers; or livestock owner, trader, and herder (Turner, 2004, p. 872). Thus, solely reducing the conflict to an ethnic climate war, may be simplistic. However, although the different groups are dependent on each other, Breusers et al. (1998, cited in Shettima & Tar, 2008) find that droughts in the 1970s and 1980s resulted in a breakdown in the balance between farmers and pastoralists. Previously, farmers had depended on meat and milk from pastoralists, and pastoralist had depended on crops and vegetables. During the droughts farmers started cattle breeding and pastoralists started cultivating. Thereby leading to the disappearance of the interdependency between the groups and larger competition between them. These new groups are often referred to as agropastoralists.

Other scholars have argued that farmer-herder conflicts often have a political origin associated with an ongoing process of pastoral marginalization (Benjaminsen, Alinon, Buhaug & Buseth, 2012; Benjaminsen & Ba, 2009). In many African countries, national legislation tends to favour farmers in farmer-herder disputes, as farmers are stationary and add 'productive value' to the land, whereas pastoralists are non-stationary and exploit the land (Benjaminsen et al., 2009, p. 424). In the Kilosa District in Tanzania, there is a narrative shared by farmers and local authorities that the main reason for farmer-herder conflicts is that herders overgraze their own land and subsequently enter farmers' land (Benjaminsen et al., 2009, p. 433).

Hence, as suggested during the expert interviews by Mach et al. (2019), political marginalisation may be a more important contributor to the conflicts than climatic factors. Moreover, in rural areas of Sub-Saharan Africa, institutions are usually scarcely developed. Rent-seeking and corruption among government officials have undermined people's trust in institutions (Benjaminsen et al., 2012, p. 109). As there are trust in institutions to settle these disputes, people often tend to take the matters into their own hands.

Thus, whether or not climatic factors have played a causal role in these conflicts is difficult to say. In some conflicts, there might be a combination of both strands where there are political and economic underlying tensions and a drought may act as a triggering element of these grievances. In this thesis, I am not trying to determine to what extent the drought causes the conflict or whether there are underlying factors triggered by the drought. Instead, I am trying to unveil whether droughts are associated with more communal conflict incidences.

In this chapter I have outlined the literature on both climate variability and conflicts, as well as communal conflicts in particular. The literature on climate variability and conflict has gradually developed from focusing on the direct linkages to more complex and conditional relationships between climate and conflict. However, there is still a problematic paradigm — most researchers only measure meteorological drought. As I will discuss further in the next chapter, this is in most cases not compatible with the researchers' own theoretic claims. The literature on communal conflicts, on the other hand, has been characterised by a hard line between those who believe climate variability have played a role in the conflicts, and those who believe that all conflicts are essentially a result of social processes and politics. In the next chapter I discuss how this thesis relates to this debate, and argue why communal conflict is the most likely conflict type during drought.

Chapter 3

Theoretical Framework

It has been argued that the elusiveness of the concepts *climate* and *conflict* can explain why there is a scientific conundrum in the climate-conflict field (Salehyan, 2014, p. 1). Similarly, von Uexkull et al. (2016, p. 1) argue that the failure to properly specify the political and socioeconomic context in which climate extremes can aggravate tensions, may lead to ambiguous findings. Hence, specifying exactly what is meant by a drought and how the drought is expected to affect the likelihood of conflict, is of particular importance in this thesis. In order to carefully specify these conditions, I divide this chapter in two. I begin the chapter by defining the concepts of drought and communal conflict. In this part I go through commonly used definitions of the concepts, but the main aim is to provide a clear picture of what I put into these two concepts in this thesis. In the second part, I discuss how drought may lead to conflict and outline the scientific value of this thesis arguing why this research is crucial in a research field dominated by ambiguous findings.

3.1 Key Concepts

Before starting an in-depth analysis of the relationship between drought and conflict, it is essential to provide clear and concise definitions of these two concepts.

3.1.1 Drought

What is a drought? People tend to have a relatively clear idea of what a drought is, namely the absence of water. This coincides with Wilhite's (2000) short definition of a drought as a "shortage of water to meet essential needs". But what causes this shortage of water? Is it caused by less precipitation than normal? Is it due to higher temperatures causing more evapotranspiration? Or is it rather a result of over-exploitation of available water reserves or other forms of mismanagement of existing water resources? And is there a difference between shortage of *water* and shortage of *precipitation*?

A drought is often described as a slow onset hazard (Wilhite, 2000). Unlike sudden weather events such as tornadoes, earthquakes and tsunamis, it can be difficult to determine the exact start and end point of a drought. Hence, drought has often been characterised as a "creeping phenomenon" (Tannehill, 1947). Moreover, drought is a highly complex phenomenon as it may last for days, weeks, months or even years, and both the direct and indirect effects of a drought can differ between being slightly noticeable to being highly fatal. Thus, the harmful effects of a drought may be difficult to identify.

Although "drought" is a widely used concept, the exact definitions tend to vary. Wilhite and Glantz (1985) found over 150 different definitions of droughts when examining published articles on drought. To systematise this, they created four different overarching categories: (1) meteorological drought, (2) agricultural drought, (3) hydrologic drought and (4) socio-economic drought. In this thesis I use three different measures of drought: SPEI, NDVI and EM-DAT. In order to theoretically conceptualise the use of these drought indicators, I argue that SPEI can be labelled as a measure of meteorological drought, NDVI as a measure of agricultural drought and EM-DAT as a measure of socio-economic drought.[6]

A meteorological drought is based on precipitation anomalies (Wilhite & Glantz, 1985). Usually this is measured by comparing the rainfall from one period against the average of the same period within an area. For instance, in 2014, The Democratic Republic of Congo had an average rainfall of 1543mm per year, while Niger only had 151mm (FAO, 2016). This means that a meteorological drought in DR Congo would require a much larger deviation in precipitation than in Niger, and hence comparing absolute deviations in rainfall does usually not make sense. This also highlights the

[6] I do not have a measure capturing the concept of hydrologic drought as I started with the data before I linked the measures to the theoretical concepts. For a thorough analysis of the relationship between hydrologic drought and conflict see Döring (2020).

need to look at drought by comparing local anomalies to the local normal as people have adapted to these local living conditions for centuries. In particular, farmers tend to cultivate in crops that benefit from the local environment, both in terms of precipitation and temperature, but also in terms of soil condition. Hence, measuring absolute rainfall deviations is not comparable across areas. Since SPEI is a measure mainly based on the lack of rainfall, I argue that SPEI can be perceived as a measure of meteorological drought.

An agricultural drought, on the other hand, refers to the amount of water needed by crops during different growing stages (Wilhite & Glantz, 1985). A plant's demand for water is not constant, but dependent on the biological characteristics of each plant and their respective growing season. For example, the root system of sweet potatoes has a big surface allowing it to easily access available soil water, contrary to other staple crops such as maize, cowpeas and rice (Loebenstein and Thottappilly, 2009). This makes the sweet potato more resilient to rainfall deviations than other crops. In order to be characterised as an agricultural drought, the drought needs to affect the crops. Just like the example of DR Congo and Niger, a field of sweet potatoes would require larger rainfall deviations than a field of maize for it to be characterised as an agricultural drought. Similarly, areas covered by rainforest are much more resilient to rainfall deviations than grassland or cropland. An agricultural drought also depends on the timing of the precipitation deviation. A meteorological drought occurring during the winter would most likely not result in an agricultural drought as it does not affect the plants' growing season. Since NDVI measures the quality of vegetation, it arguably captures these effects to a larger extent than SPEI. Hence, I argue that NDVI can be labelled a measure of agricultural drought.

The third drought type is hydrologic drought. A hydrologic drought occurs when the drought affects the water volumes in streams, rivers and water reservoirs. This is often out of sync with meteorological and agricultural droughts, as this tends to happen several months later than the rainfall deviation (Wilhite & Glantz, 1985, p. 7). A strict interpretation of hydrologic drought only includes surface water. However, some scholars have argued for ground water drought to either be included in this category or added as an additional drought type (Mishra & Singh, 2010).

The fourth and last type of drought, socio-economic drought, occurs when the drought starts affecting humans' daily life (Wilhite & Glantz, 1985). In other words, when the water demand exceeds the supply. This could be caused by the lack of precipitation during crucial times. But this could also be caused by human activity, such as too much irrigation, over-exploitation of water resources or lack of proper water management. Lack of proper water management can either be a result of purpose, as seen in Israel and Palestine (Corradin, 2016), or as a result of a general lack of capacity (Homer-Dixon, 1999). Sometimes the lack of water management is referred to as "water scarcity" rather than "drought" (Pedro-Monzonís, Solera, Ferrer, Estrela & Paredes-Arquiola, 2015). However, I will not discuss the differences between "water scarcity" and "drought" further as this debate is outside the scope of this thesis. Instead I apply a broad concept of drought comprising more than just the lack of precipitation. Since EM-DAT is based on crisis reports, it measures droughts that affected people's daily life to a large extent. Hence, I argue that EM-DAT can be perceived as a measure of socio-economic drought.

As discussed in the previous chapter, most researchers on the climate-conflict nexus tend to use the definition of a meteorological drought without an explicit discussion of different drought types. When measuring meteorological droughts, however, we are not able to properly determine whether the drought actually affected the human life. Is a drought expected to carry conflict potential even though it does not affect the people living in the respective area? This is a challenging question this thesis seeks to provide a better understanding of.

The role of climate change

The growing focus on climate's effect on human behaviour has emerged as a consequence of the human induced climate change. Climate is usually defined as the mean and variability of temperature, precipitation and wind over a longer period — often 30 years (Cubash et al., 2013). Climate change, in particular, is defined as a change in this mean over an extended period, typically decades or longer. If we were to use data on the mean of climate over decades or longer, it would leave us with very few data points and a limited variation in order to explain causality. Hence, researchers have put the focus on climate variability instead. In conflict literature, climate variability refers to shorter-term descriptions of weather, such as weeks, months or years, as opposed to the decade long time periods of climate change (Vestby, 2018, p. 11).

Since 1880 the average global surface temperature on Earth has increased by 0.85°C (Hartmann et al., 2013). This increase is further expected to exceed at least 2°C compared to the 1850-1900 average (Collins et al., 2013).[7] Even if countries comply with their nationally determined obligations in the Paris Agreement, the mean temperature is still expected to exceed 3-4°C, even though the overarching goal of the Paris Agreement is to prevent the temperature of exceeding 1.5°C (Young, 2016).

Although the exact implications of climate change are still relatively uncertain, there is evidence that there will be more climate extremes (Collins et al., 2013). These trends are complex and diffuse and climate extremes will most likely affect areas differently. For instance, there are indications that there will be more heavy precipitation in North America, Central America and Europe, while there will be less precipitation in other regions, such as Southern Australia and Western Asia. Likewise, there exist indications that the number of droughts will increase in some regions, such as the Mediterranean and parts of Africa, and decrease in others, such as central North America and Australia (Collins et al., 2013).

With increased temperature, potential evapotranspiration is predicted to increase by 1.5%–4% per °C warming (Scheff & Frierson, 2014). This leads to droughts being projected to happen five to ten times more frequent in Africa, and current 100-year events could occur every two to five years under 3°C of warming (Naumann et al., 2018). Furthermore, water supply-demand deficits could become fivefold in size for most of Africa (Naumann et al., 2018). These effects have gradually started to take place. In a recent round of Afrobarometer (2018), 48.3% of the population in Sub-Saharan Africa said droughts had been much more or somewhat more severe the last ten years, as opposed to 28.3% stating they were much or somewhat less severe. Since droughts are expected to be more frequent in the future, there is a

[7] In the Fifth Assessment Report of IPCC (AR5) four different climate change projections are used. These four projections are called Representative Concentration Pathways (RCP) and are based on the radiative forcing in year 2100 relative to 1750. These four RCPs include one mitigation scenario with low forcing level (RCP2.6), two stabilisation scenarios (RCP4.5 and RCP6.0) and one scenario with high greenhouse gas emissions (RCP8.5). Temperature is unlikely to exceed 2°C for RCP2.6, while as likely as not to exceed 2°C for RCP4.5, and likely to exceed 2°C for RCP6.0 and RCP8.5 (Collins et al., 2013).

growing demand for scientific knowledge on the exact causes and future implications of droughts in particular, and other implications of climate change in general. One of these implications may be violent conflict.

Impacts of drought

The impacts of drought can be devastating. Typical characteristics for drought impacts are that they are non-structural and often spread over large geographical areas (Wilhite, 2000). Thus, they may spand over large areas, but affect these areas differently. As discussed previously, there can easily be a meteorological drought without any socio-economic consequences. This could be if the meteorological drought occurs outside the growing season, or if the meteorological drought takes places where there are few inhabitants or natural resources dependent on rainfall. In rural areas dominated by rainfed agriculture, societies are particularly vulnerable to drought (Cooper et al., 2008). These are areas without irrigation, meaning that agriculture is solely dependent on rainfall and surface runoff. In Sub-Saharan Africa, more than 95% of all agriculture are rainfed (International Water Management Institute, 2010). Combined with the fact that more than half of all workers in Sub-Saharan Africa are employed in the field (Dercon & Gollin, 2014), this region is extremely vulnerable to drought.

When farmers are dependent on rainfall, a drought can lead to severe losses of income and livelihood. This is not only true for farmers, but also for pastoralists who rely on drinking water and grazing land for cattle. In periods with severe droughts this may fatally deteriorate their living conditions. If large societies are dependent on rainfed agriculture, droughts may lead to loss of food production and eventually result in famines. In these cases, the drought does not only affect the workers dependent on rainfall such as farmers or herders, but whole populations. In the period 2010-2012 the Horn of Africa experienced one of the worst famines in recent years, putting more than 12 million people in need of urgent assistance as a result of drought (FAO, n.d.). The exact number of deaths vary largely depending on the source. While the EM-DAT database reports 20.000 fatalities (Guha-Sapir, Below & Hoyois, 2016; Rosvold & Buhaug, 2020), Checchi and Robinson (2013) estimate that approximately 250.000 lost their lives as a consequence of the drought.

Although the direct impacts of drought, such as famines, water shortage and loss of livelihood are well documented in the literature, there is still large uncertainty on the conflict potential carried by drought.

3.1.2 Communal Conflict

Another central concept that needs a clear definition is *communal conflict*. In this thesis I apply the definition used by the Uppsala Conflict Data Program (UCDP), which defines a communal conflict as a “violent conflict between non-state groups that are organised along a shared communal identity” (Brosché & Elfversson, 2012, p. 35). This definition consists of three main parts which deserve further clarification. First, a *violent conflict* refers to the fact that parties use lethal violence to gain control over resources. This violence can be conducted with modern weapons such as guns, or with less technical equipment such as sticks and stones. The resources they fight over can either be physical resources, such as land areas, or more abstract resources such as political power. Second, the actors are *non-state groups*. This means that the groups are neither representing the state, nor the armed forces. Finally, these groups are *organised along a shared communal identity*, which means that the members share a common identity and the groups are not formally organised rebel groups. Communal identity is sometimes referred to as a common ethnic, religious or national identification. However, similarly to Brosché and Elfversson (2012), I leave this definition purposefully more open, as group identification can be based on more than just ethnicity, religion or nationality, such as for instance common history, culture or core values.

Farmer-herder conflicts

As mentioned previously, a large share of communal conflicts in Sub-Saharan Africa consists of farmer-herder conflicts. In an in-depth study of non-state conflict data, von Uexkull and Pettersson (2018) find that most communal conflicts in Africa between 1989 and 2011 were between pastoralists and pastoralists (23%), pastoralists and agropastoralists (17.5%) or farmers and pastoralists (13.6%). Less than 10% of all communal conflicts in Africa contained neither farmers, nor pastoralists. Figure 3.1 shows a pie chart of the livelihood of the different actors in communal conflicts.

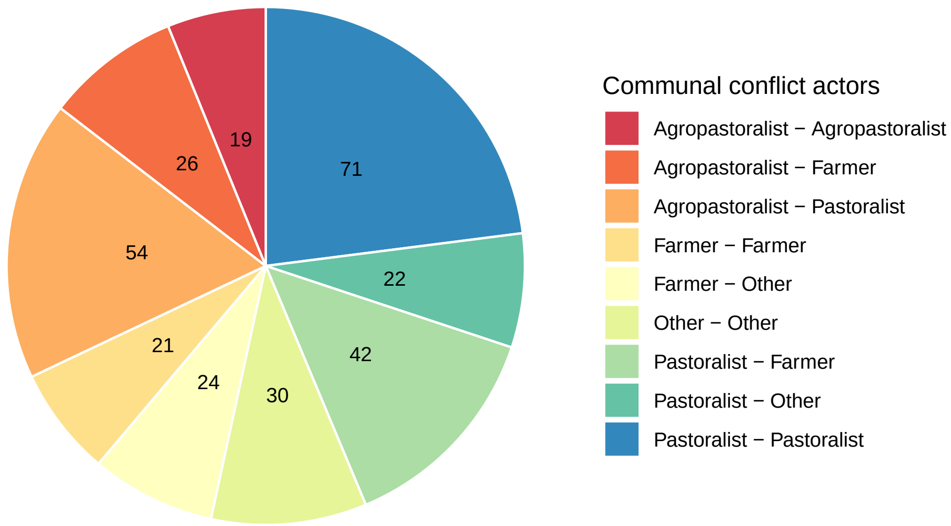


Figure 3.1: Actors in communal conflicts

Hence, although I include all communal conflicts in the analysis, the main theoretical focus will be on farmer-herder conflicts since these comprise a large share of the conflicts included in the dataset.

Farmer-herder conflicts frequently erupt over the use farmland, grazing areas, stock routes or access to water points (McDougal, Hagerty, Inks, Dowd & Conroy, 2015). Von Uexkull and Pettersson (2018) find that in 56.4% of all communal conflicts either agricultural land, water resources or livestock were seen as an issue. The rest of the conflicts were made up of other territorial issues (35.3%) such as border demarcation that do not fall into agricultural land/water sub-issue category; authority issues (26.5%) such as control of the local administration and influence in the state administration; religion (6.1%) or other issues (13.6%). These categories are not mutually exclusive, and since farmers and herders often are tied to different ethnic groups, it can be hard to determine whether a conflict should be labelled as a farmer-herder conflict, ethnic conflict or religious conflict (Turner et al., 2011). One example of this is the ongoing conflict between Muslim herders and Christian farmers in Nigeria. This conflict has been going on for decades, consisting of many small-scale conflicts. These conflicts have commonly been portrayed by the media as triggered by frequent droughts and the disappearance of grazing land and water sources (Blomfield, 2018). However, in recent years terrorism by Boko Haram have been fuelling these tensions (Onapajo & Usman, 2015) and hence it might be even harder to detect the cause of each conflict.

3.2 Theoretical Argument

The aim of this study is twofold. First, it is *explanatory* as I try to determine the causal relationship between drought and communal conflict in Sub-Saharan Africa. Second, it is *exploratory* in the sense that I use new ways of measuring drought than previously used on the climate-conflict nexus.

3.2.1 Explanatory Argument

As pointed out previously, the exact pathways between drought and violence are still not fully understood. One proposed way in which drought may contribute to violence is through resource scarcity (Homer-Dixon, 1999). Scarcity of water resources, arable land and grazing land may all contribute to disagreements between groups basing their livelihood and income on these resources. When farmers lose crops as a result of drought, they may expand their cultivation and end up in clashes over arable land with other groups, such as Moorehead (1989, cited in Shettima & Tar, 2008) found in the Niger River delta. Moreover, for pastoralists, droughts frequently drive migration to well sites and rivers that members from other ethnic communities use (Theisen, 2012). In Kenya, Detges (2014) found that communal violence is more likely close to these well sites and Döring (2020) found that areas with lower water capacity see more conflicts in general. Since pastoralists are nomadic herders migrating with their livestock during the year, they usually follow the same routes and cycles from year to year. However, if droughts deteriorate the grazing land on these routes, they need to change routes. By moving to areas with richer water resources and pasture, they may end up closer to other groups, making them more likely to get involved in such disputes (Van Baalen & Mobjörk, 2016).

In the examples above, communal conflict is the most probable response to drought in order to secure access to livelihood essentials such as grazing land, water holes or agricultural land (Fjelde & von Uexkull, 2012; Hendrix & Salehyan, 2012). While a state's lack of response may spark grievances for both farmers and pastoralists towards the state, the most rewarding short-term response is often attacking other societal groups and not the state (Fjelde & von Uexkull, 2012, p. 446).

Hence, another hypothesised pathway in which drought may cause conflict is either through the loss of food production and higher food prices or through the lack of responsiveness to the drought, which in turn may spark popular riots or driving people to take up arms against the government. However, in line with Salehyan (2014) and Seter (2016), I argue that researchers are not able to capture both these causal pathways in one analysis as the outcome variable (the conflict type) and the temporal and spatial scale of measurement will differ between these two pathways. Whilst communal conflicts often take place close to the drought (Fjelde & Østby, 2014, p. 746), popular riots often take place in larger cities. Moreover, while a communal conflict is assumed to take place in the same year as the drought is occurring, popular riots should be measured with a temporal lag as it often takes months or even a year for the drought to affect the food prices.

To sum up my explanatory argument, I systematise the argument according to Seter's (2016) four key elements when evaluating causal mechanisms on climate variability and conflict. The first element refers to the most relevant actors. As I have argued, since rainfall is a critical element for pastoralism and farming (Sulieman & Young, 2019, p. 12) and more than two-thirds of the working force in Sub-Saharan Africa are engaged in these activities (Stern, 2006, cited in Fjelde & von Uexkull, 2012, p. 445), farmers and pastoralists are the most likely actors. The second element is determining the type of climate variability most likely to affect these actors (Seter, 2016). First, droughts is by far the most common and harmful disaster for both livestock and agriculture (FAO, 2017). Second, I argue that a problematic aspect in the literature is how to measure drought. Hence, I choose to focus on drought as the type of climate variability associated with communal conflicts. This leads to the third key element which is the most likely conflict type (Seter, 2016). As argued, communal conflicts between groups fighting over scarce resources are assumed to be one of the most likely outcomes of drought (Fjelde & von Uexkull, 2012; Nordkvelle et al., 2017). The fourth, and last, element Seter (2016) outlines is the temporal and spatial scale for measuring this causal claim. As I have briefly touched upon, communal conflicts tend to be small-scaled conflicts in close proximity to the climate variability both in terms of temporal and spatial distance. Hence, measuring how drought affects these conflicts should be done within a disaggregated statistical analysis. I will get back to this in the next chapter.

3.2.2 Exploratory Argument

The second aim of this thesis is exploratory: Do different concepts of drought carry different conflict potential? And do they represent different types of drought or rather different sequences in the causal chain of a drought? As outlined in the literature review, almost all former quantitative studies have relied on precipitation measures as a way of measuring drought, either through raw measurements of precipitation or through indices such as SPI or SPEI. As discussed in this chapter, these measures mainly capture the concept of meteorological drought. However, meteorological drought is not necessarily assumed to carry conflict potential — especially if it does not deteriorate living conditions. The rainforest in DR Congo is not affected to the same extent by the lack of rainfall as the grassland and cropland in the Sahel. Yet researchers commonly assume that a meteorological drought carry the same conflict potential in both places. Hence, the main aim of this thesis is to use other operationalisations of drought to try to get closer to the actual mechanisms — where droughts causing loss of livelihood are assumed to be associated with more conflicts. There are some obvious methodological challenges with trying to measure the actual impact of drought. I will discuss these challenges further in next chapter.

3.2.3 Hypotheses

By combining my explanatory and exploratory argument I construct two testable hypotheses. First, I expect there to be a positive relationship between all three drought indices and the likelihood of communal conflict. This is formulated in Hypothesis 1 (H1):

Hypothesis 1 (H1): *Droughts are associated with a higher likelihood of communal conflict.*

However, I do not expect the effects to be similar across the three drought indices. More specifically, I assume that the larger physical impact of the drought, the higher

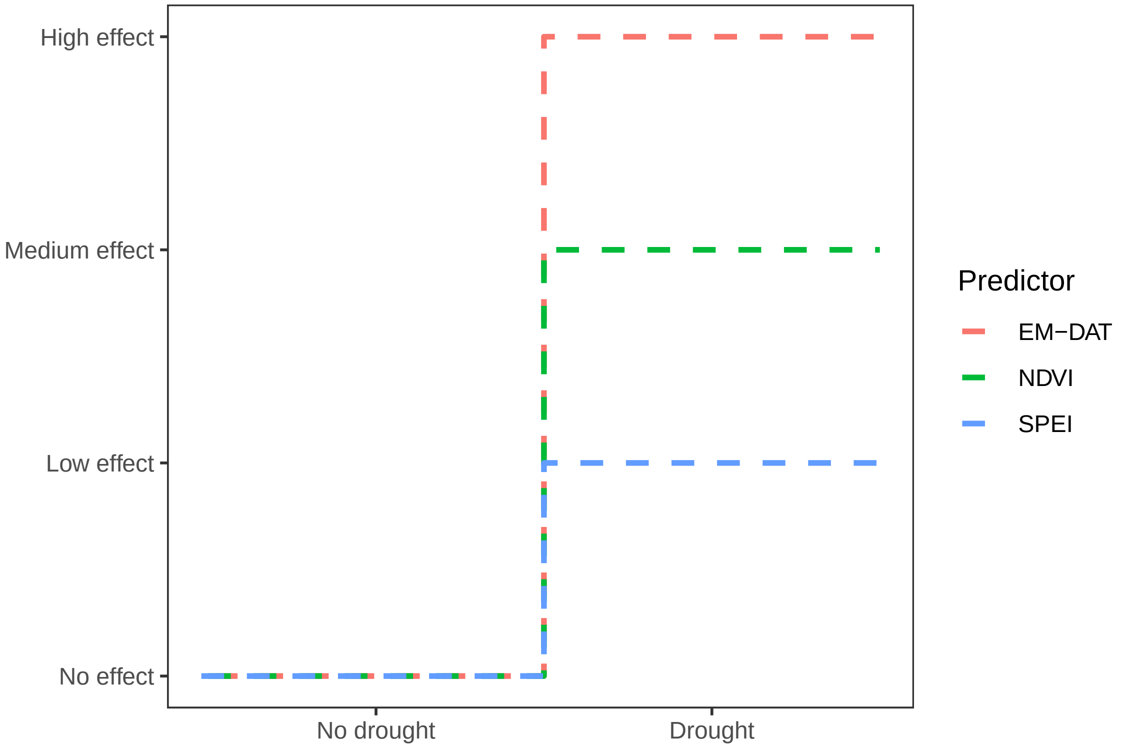


Figure 3.2: Illustration of hypotheses

the likelihood of communal conflict. This is reflected in the explanatory argument and formulated in Hypothesis 2 (H2):

Hypothesis 2 (H2): *The higher physical impact of the drought, the higher the likelihood of communal conflict.*

In Figure 3.2 I illustrate both hypotheses. With no drought, the conflict potential carried by either indicator is (of course) zero. When there is a drought, however, I expect all indicators to be positive, but I expect SPEI to have a relatively low effect, NDVI to have a medium effect and EM-DAT to have the strongest effect. This is shown in the Figure 3.2 with SPEI as the blue line, NDVI as the green line and EM-DAT as the red line.

To sum up my theoretical argument, I argue that the human consequences of a drought is not adequately measured by precipitation-based measurements. Hence, there is a compelling need for testing other measures of drought. This is what I refer to as this thesis' *exploratory* argument. I also argue that farmers and herders are the most vulnerable groups to climate variability and hence focusing on these groups and the conflicts within or between them is the most likely conflict affect by drought. Testing these claims is what I refer to as my *explanatory* argument.

Chapter 4

Data

Simultaneously with a growing focus on the climatic-conflict relationship, the technical side of conflict data has been focusing on disaggregated data, with the use of smaller entities. Traditionally, conflicts have been studied at the country level (Cederman & Gleditsch, 2009). Yet most of the conflicts in the world today are local conflicts rarely engulfing the whole country (Buhaug & Rød, 2006). Since these conflicts, and communal conflicts particularly, are local phenomena, the relevant local characteristics cannot be captured by using country-level measures. As a result of this, researchers have started using smaller entities such as sub-national levels (see Fjelde & von Uexkull, 2012; Meier et al., 2007) or grid cells (see Theisen et al., 2011; von Uexkull, 2014).

Not only do the use of disaggregated data capture the local characteristics better, but it also leads the researcher closer to the actual causal mechanisms. The proposed causal mechanism in this thesis, that droughts affect the likelihood of communal conflict by driving confrontations between farmers and herders, is expected to be local in nature. Failing to acknowledge this and hence using country-year data to measure the relationship would impose large amounts of noise to the results. For instance, a local drought in the southern part of DR Congo could falsely be associated with a conflict in the northern part over 2,000km away. Hence, the use of disaggregation seems appropriate in this study. Moreover, sub-national units such as first order administrative units vary largely in size across different countries in Sub-Saharan Africa and they may also be analytically interesting in explaining violence themselves (Theisen, 2012, p. 87). As a result of this, I choose to use grid cells as units in my analysis.

There are clear benefits of using grid cells as units. All cells are similar in size and they all have the same borders across the time series. This means that we compare each cell to the exact same cell every year, contrary to using country or sub-national units where borders may change throughout time (Tollefsen, Strand & Buhaug, 2012). On the other hand, the use of grid cells imposes a new problem: the modifiable areal unit problem (MAUP). MAUP is defined as the "sensitivity of analytical results to the definition of units for which data are collected" (Fotheringham & Wong, 1991, p. 1025). This means that using cells being 0.5×0.5 degrees may yield different results than cells being 1×1 degrees or 2×2 degrees. Even the geometric shape of the cells, whether they are squares, circles or triangles, may lead to different results.[8] Furthermore, the borders of the grid cells are arbitrary, such that they are irrespective of country borders or ethnic settlements. This can be beneficial when using variables that are not correlated with its respective country, such as rainfall, but unsuitable when using variables that correlates with the respective country, such as democracy or economic development. This is because, as I will get back to later, most regression estimators treat all observations as independent of each other.

Another challenge with using grid cells is that it requires high spatial precision in the data. There is no point in using grid cells if the data are not precise enough. As I will discuss later, this is not a problem for remotely sensed data, but for variables coded based on news reports, such as conflicts, the precision may vary.

In an attempt to create a unified spatial data structure, Tollefsen et al. (2012) released the PRIO-GRID dataset in 2012. In this dataset the whole world is divided into grid cells being 0.5 decimal degrees in longitude and latitude (Tollefsen et al., 2012, p. 367). This corresponds to approximately 50×50 km at the equator. Each cell is further assigned to its belonging country. In cases where grid cells exceed country borders the cell is assigned to the country containing the largest part of the cell (Tollefsen et al., 2012, p. 368). For simplicity, I choose to use this spatial resolution as both conflict data and various control variables are computed based on these grid cells. However, I keep in mind the MAUP and I run robustness tests

[8] The size of the cells is often referred to as the *scale effect*, whereas the shape of the cells are referred to as the *zoning effect* (Fotheringham & Wong, 1991).

with other spatial resolutions in the analysis.

Using PRIO-GRID cells, Sub-Saharan Africa consists of 8,410 grid cells. By focusing on the time-period between 1989-2014 my dataset contains $8,410 \times 26 = 218,660$ possible observations in total.

4.1 Dependent Variable: Communal Conflict

The dependent variable in this thesis is *communal conflict*. The Uppsala Conflict Data Program (UCDP) is one of the world's main providers of data on organised violence and it is the oldest ongoing data collection project for civil war (European Commission, 2018). UCDP has developed various datasets based on different actors, conflict issues and spatial resolutions. The possibility to combine these datasets by identifying relevant conflicts and retrieving the respective locations of the conflicts makes the data well suited to use in this thesis.

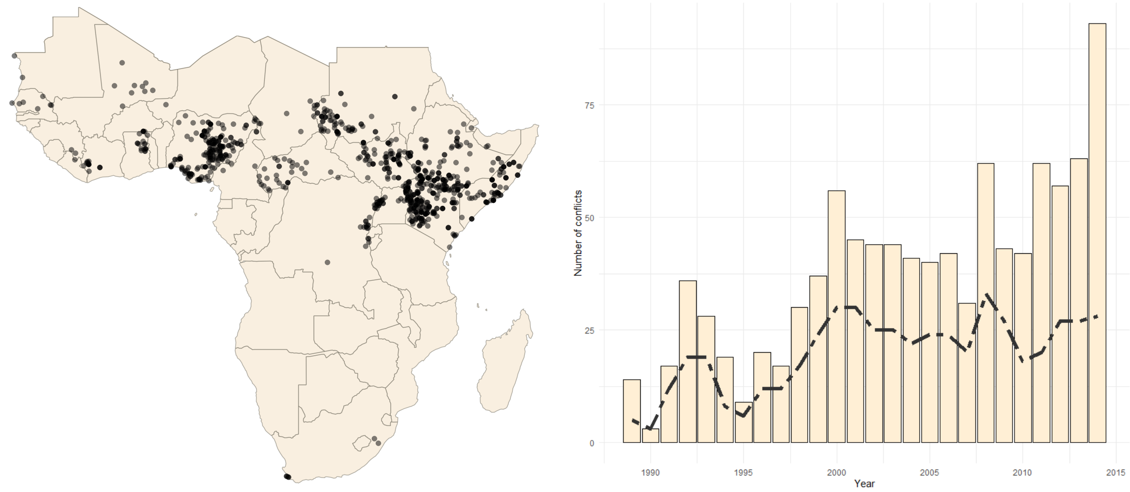
To identify communal conflicts, I rely on the UCDP Non-State Conflict Dataset v.19.1 (Sundberg, Eck & Kreutz, 2012; Pettersson, Högbladh & Öberg, 2019). This dataset contains information on all non-state conflicts in the world between 1989 and 2018. A non-state conflict is here defined as "the use of armed force between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year." (Pettersson et al., 2019). In this definition "armed force" means the use of arms, which can be both manufactured weapons, but also non-technological weapons such as sticks and stones. To be included in the dataset a conflict needs to exceed the threshold of 25 deaths directly related to the use of armed force between the warring groups. For example, this means a conflict causing 50 direct deaths in 2004 and 24 deaths in 2005, will only be coded as a conflict in 2004.

The UCDP Non-State Conflict Dataset divides non-state conflicts into three subcategories depending on the level of organisation of the warring parties (Sundberg et al., 2012, p. 353). The first level contains actors with a high level of organisation, such as paramilitary groups, guerrilla groups or competing rebel groups. The second level includes fighting between political actors that are not permanently organised for combat. The third level captures fighting between identity groups, such as ethnic, clan, religious, national or tribal groups. This third level is what is commonly referred to as communal conflicts and coincides with my definition outlined in part 3.1.2.

A shortcoming with the UCDP Non-State Conflict Dataset relates to its spatial precision. The dataset only contains information on the country in which the conflict occurred. Therefore, I combine this data with data from the UCDP Georeferenced Event Dataset Global v.19.1 (UCDP-GED). UCDP-GED contains geo-coded and fine-grained information on all conflicts (Sundberg & Melander, 2013). All geo-coded conflicts are further attached to its corresponding PRIO-GRID cell. Since both the UCDP Non-State Conflict Dataset and the UCDP-GED contains the same unique conflict-id's, I merge these datasets together to identify communal conflicts and their respective grid cells.

While conflicts embedded in the UCDP Non-State Conflict Dataset need to exceed the threshold of 25 battle deaths each year, the threshold for inclusion in UCDP-GED is that the conflict exceeded 25 battle deaths in at least one year. Thus, the hypothetical example from above where a conflict causing 50 direct deaths in 2004 and 24 deaths in 2005, will be coded as a conflict in both 2004 and 2005 in the UCDP-GED. Since communal conflicts are relatively rare and often small-scaled, I proceed with the latter definition where all years with at least one battle death are noted as conflict-years — as long as the conflict surpasses the threshold of 25 battle deaths in at least one year.

Based on this data, I create a variable *conflict incidence*. This variable is binary, taking the value (1) if a cell experienced a communal conflict in a given year, and (0) otherwise. Some scholars argue that if we are interested in knowing the causes of war, we should only measure conflict onsets — as using conflict incidence will tell us what prolongs a conflict, but not what starts it (van Weezel, 2015). However, scholars disagree as to whether the onset of conflict has a different causation than continuation of conflict (Gleditsch, Wallensteen, Eriksson, Sollenberg & Strand, 2002, p. 620). In contrast to state-based conflicts, communal conflicts tend to be sporadic, short-lived with a few days of violent clashes (Pettersson, 2010). Moreover, identifying the onset of these small-scaled conflicts may be a challenging task and often the spatial precision of the earliest noted event in the UCDP-GED



(a) By location (b) By year

Figure 4.1: Communal Conflicts in Sub-Saharan Africa (1989-2014)

dataset has a low spatial precision. As a result of this I find it suitable to carry out the main analysis with conflict incidence as the dependent variable.

4.1.1 Descriptive Statistics

Between 1989 and 2014 there were 995 communal conflict incidences in Sub-Saharan Africa. The geo-locations of the different conflict events are shown in Figure 4.1a. From this map it is evident that most conflict events took place either in Nigeria or in the border region between Kenya, Uganda, Ethiopia and South Sudan. Almost no communal conflicts took place south of Tanzania. A likely reason for this is that most communal conflicts are farmer-herders conflicts and these tend to erupt in the semi-arid regions where most farmers and herders live.

As briefly touched upon in the previous section, the geo-precision in the data may be a challenge when using grid cells. Of the 995 conflict observations in this dataset, 422 (42.4%) are coded on the exact location. 187 (18.8%) observations are coded within a 25km radius of the exact location. Since grid cells are approximately 50km × 50km these would mostly be coded in the right cell, but may be coded in one of the neighbouring cells. For the remaining conflict observations (38.8%) only the second order administrative unit or higher administrative levels were known.[9] Since this may be problematic I run a robustness test only including the conflict

[9]For a detailed description of the spatial resolution of conflict observations see Table A.1 in appendix.

observations where the exact coordinates were known.

Figure 4.1b shows the time trend of communal conflicts. The bars show the number of cells experiencing communal conflict each year, while the dashed line shows the total number of unique conflicts per year. If there are large differences between the bar and the dashed line, as for 2014, it means the conflicts are spatially spread out. From the bar chart it may seem as there is a solid time trend for communal conflicts with more conflicts taking place in recent years, when in fact the main explanation is that most communal conflicts are spatially confined, while some conflicts are more spread out and hence influence the bar chart severely.

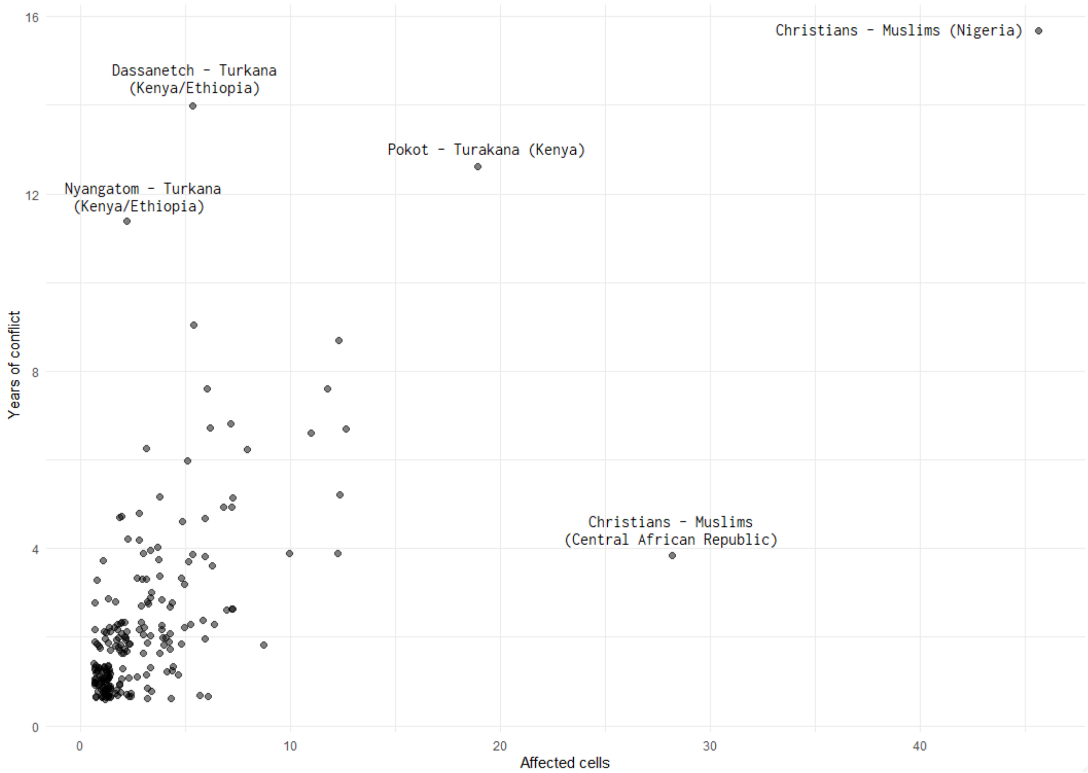


Figure 4.2: Scatter plot of communal conflicts

To examine whether the use of disaggregated spatial units is beneficial, I scatter all unique conflicts by their spatial and temporal length in Figure 4.2. From this figure it is evident that most communal conflicts tend to be relatively confined both in terms of spatial and temporal length. However, a few conflicts are spread out and comprise large parts of the conflict observations. For instance, the conflict between Christians and Muslims in Nigeria represents 71 cell-years, and hence 7.1% of all conflict observations in the dataset. While there are 215 unique conflicts, the ten largest conflicts make up over 30% of all conflict observations. This may be problematic as the results from the regression analysis may be driven by these few conflicts. I discuss this problematic aspect more in detail in the next chapter and in the analysis I run a robustness test excluding the conflict between Christians and Muslims in Nigeria.

4.2 Explanatory Variable: Drought

The explanatory variable in this thesis is *drought*. As discussed in the previous chapters, SPEI and its predecessor SPI, have been the most widely used drought indices in the climate-conflict literature. However, both the validity and reliability of SPEI as a drought measure have been questioned (Homdee, Pongput & Kanae, 2016). How certain are we that SPEI measures the physical impacts commonly assumed to carry conflict potential?

As I will discuss in this section, there are clear technical benefits of using SPEI, but the operational validity can be challenged. Does SPEI measure what we want it to measure? In order to examine this further, I test whether using other operationalisations of drought yields different results. As previously outlined, I rely on three different indicators: SPEI, NDVI and EM-DAT. These are both assumed to capture different definitions of drought (meteorological, agricultural and socio-economic), as well as different stages in the the causal chain of a drought.

4.2.1 SPEI - Standardized Precipitation Evapotranspiration Index

The most common way to measure drought in the climate-conflict literature has been to use measurements of precipitation and temperature (see, e.g., Hendrix & Salehyan, 2012; Miguel et al., 2004). Researchers are often interested in changes and anomalies in climate variability and therefore often standardise the measurement, meaning they subtract the mean and divide by the standard deviation. For a long time, the Standardized Precipitation Index (SPI) was the most used way of measuring drought on the climate-conflict nexus. SPI measures the standardised precipitation for each month by using the historical mean of the time-series and its standard deviation (McKee, Doesken & Kleist, 1993). Thus, SPI only measures precipitation anomalies. Large negative precipitation anomalies can be characterised as droughts, whereas large positive anomalies can be characterised as floods.

The last couple of years researchers have started to use the Standardized Precipitation Evapotranspiration Index (SPEI) instead of SPI (see e.g., Döring, 2020; Harari & Ferrara, 2018; von Uexkull et al., 2016). While SPI only focuses on the level of precipitation anomalies, SPEI includes temperature to the equation through an estimate of potential evapotranspiration (PET) (Vicente-Serrano, Beguería & López-Moreno, 2010). By including potential evapotranspiration SPEI arguably measures droughts better than the SPI, as it captures more than just the lack of rainfall (Vestby, 2018, p. 13). Similarly to SPI, large negative SPEI values can be interpreted as droughts, whereas large positive values can be perceived as floods. A usual interpretation is that SPEI values between -1 and 1 can be characterised as normal weather. SPEI values between -1 and -1.5 are dry, between -1.5 and -2 are considered moderately dry and below -2 are considered extremely dry. Usually researchers put a threshold at SPEI values < -1.5 to distinguish between droughts and non-droughts (see e.g. Nordkvelle et al., 2017; Vestby, 2018). Both SPI and SPEI can be used with different lengths of intervals. For instance, SPEI-1 is calculated based on the average climate variability during one month, while SPEI-6 uses the average over a given six-month interval, rather than only one month (Vestby, 2018, p. 13). This makes the measurement flexible. It is often preferable to measure SPEI during growing season or rainy season, as a drought arguably create more damage on crops and supposedly carry a larger conflict potential during these seasons.

There are some obvious strengths and weaknesses with using SPEI. One benefit is that SPEI is based on meteorological data. Thus, we can assume that SPEI is exogenous to both the dependent variable, which is conflict, but also the other independent variables such as poverty and population. Second, a technical benefit of using SPEI is that it is a function of two independent variables (precipitation and evapotranspiration) (Vestby, 2018, p. 124). As the central limit theorem states that any sum of two or more independent variables will converge towards a normal distribution, we can expect SPEI to be normally distributed. Using a normally distributed variable is preferable as it reduces the changes of large outliers affecting the results. Further in order to best identify a causal effect, the independent variable must be both exogenous and random (Vestby, 2019, p. 119).

A small caveat with using SPEI is outlined by Vestby (2018, p. 13). Since temperature is increasing slowly over time, and SPEI uses the average over the whole time-series, this may lead to higher values (wet and cold) being more likely early in long time-series and low values (dry and hot) being more likely later in the time-series. One way to solve this problem is to use a moving average instead of the mean.

The largest problem with using SPEI is that we do not know how well it measures actual droughts. Although each unit is compared to its own mean and thus the standard deviation gives approximately same hypothesised effect of meteorological drought to each unit, we do not know the actual effect on the ground. For instance, in some areas 1.5 standard deviations less rainfall may not have any effect on agriculture or water scarcity. While other places only a small deviation may have devastating impacts. Similarly, neither SPI, nor SPEI, consider the intensity of precipitation or the impacts on streamflow, runoff or water availability in the area of interest (Keyantash et al., 2018).

Another limitation with SPEI is that the results are sensitive to the method of calculating the potential evapotranspiration. The original SPEI was calculated with the Thorntwaite method to calculate potential evapotranspiration. This method only includes temperature as a proxy for potential evapotranspiration (Trenberth et al., 2014). Newer versions of SPEI are calculated based on the Penman-Monteith method which also incorporates effects of wind, humidity, and solar and longwave radiation (Trenberth et al., 2014, p. 18). Thus, the Penman-Monteith method is generally considered better than the Thorntwaite method. However, since the Thorntwaite method only considers temperature to calculate PET, it requires less data than the Penman-Monteith method and are therefore often used when studying areas where data availability is limited to temperature (Shiru, Shahid, Alias & Chung, 2018). Hence, the validity of SPEI as a drought measure is higher with the Penman-Monteith method, but the reliability is higher using the Thorntwaite method.

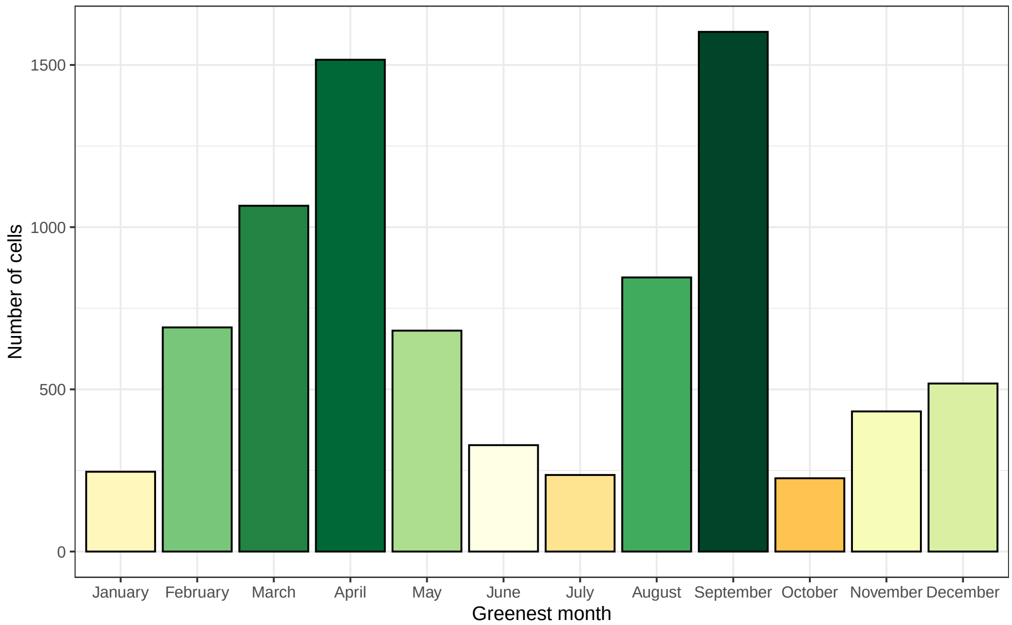
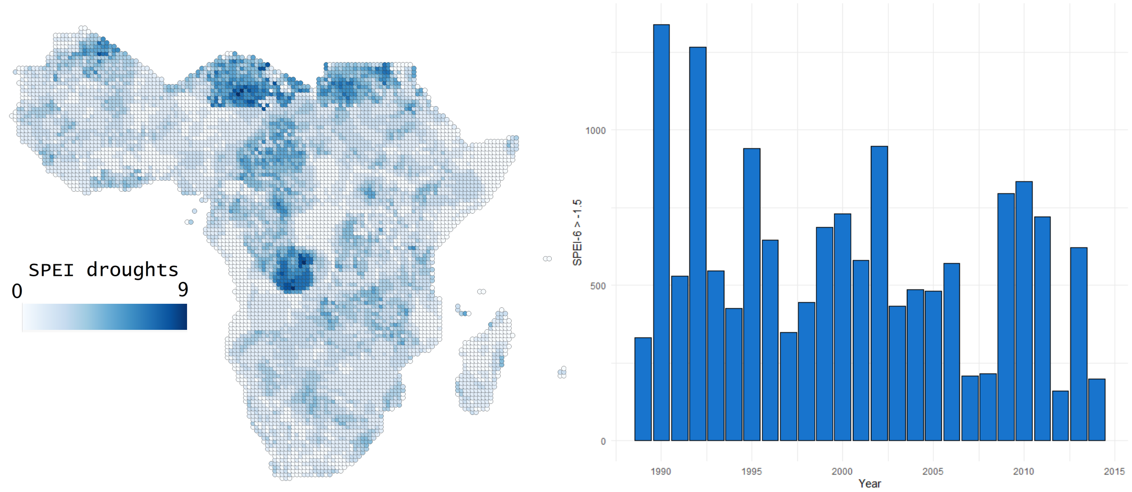


Figure 4.3: Greenest month

Operationalisation and descriptive statistics

I retrieve raw SPEI data from the unpublished PRIO-GRID v.3.0 currently in development (Tollefsen, Vestby, Landsverk, Larsen & Bahgat, 2020). This SPEI variable is based on precipitation data from The Climate Research Unit (CRU) version 4.03 and is computed using the Penman-Monteith method. The SPEI variable contains data for every month. However, using yearly averages are not beneficial as a year may contain both droughts and floods and hence this may mask a drought. Instead researchers often use SPEI data at the start of rainy season, end of rainy season or during growing season. In this thesis I offer a new way to determine which month to focus on. Since I use NDVI as a measurement (see more in section 4.2.2) I am able to identify the typically greenest month within each cell.[10] This is suitable as the greenest month is arguably one of the most important months for agriculture and grazing. Similarly, I focus on the greenest month when measuring NDVI. Hence, the comparison between SPEI and NDVI will not only be of the same year, but also based on the same month. The distribution of the greenest months is shown in Figure 4.3. From this figure it is evident that most cells have either September or April as their greenest months. Whereas fewest cells have October, July or January

[10] I define the greenest month as the month with the highest average NDVI value across the time-series within each cell. I elaborate more on this in section 4.2.2.



(a) By cell (b) By year

Figure 4.4: SPEI droughts

as their greenest months.

PRIO-GRID v.3.0 contains four different SPEI variables: SPEI-1, SPEI-3, SPEI-6 and SPEI-12. I choose to focus on SPEI-6 for two different reasons. First, it is the most commonly used SPEI variable used by researchers on the climate-conflict nexus. Second, it is, together with SPEI-3, the SPEI variable with the highest correlation with NDVI data.[11] To make the variable binary, I impose a cutoff at SPEI values < -1.5. This is considered a "severe drought" or "moderate drought", and gives approximately the same proportion of droughts as NDVI and EM-DAT.[12] Figure 4.4a shows a map of the most frequent SPEI drought locations. From this map it is evident that most SPEI droughts have appeared in the border region between Niger and Chad, as well as in the western part of DR Congo. Figure 4.4b shows the time trend. Interestingly, there seems to be a weak decline in the number of SPEI droughts in recent years. This is counter intuitive. Since SPEI uses the long-term mean from the whole time-series, and the mean temperature has risen across this time period, we would expect to see more SPEI droughts in recent years. Moreover, evidence from Afrobarometer (2018) also suggests that there would be more droughts later years.

A limitation with this SPEI variable is that it is based on precipitation measures

[11] Correlation between SPEI and NDVI: SPEI-1 (11.9%), SPEI-3 (21.9%, SPEI-6 (20.3%) and SPEI-12 (18.4%).

[12] SPEI (7.17%), NDVI (8.53%) and EM-DAT (8.23%).

from weather stations on the ground. According to Trenberth et al. (2014) precipitation data from CRU in the 2000's are retrieved from approximately 2400 climate stations across the world. When my dataset contains 8410 cells in Sub-Saharan Africa, it means that precipitation is not physically measured in each cell. Therefore, CRU uses interpolation methods to create data points. Thus, both the validity and reliability of the precipitation data may be questionable. However, Trenberth et al. (2014) argue that more stations do not necessarily mean improved coverage if the extra stations are all in the same area. Additionally, the rationale behind creating the CRU dataset has been that fewer, more homogeneous, records may provide more a reliable time series.

The main reason for using SPEI to measure drought is due to its exogenous feature. Rainfall may affect conflict, but conflict does not affect rainfall. However, when precipitation is measured on the ground by weather stations, the precipitation *estimates* may not be exogenous to conflict. Schultz and Mankin (2019) find that during conflict, climate stations were destroyed and other stopped reporting. Moreover, they find that the density of weather stations in Sub-Saharan Africa are negatively correlated with the conflict risk. Meaning that areas with higher conflict risk have fewer weather stations. Another problematic aspect they highlight is that one of CRU's requirements is that weather stations have a record of coverage between 1961 and 1990, as this is used as a baseline comparison. The problem with this criterion is that it makes it hard to establish new stations. As a result of this there has been a notable decline in the number of CRU climate stations over the last years (Schultz & Mankin, 2019).

Thus, the fact that precipitation and rainfall are exogenous to conflict does not necessarily mean that the precipitation *estimates* are exogenous to conflict. One way to avoid that the data are influenced by political and social conditions on the ground, such as conflict, is by using satellite-based measures.

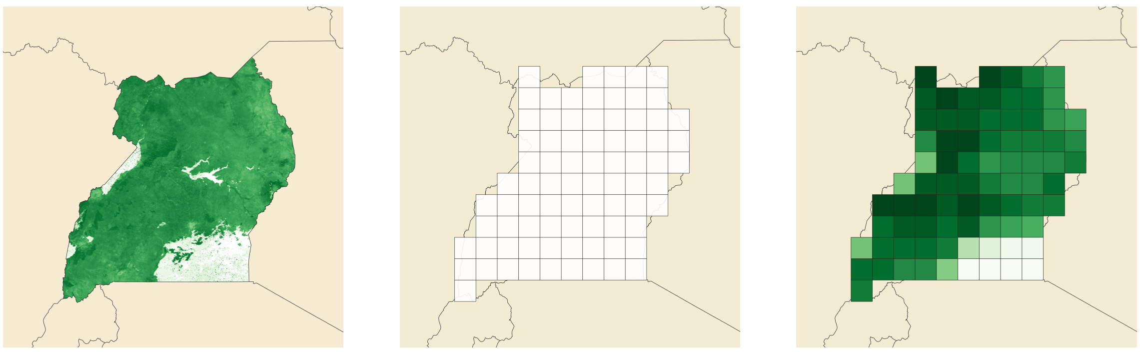
4.2.2 NDVI - Normalized Differences Vegetation Index

The second way I measure drought is by using satellite data on vegetation through the Normalized Difference Vegetation Index (NDVI). Early tools of NDVI were developed and computed by Rouse, Haas and Deering (1973) to measure the quality and quantity of vegetation. Thus, measuring drought is not its original intent and this should be clearly stated. NDVI is computed based on remote sensing, meaning data is acquired without making physical contact with the object. Usually, as for NDVI, this means that data are retrieved by using satellites (Lillesand, Kiefer & Chipman, 2004).

NDVI is based on two indicators: Near infrared light (NIR) and infrared light (Red) (Gamon et al., 1995). Put simply, during photosynthesis plants both absorb and reflect different wavebands of light. Green and healthy vegetation absorbs most of the infrared light and reflect more of the near infrared light. Conversely, if vegetation is less healthy and sparse there is a decrease in the near infrared reflectance and an increase in the infrared reflectance as there is less chlorophyll to absorb the red light (Dick, 2009). The NDVI is calculated based on the reflectance of these two wavebands with the following formula:

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

If the reflectance of NIR is high and Red is low, the NDVI will tend towards 1. For healthy vegetation the NDVI value will vary between 0.7 and 1. While stressed vegetation will converge towards 0. Negative NDVI values indicates surfaces such as water, ice or snow (Lillesand et al., 2004).



(a) NDVI data (b) PRIO-GRID cells (c) Merged data

Figure 4.5: Merging process

Operationalisation and descriptive statistics

I retrieve NDVI data from Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Indices (MOD13A3) Version 6 (Didan, Munoz, Solano &

Huete, 2015). The NDVI data are provided monthly at a 1-kilometre spatial resolution. I aggregate the data to PRIO-GRID cells using geographical information systems (GIS). A graphical example of the merging process can be seen in Figure 4.5 with Uganda as an example.

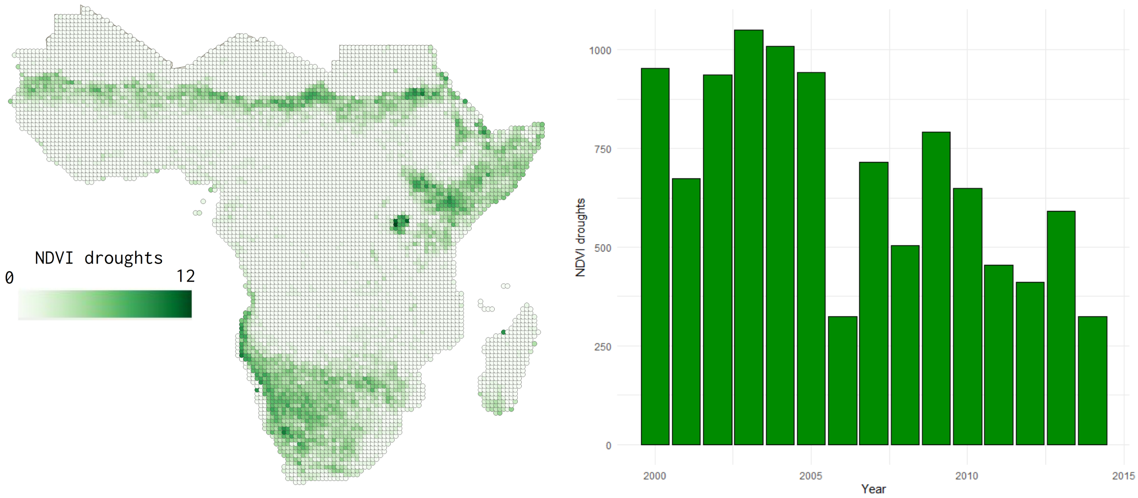
Similar to SPEI, NDVI contains data for every month. I identify the greenest month, the month within each cell containing the highest average NDVI value, and use this month as a representation of the NDVI value for the given year. Thus, I use the same month for SPEI and NDVI values, which should be a good way of comparing the different measures. Based on the NDVI values for the greenest month in each cell, I calculate an NDVI anomaly for each cell with the following formula:

$$NDVIanomaly = NDVI_{it} - \mu NDVI_i$$

In this formula $NDVI_{it}$ represents the NDVI value for the greenest month in cell i in year t. While $\mu NDVI_i$ is the average NDVI value for the greenest month in cell i across the whole time series. Comparing the same month, for instance September in 2009 against September in 2010, is beneficial as vegetation varies greatly between seasons in Africa. Failing to take this into account could lead to false inferences, which happened in December 2019 when newspapers reported that the Victoria Falls were drying up, when in fact they compared pictures of different months (Dube, 2019)

In contrast to SPEI, there are no universal cutoffs for determining whether there has been a drought according to NDVI. A possible cutoff could be saying there is a drought if vegetation quality is lower than one standard deviation of the mean, meaning *NDVI anomaly* $< \mu - 1\sigma$. However, the standard deviation of NDVI is 0.26. Using this cutoff would mean that areas with a higher mean value of NDVI (say 0.8) have a higher likelihood of experiencing drought, whereas areas with a lower mean value (say 0.2) are not able to experience a drought, since, practically, the NDVI value cannot go from positive to negative.

A second possibility is to create the cutoff based on each cell's own standard deviation. This is not feasible either, since we do not have the long-term mean. Applying this cutoff would result in all cells experience drought at some point, even cells with limited vegetational variety.



(a) By cell (b) By year

Figure 4.6: NDVI droughts

Instead, I choose to set the cutoff at 10% below the cell mean. This makes the cutoff relative to each cell. Moreover, this means that cells with a mean NDVI value of 0.8 would require an NDVI value lower than 0.72, whereas a cell with mean value of 0.2 would require a much lower deviation (<0.18). Applying this cutoff results in approximately the same proportion of droughts for all three indicators.

Figure 4.6a shows a map of areas experiencing the most NDVI droughts. It is evident that the number of NDVI droughts are spatially clustered and not spread out in the same way as SPEI droughts. The areas experiencing NDVI droughts seem to be the Sahel region, Eastern Africa, and the South-Western parts of Africa surrounding the Kalahari Desert. One reason for this is that the changes in the raw NDVI values seem to be dependent on certain land types. This also makes sense as there is little variation in the vegetation quality in the rainforest, whereas in areas vulnerable to climatic hazards, such as cropland and grassland, there will be a higher variation.[13]

In Figure 4.6b I show the distribution of the number of NDVI droughts per year. Similar to the SPEI droughts there seems to be a negative trend with fewer droughts in the recent years.

There are some clear benefits of using NDVI or other satellite data. One important aspect is that we know that there has been a change in vegetation on the

[13]For a map of the NDVI variation for each cell, see Figure A.2 in the appendix.

ground in the respective cell-year. Additionally, remotely sensed data does not require physical climate stations, and both the spatial and temporal resolution can be made much higher and better. Therefore, some researchers have argued that satellite data should be an integral part of monitoring drought (Wan, Wang & Li, 2004, p. 62).

Although the benefits of using remotely sensed data are evident, there are clearly some limitations with using NDVI to measure drought. The first potential problem is endogeneity. When conducting quantitative analysis and drawing causal inferences the assumption of exogeneity is pivotal (King, Keohane & Verba, 1994). In this case, exogeneity means that the drought may affect the conflict, but the conflict may not affect the drought. This assumption is assumed to be fulfilled when using meteorological data such as SPEI, although, as discussed previously, this might not always be the case. When using NDVI, this is assumption is not fulfilled as conflict may affect the level and quality of vegetation.

A second caveat is that the causes of these NDVI anomalies are unknown and hard to distinguish. As NDVI only measures the "level of greenness", it does not solely measure drought. Deforestation, for instance, imposes a severe limitation on using NDVI as a measurement of drought.

There exists a lot more complicated and advanced indices based on vegetation which arguably measures droughts better than NDVI. The reason for using NDVI is its established robustness and that it is relatively easy to understand. Moreover, the NDVI is used extensively in global, regional and local studies and is by far the most commonly used vegetation index.

4.2.3 EM-DAT - Emergency Events Database

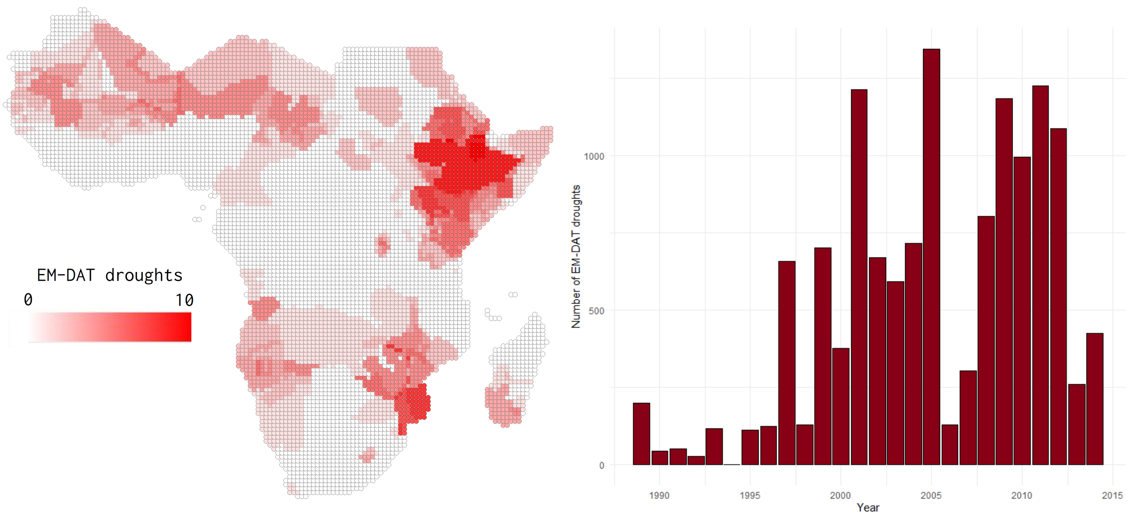
The third way I measure drought in this thesis is by measuring events in which there have been explicitly reported an experienced drought. The Emergency Events Database (EM-DAT) is a global database created by The Centre for Research on the Epidemiology of Disasters (CRED) and contains data and information on natural and technological disasters since 1900 (Guha-Sapir et al., 2016). For a disaster to be included in the EM-DAT database, at least one of the following four criteria must be fulfilled: 10 or more people dead; 100 or more people affected; the declaration of a state of emergency; or a call for international assistance (Guha-Sapir et al., 2016). The database is compiled of information from various sources, including NGOs, UN agencies, research institutes and press agencies.

The original database contains only information on the country-level. However, a recent extension to the database has been created by Rosvold and Buhaug (2020) where they have connected each disaster to the affected first level administrative unit(s) within each respective country. This provides the database with a much more precise geolocation and makes it more suitable to use in spatially disaggregated studies.

There are certain benefits of using EM-DAT to measure droughts. First, the database only includes droughts which have made a clear impact on the livelihood of people, as opposed to using SPEI. Unlike NDVI, by using EM-DAT we know the physical drought not only degraded the vegetation, but also affected the people in the respective area. Ultimately, it is through the deterioration of living conditions we expect drought to contain a conflict potential (Homer-Dixon, 1999), not through the lack of precipitation or vegetation.

However, there are some clear limitations with using EM-DAT to measure droughts. First is the potential problem of endogeneity. When using crisis data such as EM-DAT the assumption of exogeneity is no longer upheld. Particularly, since one of the thresholds to be included in the EM-DAT database is a minimum of 10 fatalities, violent conflict may bias this. Often it is impossible to determine whether a death was a direct consequence of a conflict or a drought. What if there was a drought creating water scarcity, but a violent conflict enhanced this scarcity or deteriorated the only water left? The same problem applies to the threshold of '100 people affected'. It is hard to determine whether the people were directly affected by a drought or if they were affected by a nearby conflict or both.

Another potential problem is that there may be systematic reporting differences between countries. For instance, some countries may declare a state of emergency or call for international humanitarian assistance more frequent than other states. This means that many drought events causing monetary losses without exceeding local response capacities are not recorded in the database. Thus, this database is biased toward economically catastrophic and deadly events (Gall, Borden & Cutter, 2009,



(a) By cell (b) By year

Figure 4.7: EM-DAT droughts

p. 804).

Similarly, larger areas both in terms of area and population may experience a higher frequency of droughts than smaller areas. This is important to remember when using grid cells as units in the analysis, because EM-DAT droughts are not measured on the grid cell level, but rather on the first order administrative level. This means that all cells belonging to a region experiencing drought will be coded as experiencing drought. Thus, the results may be biased such that some grid cells which have not experienced drought will be coded as if they had. Hence, the likelihood of experiencing a drought would not be the same for all cells.

Another challenge is the role of temporal bias. For instance, the growing wealth in the recent years among countries means there are more money to be lost during droughts. Hence, calls for international assistance may be more frequent in recent years. On the other hand, more areas are no longer solely dependent on rainfed agriculture and use irrigated agriculture instead. Thus, more areas are more drought resilient now than before. However, this temporal aspect may create biases in both directions, and it is impossible to determine the counterfactual: Would this drought have been reported as a disaster if it happened ten years earlier or ten years later?

Descriptive statistics

Between 1989 and 2014 there were a total of 167 unique droughts in Sub-Saharan Africa. The worst drought in terms of causalities took place in Somalia in 2010, where there was estimated more than 20.000 deaths as a direct result of the drought (Guha-Sapir et al., 2016).

Contrary to SPEI and NDVI, EM-DAT droughts see a different time trend with droughts being more frequent in later years. This is shown in Figure 4.7b. As discussed above, the exact reason behind this trend is difficult to determine. It might be the fact that there actually have been more droughts in recent years, but it might also be due to reporting biases or economies being more vulnerable to drought in later years. Another peculiar trend is evident from the map in Figure 4.7a. There are reported several droughts in Kenya and Mozambique between 1989 and 2014, but none in Tanzania. This may suggest that there are reporting biases among countries and this should therefore be considered when conducting the statistical analysis.

4.2.4 Relationship Between SPEI, NDVI and EM-DAT

The main interest in comparing different drought indices is the assumption that SPEI does not measure the *impact* of the lack of precipitation and the potential evapotranspiration. Thus, we know that there has been less precipitation than normal, but we cannot for certain tell whether the drought has impacted livelihood or water availability on the ground. Further challenges with using SPEI relates to whereas the validity of SPEI can be made relatively high (with using the Penman-Monteith method of calculating PET, rather than the Thornthwaite method), it comes at the expense of the reliability, as it requires more precise data (e.g. wind speed) which in many places are of questionable quality or do not exist at all (Trenberth et al., 2014). Furthermore, the reliability of in situ climate stations may also be questioned. Particularly, as there exist far more cells than climate stations on the ground, how can we for certain tell that the interpolation techniques used by the Climate Research Unit (CRU) are good enough to provide an accurate picture of precipitation and drought in Sub-Saharan Africa? Moreover, these weather stations may also be affected by conflicts themselves.

Thus, the idea is to test whether the relationship between drought and conflict varies depending in the way we measure drought. If there are large discrepancies between the results constructed by SPEI, NDVI and EM-DAT it may suggest that other measurements than SPEI would be beneficial in exploring the causal mechanisms between drought and conflict in the future.

Correlation

Although the three drought indicators are assumed to grasp different drought-related aspects, they are still assumed to be relatively correlated. Since I have dichotomised SPEI and NDVI, I run two separate correlation analyses showing the correlation between the indices with both the binary and the continuous values. This is done as we often lose a lot of valuable information by dichotomising variables.

In Table 4.1 I show two correlation analyses using Pearson's R. In Table 4.1a both SPEI and NDVI are used as continuous variables. The table shows that the correlation between SPEI and NDVI is 20%. This means that there seems to be a moderate amount of NDVI droughts being captured by the SPEI variable.[14] The correlations between SPEI and EM-DAT (-2.8%), and NDVI and EM-DAT (-7.1%) are both negative. This is quite counter intuitive. One reason for this may be that it takes time for the meteorological and agricultural drought to cause socio-economic impacts. To test this, I lag both the SPEI and the NDVI variables one year. Now both correlation coefficients are positive, being 4.4% and 2.9% respectively. However, the correlations are still relatively low. Another reason for the low correlation may be that the administrative units are too large, and that we see a lot of local variations within each cell, which is not reflected in the EM-DAT data.

In Table 4.1b I show the correlations where all three variables are made binary. This is how I treat the variables in the analysis. Now the correlation between SPEI and NDVI has diminished. Imposing new cutoffs (SPEI < -1 or NDVI $< -15\%$) does not yield substantial changes. This points to the fact that by dichotomising variables we lose a lot of valuable information. The correlation between SPEI and EM-DAT are still very low (1.9%), but the correlation between NDVI and EM-DAT are now positive and higher than before (7.7%). Thus, although I run my

[14] For other operationalisations of SPEI, the correlation with NDVI is the following: SPEI-1 (11,9%), SPEI-3 (21,9%) and SPEI-12 (18,4%).

Table 4.1: Correlations: Pearson's R

(a) Continuous SPEI and NDVI

	SPEI	NDVI
SPEI	-	-
NDVI	0.2033***	-
EM-DAT	-0.0282***	-0.0708***

(b) Binary SPEI and NDVI

	SPEI	NDVI
SPEI	-	-
NDVI	-0.0049	-
EM-DAT	0.0194***	0.0766***

main model with binary variables, this suggests I should run a robustness test with continuous variables as this may yield different results.

The low correlation is also evident when making frequency plots of the different drought indices based on countries as seen from Figure 4.8. In this figure I use the binary indicators and divide the number of droughts by the number of observations for each country. Otherwise large countries such as DR Congo and Sudan would dominate the picture since they have more cells and more observations. Hence, they have a greater likelihood of containing cells experiencing droughts.

From Figure 4.8 it is evident that the three indices do not correlate on a high level. Chad, for instance, see the most SPEI droughts, whereas it experiences relatively few NDVI and EM-DAT droughts. The same is true for Botswana and Namibia, the countries experiencing the most NDVI droughts, and Ethiopia, the country seeing most EM-DAT droughts. Kenya, for instance, see quite frequent EM-DAT and NDVI droughts, but relatively few SPEI droughts. In countries with few observations, such as Cape Verde, Seychelles, Togo and South-Sudan, there seems to be a high correlation among the indices with all three reporting very few droughts.

Case testing: Somalia

One explanation why the correlations between the three indices are relatively low, may be that NDVI droughts are caused by deforestation rather than actual droughts. This is highly problematic and may lead to false inferences. To investigate this possible effect, I use a case from Somalia.

During the regime of Al-Shabaab in the Kismayo district in Somalia, there was a rapid expansion in wood charcoal production. As charcoal is produced by wood, deforestation is a necessary condition in the production. Rembold, Oduori, Gadain and Toselli (2013) find a large increase in the charcoal production between 2006

Figure 4.8: Droughts by country

and 2012 when Al-Shabaab controlled the Southern Somalia. Additionally, burning trees to produce charcoal also involves using biomass, such as dry grass, to start the fire (Robinson, 1988, cited in Rembold et al., 2013).

To examine these potential problems, I run a quick check of the NDVI and SPEI values in the Southern Somalia for all years between 2000 and 2014 and highlight the years when Al-Shabaab controlled the region. Figure 4.9 shows a graph containing the 15 cells of the Lower Juba region in Southern Somalia and their NDVI and SPEI values between 2000 and 2014. The years in which Al-Shabaab controlled the region and supposedly deforested large parts of the region are marked in purple. From the graph it is evident that the years between 2008 and 2012 saw a rapid decline in NDVI values, whereas the SPEI values have mostly been positive during the same period.

However, to make the situation more complex, Somalia was part of one of the worst droughts in recent years with between 20.000 and 200.000 casualties between

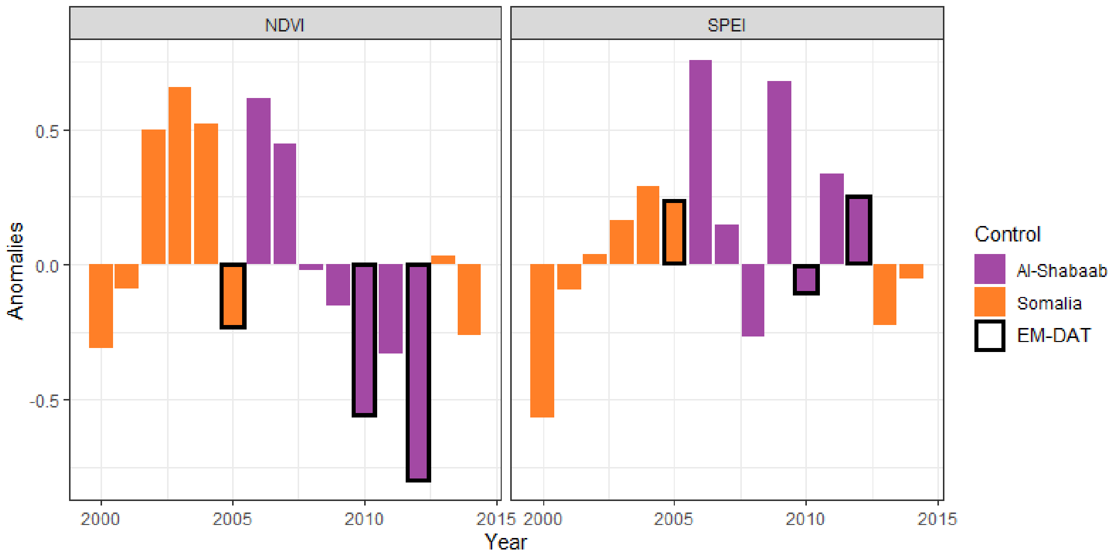


Figure 4.9: Droughts in Southern Somalia

2010 and 2012 (Guha-Sapir et al., 2016). Therefore, I have also embedded EM-DAT into the graph in Figure 4.9. The three years where all fifteen cells in the Lower Juba region experienced a socio-economic drought (2005, 2010 and 2012) are outlined with a bold frame. Thus, it might not be the deforestation causing the large NDVI anomalies, but rather an actual drought. This points to a problematic aspect with using NDVI as a drought measure, but also with large-N statistical studies in general. Without qualitative in-depth studies it is impossible to determine the cause of the vegetation loss for certain.

Figure 4.9 also shows some surprising trends when it comes to the SPEI data. Surprisingly, during the 2010-2012 drought there have actually been more precipitation than normal. This can imply several things. First, this may be a sign that this SPEI variable does not measure actual droughts precisely. However, on the other hand, this may be the result of a combination where the southernmost part of Somalia did not experience a meteorological drought, but the lack of rainfall in other parts of Somalia affected the Southern areas.

Another reason for this may be that the SPEI variable included in this thesis only considers the greenest month and sixth months prior to this. In the Lower Juba region $\frac{8}{15}$ cells have November as their greenest month, $\frac{6}{15}$ cells have June and $\frac{1}{15}$ cell has July. The 2010 drought has been argued to be a result of failed rain during the 2010 Deyr period, which is the minor rainy season of Somalia and usually takes place between October and December (Maxwell & Fitzpatrick, 2012). However, this is barely captured by the SPEI data. Digging deeper into the data, I check the SPEI-6 values for all months in 2010 for these cells. Surprisingly, none of the 15 cells had SPEI-6 values lower than -0.6 in 2010, which corresponds to normal weather.

To further investigate this scenario, I test whether the other operationalisations of SPEI (SPEI-1, SPEI-3 and SPEI-12) show other types of decline. Surprisingly, none of the 15 cells have SPEI-1, SPEI-3, SPEI-6 or SPEI-12 values below -1 between 2009 and 2012.

4.2.5 Spatial Lag of Drought

So far, I have only considered drought as having a direct effect, meaning the effect of drought in cell_{it} may be associated with conflict in cell_{it}. However, drought is a highly complex natural phenomenon and its impacts may often extend outside its defined area (Wilhite & Glantz, 1985, p. 13). For example, the lack of precipitation in one cell could lead to reduced vegetation in neighbouring cells if that cell contains important water sources (e.g. headwaters for rivers) for vegetation in neighbouring areas. To account for this, I compute other operationalisations of drought to be used in my main model.

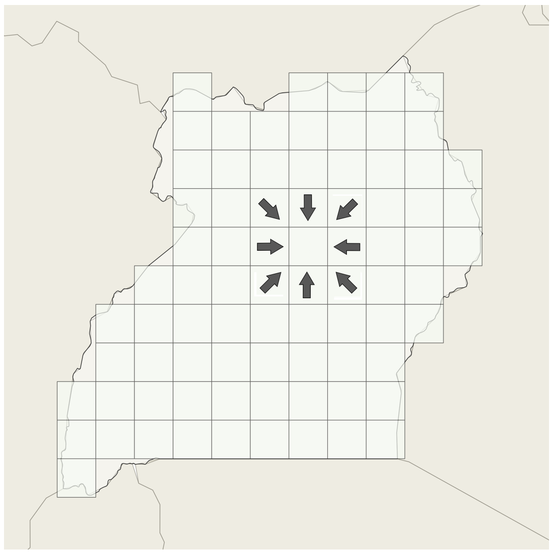


Figure 4.10: Queen's contiguity matrix

First, I create spatial weights of the different drought variables making use of a queen's contiguity weights matrix. This provides each cell with a maximum of eight neighbouring cells. Cells bordering the sea or other outskirts of the Sub-Saharan Africa will naturally have less than eight neighbours. An illustration of this weight matrix is shown with Uganda as an example in Figure 4.10.

Based on these weights I create two new variables: (1) indirect effect and (2) total effect. The indirect effect denotes the proportion of neighbouring cells experiencing droughts. This means if one of the eight neighbouring cells experience a drought in year t, the cell in question will obtain a value of $\frac{1}{8} = 0.125$. If all eight neighbours experience drought the cell will acquire the value $\frac{8}{8} = 1$.[15]

Since the indirect effect only takes into account the neighbouring cells and not the cell in question, I combine this variable with the direct effect to create a variable denoting the *total* effect. In order to make coefficients comparable in the analysis, I compute this variable to range between 0 and 1 by adding the direct and indirect effect together and dividing the sum by two. This is shown in equation 4.1 below:

$$\text{Total effect} = \frac{\text{Direct effect} + \text{indirect effect}}{2} \tag{4.1}$$

If all neighbouring cells and the cell in question experience drought, the total effect will be 1. If all neighbouring cells experience drought but not the main cell, or the opposite, only the cell in question but none of the surrounding cells, the variable will take the value of 0.5. In the dataset there are very few cases where all neighbouring cells see drought, but not the cell in question.[16]

4.3 Control Variables

Although the use of grid cells makes the area of each unit similar, the cells may be inherently different in terms of other contextual factors. In order to make these cells as comparable as possible, and to try to isolate the effect of drought on conflict as best as possible, I apply different control variables. This is particularly important as failing to include variables correlated with both the explanatory variable and the dependent variable, will bias the estimates (King et al., 1994, p. 190). For instance, let us say I do not include population in the analysis. This would bias the estimates as the conflict potential induced by drought is assumed to be higher where there are more people (see e.g. Döring, 2020). Moreover, cells with zero or a low number of inhabitants may experience droughts, but not conflict. Hence, the effect of drought on conflict would likely be underestimated.

[15]Graphical examples can be found in Figure A.4 in the appendix.
[16]Graphical examples can be found in Figure A.5 in the appendix.

On the other hand, however, researchers should not include every possible causal influence in their analysis as this would reduce the efficiency of the estimates (King et al., 1994, pp. 182-183). The more correlated the key causal variable is with the irrelevant independent variable, the less efficient the estimates of the causal effect. Similarly, we should not include a variable that is in part a consequence of the explanatory variable (King et al., 1994, p. 173). One example of this would be to include vegetation levels (NDVI) when measuring the effect of precipitation (SPI/SPEI) on conflict, as the level of vegetation is affected by the level of precipitation. The easiest way to identify relevant control variables is by consulting theory and former literature.

However, what is considered a mediating variable and what should be considered a confounding variable has been subject for debate. This has been particularly contested on the climate-conflict nexus. Whilst what Busby (2018, p. 340) denotes as 'Berkeley scholars' have argued that the effect of climate variability on conflict is mediated by socio-economic factors, i.e. droughts lead to lower development which may contribute to conflict, and hence these should not be included as control variables; 'PRIO-scholars' have argued that these socio-economic conditions make areas less drought resilient in the first place, resulting in a higher baseline probability of experiencing conflict. Thus, there is a question regarding whether drought may affect conflict through the deterioration of socio-economic factors or whether drought may contribute to conflict in areas with lower coping capacity. Since the baseline risk of experiencing communal conflict, as well as the risk of experiencing NDVI and EM-DAT droughts, are clustered within certain geographical areas, it seems favourable to include control variables in this study. The control variables I apply are first and foremost correlated with conflict in general. Some of the variables are, however, particularly associated with communal conflict or the climate-conflict relationship.

Population

First, I consider the potential impact of population and population density. Findings suggest that higher population density makes conflict more likely (Homer-Dixon, 1999; Döring, 2020). To account for population pressure in a given cell during a given year I use the variable *population* retrieved from CISEIN & CIAT (2005) through PRIO-GRID v.2.0 (Tollefsen et al., 2012). This cell-level variable is measured with five-year intervals (1990, 1995, 2000 and 2005) between 1990 and 2005. To get an approximate value of population, I use linear interpolation and extrapolation to create data points for missing years. Furthermore, I log-transform the variable as the social effects of receiving 1.000 immigrants would be different in an area consisting of 1.000 inhabitants and an area consisting of 100.000 inhabitants. Finally, I exclude grid cells highly unlikely to experience conflict. Tollefsen and Buhaug (2015) exclude cells with a population density < 1 per km^2 and cells with < 100km^2 as there is a low chance of seeing conflict in these cells. Applying a similar threshold in this thesis, however, leads to the exclusion of eight conflict observations.[17] Instead I exclude cells with < 200 people and < 100km^2. This returns a valid sample of 7,986 cells in Sub-Saharan Africa, as opposed to 8,410 without exclusion.

Subnational Human Development Index

Another frequent correlate with conflict is economic development. To control for economic development, I make use of the newly developed Subnational Human Development Index (SHDI) Version 3.0 (Smits & Permanyer, 2019). The SHDI is based on three dimensions: education, health and standard of living. The variable ranges on a scale from 0 to 1.

Like EM-DAT, the SHDI data are not coded on the grid cell level, but rather on the first order administrative unit. This is more precise than using HDI data on the country level as it allows for variation within countries. Through GIS-tools I assign each cell with the value of its respective administrative unit. For cells containing borders, a weighted average is used. Some administrative units miss data for early years. For instance, in Nigeria there are no data available before 2003. Using similar linear extrapolation techniques used for population is not feasible in this regard, as this provides some units with negative SHDI values — which is not possible. Instead I use the "next observation carried backward"-technique by assigning the same value

[17] There are large discrepancies across population datasets. Some cells may fall below the threshold of < 1 per km^2 when using data from CISEIN and CIAT (2005), whereas they would not fall below the threshold when using other datasets and vice versa (see e.g. data from the History Database of the Global Environment (HYDE) version 3.1. (Klein Goldewijk, Beusen, de Vos & van Drecht, 2011; Klein Goldewijk, Beusen & Janssen, 2010)).

of the first observation to all preceding years. For example, with this technique the Sokoto region in Nigeria holds the value 0.235 for all years between 1989 and 2003, instead of having negative values between 1989 and 1994 as it would with linear extrapolation.[18]

Democracy

It is well documented that both democracies and harsh autocracies experience few civil conflicts (Hegre, Ellingsen, Gates & Gleditsch, 2001). Instead, anocracies, states that are neither a strong democracy nor a strict autocracy, are found to be most prone to experiencing conflicts. This is also found to be typical for communal conflicts (Brosché & Elfversson, 2012).

To measure democracy, I use the *liberal democracy index* from the Varieties of Democracy (V-Dem) dataset v9 (Coppedge et al., 2019). The liberal democracy index combines both liberal and electoral principles of democracy. The liberal principles emphasise core values like the protection of minority rights, constraints on the executive power and equality before the law. The electoral principles consist of the level of free and fair elections and extensive suffrage. The index ranges on an interval scale from 0 to 1, where 0 denotes a closed autocracy and 1 denotes a liberal democracy (Coppedge et al., 2019). This variable is measured on the country level and in cases where cells contain borders I have applied the majority rule — the country with the largest share of the cell is used.

Ethnic Exclusion

A fourth variable often assumed to carry conflict potential is ethnic exclusion (Cederman, Gleditsch & Buhaug, 2013). Communal conflicts, in particular, are more likely in areas with ethnic exclusion, as they are fought between groups organised along a shared communal identity (Brosché & Elfversson, 2012). Some typical communal conflicts have been over agricultural land, such as between the Dinka and the Nuer in South Sudan, or the Hausa (farmers) and Fulani (pastoralists) in Nigeria (Wig & Kromrey, 2018; Abbass, 2012).

[18] For a graphical example of the differences between these two extrapolation techniques see Figure A.1 in appendix.

To account for ethnic exclusion, I include whether there are discriminated or powerless groups in each cell as defined in the Geo-referencing Ethnic Power Relations (GeoEPR) 2014 dataset (Vogt et al., 2015). Benjaminsen and Ba (2009, p. 78) argue that pastoral marginalisation, where land legislation is favouring farmers on the expense of pastoralists, has been a driver for farmer-herder conflicts in West Africa, and in Mali particularly, for many years. However, in Mali, Fulanis (pastoralists) are included in a larger category of "Blacks" in the GeoEPR 2014 dataset (Vogt et al., 2015). Thus, the dataset only denotes whether a group is excluded from central political processes, not whether there are other types of discrimination such as pastoral marginalisation. Similarly, the Dinka and the Nuer are coded as senior partner and junior partners in South Sudan, while Hausa and Fulani are included in the same category and have been both senior and junior partner during the time frame in Nigeria. This has been a source of criticism as "ethnic groups in the data are often aggregated in ways that would be unrecognizable to country experts and group members themselves." (Peterson, 2016, p. 1).

Despite the criticism, I include the data as a control variable to control for ethnic exclusion at the higher level. The variable takes the value 1 if there is one or more ethnic group(s) coded as either powerless or discriminated; and 0 if there is no ethnic exclusion. Cells are missing data if there are no politically relevant groups in their area. To make the sample complete, I assign all these cells with the value 0 as there is no ethnic exclusion in these cells.

Lagged Dependent Variable

One of the best predictors of conflict is whether there was a conflict in the same cell the previous year. To control for this, I create a lagged version of the dependent variable. A caveat with this is that it limits the valid sample to only containing data between 1990 and 2014, as the variable will be missing for the year 1989. However, 25 years with 7,986 cells returns a valid sample of 199,650 observations.

Conflict spillovers are not only temporal, they can also be spatial. Areas with close spatial proximity to conflicts are also more likely to see violent events. Rebels in one area may find shelter in a neighbouring area, or a violent conflict in one area may lead to refugee flows to other areas (Grieco, Ikenberry & Mastanduno, 2019, pp. 224-225). Similarly, when using disaggregated data, conflicts may cross the borders of several cells. However, controlling for this may be problematic as these communal conflicts tend to be small-scaled and droughts may have a diffusion effect — lack of water in some cells may lead to a socio-economic drought in neighbouring cells. Hence, controlling for conflicts in neighbouring cells may not be relevant and may take away some the effect of drought. Similarly, since a drought often engulf larger areas including several cells, small-scaled conflicts unrelated to each other may erupt in neighbouring cells. Therefore, I choose not to include spatial lag of the dependent variable.

Summary Statistics

Table 4.2 shows summary statistics of the variables included in the dataset. The left columns show the original values, whereas the right columns show the recoded values. It is evident that some sort of interpolation of the population variable had to be done, otherwise the regression models would only contain 42,050 observations.

Table 4.2: Summary statistics of variables

Variables	**Version**											
	Original						**Recoded**					
	Min	Mean	Median	Max	Std	N	Min	Mean	Median	Max	Std	N
Conflict event	0	0	0	1	0.07	218,660	0	0	0	1	0.07	218,660
SPEI	-4.97	-0.12	-0.06	3.31	0.91	218,660	0	0.07	0	1	0.26	218,660
NDVI	-0.38	0	0	0.46	0.26	125,569	0	0.16	0	1	0.27	125,569
EM-DAT*	0	0.09	0	1	0.28	218,660	0	0.09	0	1	0.28	218,660
Population**	0	78,130	23,681	7,400,038	199,746	42,050	-2.12	9.79	10.11	15.9	2.06	218,660
SHDI	0.17	0.41	0.39	0.79	0.12	165,330	0.17	0.41	0.39	0.79	0.11	205,322
Democracy	0.01	0.24	0.17	0.77	0.18	217,150	0.01	0.24	0.17	0.77	0.18	217,150
Conflict event$_{t-1}$	0	0	0	1	0.06	210,250	0	0	0	1	0.06	210,250

*The variable “EM-DAT” is not recoded.
**The variable “Population” is logged in the recoded version, but not in the original version.

Chapter 5

Research Design

This chapter presents the research design used to examine the relationship between drought and communal conflict. As discussed in the previous chapter, conventional models have employed country-level data, whereas more recent data and modelling innovations have allowed researchers to examine the climate-conflict nexus at the sub-national level using high resolution data of climate variability and conflict. I start this chapter by discussing the role of causality in social science. Due to the lack of randomisation, we seldom see experiments in social science.[19] As a result of this, researchers often do their best to mimic experiments (Gerring, 2004, p. 350). Mimicking experiments requires well-specified and robust statistical models. In order to obtain such a model, I end this chapter by discussing different estimators and carefully examining certain assumptions related to the specific estimator.

5.1 Causality in Social Science

The aim of all social sciences is to draw inferences by using observed data to obtain knowledge of the unobserved (King et al., 1994, p. 46). According to King, Keohane and Verba (1994) there are two types of inferences: descriptive and causal inferences. A descriptive inference seeks to describe the existence of something. This could for instance be the number of conflicts taking place after a drought. Researchers can use this information to draw information from a sample to a population. However, this does not say anything about causality. A causal inference, on the other hand,

[19] Particularly in political science. Experiments are to some extent more frequent in psychology or economics.

is a statement about *why* something happens. This could then be the causal effect drought has on the probability of experiencing conflict.

The causal effect can be described as the difference between two descriptive inferences. In this case, what happens to a unit after a drought versus what would have happened without a drought (Epstein & King, 2002, p. 34; King, et al., 1994, pp. 81-82). This means that a causal inference is impossible without good descriptive inferences. However, a descriptive inference alone is often not satisfying (King et al., 1994, p. 75).

Estimating the causal effect, however, imposes the 'fundamental problem of causal inference': We can never determine the causality for certain, because we cannot observe the counterfactual (Holland, 1986). For any observation, we can only observe one possible outcome. It is impossible to say whether or not there would have been a conflict between the Afar and the Kereyou in Ethiopia in 2002 without there being a drought. Although clashes have been argued to be sparked by a drought, clashes over agricultural land could have happened regardless of the drought (Sundberg et al., 2012; Pettersson et al., 2019).

The uncertainty imposed by this fundamental problem will always be present in research in social sciences. Even though we can never avoid this uncertainty, we can still do our best to minimise it. Two important components of minimising this are *unit homogeneity* and *conditional independence* (King et al., 1994, p. 91).

Two units are said to be strictly homogeneous if the expected values of the dependent variable are the same for all units when the explanatory variables take on a particular value (King et al., 1994, p. 91). This would be true if we could say that all grid cells experiencing drought would experience conflict. Due to contextual variations this is obviously not the case, and this assumption is seldom upheld in social science.

Therefore, a less strict version of the unit homogeneity assumption is the assumption of *constant effect*. This assumption states that similar variation in values of the explanatory variable leads to the same causal effect in different units (King et al., 1994). This means that if two relatively equal cells experience a drought, we would expect the same effect in both cells. This is an important reason for including control variables and using grid cells, to account for the contextual variation and baseline risk induced by other factors than drought itself. Hence, we can make the units as similar as possible.

The second important component to minimise the uncertainty is the assumption of *conditional independence*. This assumption is upheld when values on the explanatory variable(s) are assigned independently of the values on the dependent variable (King et al., 1994, p. 94). If the explanatory variables are caused, at least partly, by the dependent variable, there exists a problem of endogeneity. As discussed in the previous chapter, this might be a problem with both the NDVI values, as conflict may affect the level of vegetation, and the EM-DAT data, as a conflict may aggravate an ongoing drought-related disaster. One way to account for this is to use instrumental variables regression by using an exogenous proxy for drought. This is what researchers have tried to account for when using precipitation-data. However, as I also discussed in the previous chapter, the precipitation-data may not always be exogenous to conflict. A second way to account for this is by temporally lagging the drought variables with one year. However, since I look at communal conflicts, and these often take place close to the drought occurring, I expect the conflict potential to occur within the same year. Hence, lagging the variables is not feasible either. Thus, there may be a problem of endogeneity in the data and therefore the results from the regression should be carefully interpreted.

5.2 Regression Estimator

When testing causal claims in political science, the main tool used by researchers is multiple regression (Kellstedt & Whitten, 2018, p. 236). In short, the idea of regression is to fit the 'best' possible line through a scattered plot of data points (Kellstedt & Whitten, 2018, p. 189). The fundamental question arising then is how should this line be drawn? In other words, which regression estimator should be used?

When the outcome variable is binary, logistic regression is often preferred (Christophersen, 2018). This is often the case in conflict studies as the dependent variable usually denotes whether or not there was an ongoing conflict in unit i in year t. Although regular linear regression, such as the ordinary least squares estimator

(OLS[20]) is perfectly possible to use when the dependent variable is binary, it faces certain challenges. The most common problem is that it will often provide impossible values, such as probabilities below 0 or above 1. However, the probability can never be below 0 or above 1.

Hence, researchers often tend to use logistic regression as this takes into account the binary nature of the dependent variable.[21] Whereas coefficients produced by OLS can be interpreted as one unit increase in X is associated with a β increase in Y, the coefficients produced by logistic regression tend to be a bit more complex. Coefficients produced by the logit model are often referred to as "logit" or "log-odds" and can be hard to interpret. Instead researchers often interpret the odds ratio instead. The odds ratio can be interpreted as a one unit increase in X is associated with a β change in the odds of Y occurring. I elaborate more on the interpretation of odds and odds ratios in the next chapter.

5.2.1 Multilevel Modelling

A pivotal assumption when using regular regression (be it OLS or regular logistic regression) is that all observations are independent of each other. In this study data are nested in a panel data structure, which means that the same units have repeated observations over a certain time period. As noted in the previous chapter, after excluding cells unlikely to see conflict, there are a total of 7,986 cells measured over 26 years in this study. Using regular OLS or regular logistic regression will assume that each observation is independent of each other. This means that cell number 1000 in year 1991 is expected to be as uncorrelated with the observation of itself in 1992 as it is with cell number 7000 in year 2014. This is most likely not the case since population, SHDI and the level of democracy are not randomly assigned to each cell every year. For instance, the number of inhabitants in each cell will correlate highly from one year to another within that cell. One way to treat this problem is by including unit fixed effects to the regression equation. By doing this, researchers only measure within-variation. This means they are no longer interested in the

[20]OLS is often referred to as the Linear Probability Model (LPM) when the dependent variable is binary.

[21]The logit model is one of two commonly used binary regression models, the other being probit. These models are similar apart from logit using the quantile function of the logistic distribution, while the probit uses the quantile function of the normal distribution.

difference between cells, but only in the variation within each cell. Thus, whereas pooled regression (no fixed effects) requires independence of all observations, fixed effects regression only measures within-variation (Kennedy, 2008, pp. 283-284).

A compromise between these two extremes is multilevel modelling (Gelman & Hill, 2006, p. 251). This is a relatively new and popular way of handling clustered observations and has previously been used to analyse the relationship between climate and conflict (see e.g. von Uexkull et al., 2016; or Döring, 2020). The general aim of using a multilevel logistic model is to estimate the probability of an event occurring, while taking the nested data structure into account. In contrast to a fixed-effects model, multilevel modelling allows the researcher to consider both the within and the between variation.

A typical example of a multilevel model design could be trying to predict different students' grades. Their grades would (most likely) not be random from time to time. To illustrate this, suppose we have a dataset containing a large group of students measured over three years. Each year the students have exams determining their grades. The first, and lowest, level in this multilevel model would be the time-varying predictors. These are variables that vary within each student (across years) but not between students. This could for instance be variables such as time spent studying or hours of sleep the night before the finals. These level-1 factors could influence the test results within each student from one year to another. On the second level we have predictors that are unique for each student, but constant across time. This could for instance be gender or parents' education level. These variables are time-constant, but unit specific, meaning they vary between students but not within students.

Similar to the hypothesised example above, we could imagine grid cells instead of students, and conflicts instead of grades. Grid cells have certain time-varying features such as population, SHDI and drought. These are unique values for each cell and vary from one year to another. On the second level, the cells contain time-invariant features such as the the mean value of these time-varying variables, which country they belong to, or, in some cases, ethnic exclusion. These could be considered level-2 variables — variables that are constant within the time-series. Put simply, on the first level we measure within variation, whereas on the second level we measure between variation. This is the main benefit of the multilevel model.

It is possible to make the model more complex (a three-level model) by arguing that students are nested within classes containing some common features for all students within that class, such as characteristics of the teacher or the size of the class. Similarly, we can argue that grid cells are nested within countries which contain relevant features. However, the only country-level variable included in this analysis is “democracy”. Since the democracy score is time-varying, it will be considered a level-1 variable in the regression analysis. Since there are no other country-level variables, I proceed with a two-level mixed effects model with observations nested on the cell-level — taking into account the differences both within and between cells.

5.2.2 Fitting the Model

The main idea of using a multilevel model is to allow each group (cell) its own intercept (Sommet & Morselli, 2017). In other terms, this allows each group to have its own baseline probability of experiencing conflict. This is commonly referred to as “random intercept”. However, the term “random intercept” has been criticised. Gelman and Hill (2006) refuse to use the term as it is used inconsistently across scholars and the word “random” may be misleading. The intercepts are computed individually for each cell, which means they vary, but they are not random, they are estimated. As a result of this, I use the terminology “varying intercepts” instead of “random intercepts” throughout this thesis.

Recall from the conflict map in the previous chapter (Figure 4.1a) that conflict observations tend to be spatially clustered. While some cells experience numerous conflicts, a large number of cells does not experience conflict at all. Hence it makes sense that the baseline risk of experiencing conflicts should vary between cells.

A statistical way of testing whether the data points are nested in groups is through the calculation of the intraclass correlation coefficient (ICC) (Koch, 2014). The ICC calculates the degree of homogeneity of the outcome within the groups. This is calculated by dividing the between-group variance by a combination of the between-group variance and the within-group variance. This formula is shown in equation 5.1:

$$\text{ICC} = \frac{var(u_{0j})}{var(u_{0j}) + (\pi^2/3)} \tag{5.1}$$

In equation 5.1 $var(u_{0j})$ denotes the between-group variance. This tells whether some cells have a higher conflict-mean than others. Since the regression does not include a within-group variance it uses the standard logistic distribution ($\pi^2/3$) as the assumed level-1 variance component (Sommet & Morselli, 2017).

The ICC ranges from 0 to 1, where 0 indicates perfect independence of residuals — meaning there are no nested effects from the cells, and hence no reason to use a multilevel model. The higher the ICC, the more the observations depend on the cluster relationship (Sommet & Morselli, 2017). For instance, if the ICC = 1 it means that either a cell experiences conflict all the time or never.

$$ICC = \frac{6.96}{6.96 + 3.29} = \frac{6.96}{10.25} = 0.68 \tag{5.2}$$

The ICC for conflict in this study is 0.68, denoting a relatively high degree of intraclass correlation. What this means is that 68% of variation in conflict observations is between cells. In turn, 100% - 68% = 32% of variation in conflict is within cells. The high level of between-variation indicates that observations are in fact clustered and that some cells experience several conflict incidences, whereas other cells do not experience conflict at all. This is also evident from the map of conflict locations in Figure 4.1a. Due to the high level of ICC, a multilevel model with varying intercepts is appropriate in this case (Sommet & Morselli, 2017).

ICC for drought

It is also possible to calculate the ICC for the three drought indices. This is interesting as it tells us whether droughts are spread across units or whether some areas tend to see many droughts and other tend to see relatively few.

Table 5.1: ICC for drought

	SPEI	NDVI	EM-DAT
Between cell variance	6.4%	54.0%	20.5%
Within cell variance	94.6%	46.0%	79.5%

It is evident from Table 5.1 that between and within variances vary across the three indicators. It is no surprise that SPEI contains a large within variance, since all cells are measured against their own long-term mean. Conversely, the extent to which vegetation vary is dependent on the type of vegetation. Since areas with rainforests are not as easily affected by drought as cropland or grassland, we can see that NDVI has a relatively large between-cell variance.[22] This means that the probability of experiencing NDVI drought is not random, but rather dependent on the land type. EM-DAT is somewhere between SPEI and NDVI, with a moderate level of within-cell variance.

Varying slope?

A benefit of the multilevel model is that it allows the researcher not only to impose varying intercepts for the different groups, but also to allow varying slopes between the groups (Gelman & Hill, 2006). This means the effect drought has on conflict could be different between units. Often this will be closer to the true relationship and will therefore make the model a better fit. However, there is an ongoing debate in the literature as to whether researchers should use a maximal model with both varying intercept and varying slope (Barr, Levy, Scheepers & Tilly, 2013), or whether this imposes a risk of overparameterisation (Bates, Mächler, Bolker & Walker, 2014). According to Sommet and Morselli (2017) this should first and foremost be a theoretical question: Do we expect the effect of drought to be inherently different between cells? According Hypothesis 1, I expect drought to have a positive effect on the probability of conflict in all areas. Naturally, the strength of this effect could vary between areas, but I do not expect the effect to be positive in some areas and negative in others. Hence, it does not seem necessary to include varying slopes. Moreover, it is possible to calculate the extent in which the effect of drought varies between clusters. To test this, I run a constrained intermediate model (CIM) and an augmented intermediate model (AIM), and compare both by performing a likelihood-ratio test. The results from the tests are shown in Table 5.2 and the table shows diverging results across the three models. The SPEI model suggests that slopes are in fact not different, whereas in both the NDVI and EM-DAT

[22]The cutoff value for NDVI is thoroughly discussed in part 4.2.2. Choosing a cutoff based on each cell's standard deviation would lead to 100% within variance.

Table 5.2: Varying-slope test

Likelihood-ratio test	SPEI	NDVI	EM-DAT
LR $\chi^2(1) =$	1.85	276.02	22.41
Prob $> \chi^2 =$	0.1733	0.0000	0.0000

model slopes are different between cells. This is not very surprising given that SPEI has a low between-cell variance compared to the other indicators. Moreover, the mathematical results may also be driven by a large number of cells not experiencing drought or conflict.

Since my hypothesis states that a drought should be positively related with communal conflict across all units, and this is supported by at least one of the three models, I choose to run the original model with varying intercepts and fixed slopes.

5.2.3 Regression Assumptions and Diagnostics

In regular OLS or regular logistic regression assumptions and diagnostics are well-defined and heavily discussed (see e.g. Christophersen, 2018; Kennedy, 2008; or Stock & Watson, 2011). There exist numerous tests for controlling for omitted variable bias, multicollinearity or outliers. In multilevel modelling, however, the literature is rather scarce. Some of the same assumptions apply as for regular logistic regression, but the mathematical tests are not straight forward. Based on Christophersen (2018), Kennedy (2008) and Stock and Watson (2011) I identify four assumptions vital to this research design and outline a thorough discussion on these assumptions. These four assumptions are: (1) Linearity and model specification, (2) omitted variable bias, (3) multicollinearity and (4) influential observations.

Linearity and model specification

The logistic model is often referred to as a non-linear model (Stock & Watson, 2011, p. 387). Intuitively the model looks non-linear since the line it produces is S-shaped and not a straight line. However, although the line it produces is not linear, the model is still linear *in parameters*. This means that when one unit increase in X is associated with a β increase in the log-odds of Y, β will have the same value regardless of the value of X. Hence, although the model itself can be described as non-linear, the assumption of linearity (in parameters) is still pivotal.

Since the linearity assumption assumes that the causal effect is the same across the sample, this assumption implies that the variables should be correctly specified. For instance, I have log-transformed the population variable. The reason for log-transforming the variable can be shown with a hypothetical example. Imagine a grid cell (A) containing 200 inhabitants in the year 2000. Due to climatic hardship in surrounding areas, the cell receives 200 new inhabitants the next year. Hence, the population has doubled its size from 200 to 400. Now, imagine a much more populous grid cell (B) with 200.000 inhabitants. Receiving 200 new inhabitants in cell B would most likely not result in the same effects as in cell A. By log-transforming we basically operate with the relative frequency, meaning that for cell B to have the same estimated effect, it also needs to double its population by receiving 200.000 new inhabitants. When a variable is either log-transformed or squared, the regression equation can be described as “nonlinear in variables, but linear in parameters” (Kennedy, 2008, p. 96).

Similar argumentation applies to why I have transformed the ethnic exclusion-variable from originally ranging between 0 and 5, to only take on the binary values of 0 and 1. The theoretical justification is that having five discriminated ethnic groups in an area, is not necessarily making a grid cell five times more likely to experience conflict compared to a cell containing only one discriminated group.

Failing to take the linearity assumption into account and properly specify the variables in the regression model may impose a second problem, namely omitted variable bias.

Omitted variable bias

Omitted variable bias occurs when we fail to control for underlying variables correlated with both the independent and the dependent variable. Since omitted variables are included in the error term, it means that one of the independent variables are correlated with the error term (Christophersen, 2018, p. 73). Failing to control for omitted variable bias results in the estimated parameters being biased (Kennedy, 2008, p. 93).

The best way to prevent omitted variable bias is to include independent variables based on theory. What may affect both the level of drought and the level conflict? It is easy to think that some sort of drought resilience should be included. Democracy, SHDI and ethnic exclusion are all included to capture some of these effects. However, recall the challenging question from the previous chapter: Does a drought deteriorate the general coping capacity in areas and hence make them more likely to see conflict? Or is it rather the case that a drought happening in areas with already low coping capacity carries the conflict potential? As discussed in previous chapter, according to Busby (2018) this has been source of conundrum in the climate-conflict research. My argument for including these control variables is twofold. First, including control variables prevents the chances of inducing omitted variable bias to the model. Although including mediating variables may reduce the efficiency of the estimator, it is generally considered worse to have omitted variable bias than an inefficient model. Second, both NDVI and EM-DAT vary largely between areas, meaning that different areas have a different chance of experiencing drought in the first place, hence including these control variables seems feasible in this thesis.

Kim & Frees (2006) have developed a statistical test to test whether there exists omitted variable bias in a multilevel model. This test is based on running a multilevel model and a robust fixed effects model and comparing these with a Hausman test. If there are significant differences between the models, then the robust model should be preferred as this is less prone to omitted variables than the multilevel model. However, since I use logistic regression containing large numbers of cells not experiencing conflict, running a logistic model with fixed effects will omit all these cells from the regression model. Hence, the test is not possible to run. To cope with this limitation, I run an OLS model with cell-fixed effects as a robustness test in the analysis.

Multicollinearity

Collinearity (X_1 is linear with X_2) and multicollinearity (X_1 is linear with $X_2 + X_3 + \ldots X_k$) occur when there is a perfect or strong relationship between the independent variables. This means that we can perfectly predict one independent variable based on one or more independent variables (King et al., 1994, p. 112). This assumption is often referred to as "perfect multicollinearity" (Stock & Watson, 2011, p. 200).

Recall that with the Thorntwaite method SPEI can be computed solely based on two measurements: precipitation and temperature. By including SPEI, temperature and precipitation as three separate variables in a regression we would end up having a problem of perfect multicollinearity — the SPEI variable is a linear function of temperature and precipitation. However, a much more common problem is a strong tendency of, but not perfect, multicollinearity (Kellstedt & Whitten, 2018 p. 264). This is not as easily detectable as perfect multicollinearity.

The higher the correlation between two or more independent variables, the less information the estimator can use to calculate the parameter estimates, and thus the greater variances and lower efficiency (Kennedy, 2008, p. 194). The likelihood of multicollinearity arises when adding more independent variables to the regression model. Thus, when reducing the problem of omitted variable bias, we may induce a problem of multicollinearity.

Table 5.3: Variance Inflation Factor (VIF)

VIF	SPEI	NDVI	EM-DAT
Drought	1.01	1.02	1.01
Population	1.02	1.03	1.02
SHDI	1.28	1.31	1.29
Ethnic exclusion	1.07	1.07	1.07
Democracy	1.33	1.35	1.33
Conflict (temporal lag)	1.00	1.00	1.00

To test whether the models contain collinearity or multicollinearity I calculate the Variance Inflation Factor (VIF). The VIF is calculated by running OLS regressions several times with each independent variable as the dependent variable (Acock, 2014, p. 288). The R^2 from each regression denotes the proportion of the new dependent variable which is explained by the other variables in the regression model. A VIF above 10 is often assumed to be problematic and according to Christophersen (2018, p. 76) it could inflate the standard errors by 129.4%. When standard errors get inflated we could end up committing a “type II error”: failing to reject a false null hypothesis, and stating that there is not a statistically significant relationship, when there is. Table 5.3 shows that all variables are well below the threshold of 10 and suggests that the models do not suffer from multicollinearity.

Influential observations

The last assumption I will touch upon is that large outliers are unlikely (Stock & Watson, 2011, p. 125). Outliers are one of two types of influential observations.[23] Outliers are observations that the model predicts poorly. However, in a model where less than 0.5% of all observations are conflict observations, all these conflict observations would likely be considered outliers. Moreover, outliers are not generally considered a large problem unless they either are a result of coding mistakes or they are driving the results to a large degree. I have re-run the data processing with conflict data, SPEI data and EM-DAT data and this has resulted in the same number of observations.[24] This makes it less likely that outliers have been constructed due to my own coding mistakes, however there is always a possibility of coding mistakes in the original datasets.

In contrast to country-level studies, a problematic aspect occurring in spatially disaggregated studies is that one conflict may comprise large parts of the dataset. As outlined in the previous chapter, the conflict between Christians and Muslims in Nigeria represents 71 of the 995 conflict observations. Hence it might be that this conflict influences the results from the regression model. In order to test this, I run a robustness test excluding this particular conflict.

Summary

In sum, the models in this analysis does not fulfil all requirements of a regular OLS model. Nor do most regression models. However, it is important to be aware of which flaws the model contains in order to run robustness tests to cope with these problematic aspects.

[23]The second being observations with high "leverage" — observations influencing the model the most.

[24]I have not done this for NDVI data as the data processing took several days.

Chapter 6

Results and Analysis

In this chapter, I present the results from the various regression analyses. I start the chapter with some descriptive statistics illustrating the trends between drought and communal conflict. Second, I present the main regression analysis and discuss the results. Moreover, I run robustness tests to examine whether the results are robust across different models. I also discuss whether SPEI, NDVI and EM-DAT are three different measures of drought or whether they are all measuring the same type of drought but represents different parts of the causal chain. To conclude the chapter, I look into whether or not "drought" has been explicitly mentioned as a contributing factor in the conflicts that coincides with the various drought observations.

6.1 Descriptive Statistics

Descriptive statistics and correlations are interesting because they reveal what we should expect from the regression analysis. If there are large discrepancies between the raw correlations and the regression results, it may be a sign that confounding factors play a large role.

In Figure 6.1 I present the frequency of drought events and communal conflicts in Sub-Saharan Africa between 1989 and 2014. Since I exclude cells with limited population and land area, the total number of valid observations is 207,636.[25] Of all these observations in the dataset, only 995 cell-years experienced conflict. This means that if we were to pick a random observation in the dataset, the probability

[25] Unlike the number presented in Chapter 4 (199,650), this is with all 26 years included instead of only the 25 when including the lagged dependent variable.

that it would contain a conflict is $\frac{995}{207{,}636} = 0.48\%$. This baseline probability is shown with the horizontal dashed line in Figure 6.1.

If we only look at the observations not experiencing any type of drought, the probability of choosing a conflict observation would be lower than the baseline probability with $\frac{736}{170{,}170} = 0.43\%$.

By limiting the sample to only those observations who experienced a SPEI drought, the probability of picking a conflict observation would, surprisingly, be lower than the baseline probability with $\frac{66}{14{,}603} = 0.45\%$. This suggests that SPEI drought is not a good predictor of communal conflicts. However, this is only the raw correlation, not taking into account other relevant variables. For instance, it might be that SPEI droughts in areas with low coping capacity see more conflicts, whereas SPEI droughts have no effect in drought resilient areas.

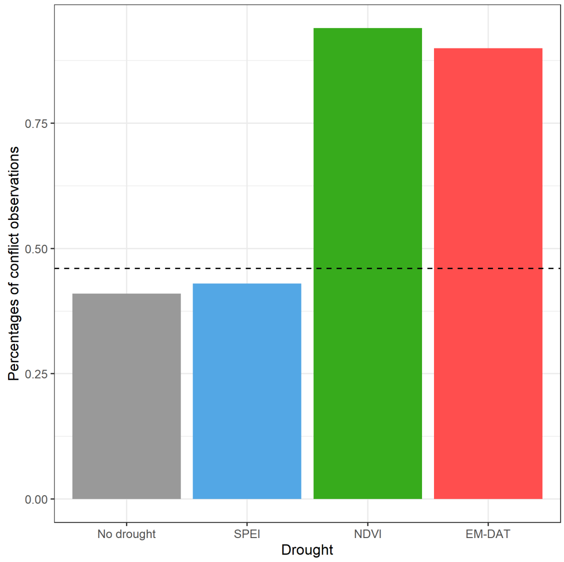


Figure 6.1: Conflict observations

While the probability of picking a conflict observation from the SPEI sample is a bit lower than the baseline probability in the whole dataset, this changes drastically for NDVI and EM-DAT droughts. For both NDVI and EM-DAT droughts the probabilities of experiencing communal conflict are over twice the size of SPEI droughts, with $\frac{98}{9{,}933} = 0.99\%$ and $\frac{169}{17{,}840} = 0.95\%$, respectively. Again, these probabilities only reflect the raw correlations, not considering confounding factors. Therefore, researchers need to make use of regression models to control for these underlying factors.

6.2 Regression Models

In Table 6.1 I present the results from the multilevel logistic regressions. Table 6.1 contains nine models in total: three models for each drought indicator, containing

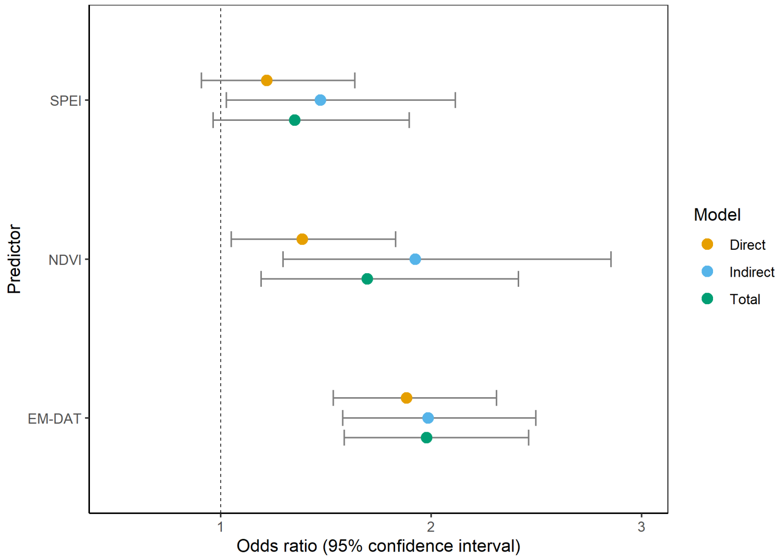


Figure 6.2: Odds ratios

the direct, indirect and the total effect. I provide the coefficients as odds ratios as these are easier to interpret than log-odds. The corresponding standard errors are provided in parentheses.

For a more intuitive interpretation of the relationship between drought and conflict I plot the odds ratios for the different drought indicators in Figure 6.2. In this figure I exclude odds ratios for the control variables as the main scientific interest is to estimate the effect of the different drought indicators. In Figure 6.2 the orange plot represents the direct model, the blue represents the indirect model and the green represents the total model. The whiskers represent the corresponding 95% confidence interval and the dashed vertical line denotes where the odds ratio is 1 (there is no effect).

The odds ratio for SPEI is positive in all three models, suggesting SPEI droughts are associated with a higher likelihood of experiencing communal conflict. This supports Hypothesis 1 that droughts in general are associated with more conflicts. However, the coefficient is only (and barely) significant in the indirect model. Additionally, in very large-N studies even small coefficients may be significant. Hence, significance does not necessarily reflect a true relationship in these very large-N studies. Failing to acknowledge this is referred to as the "large sample size fallacy" (Lantz, 2013). Thus, researchers should not only look at the significance, but also consider the size of the coefficient in order to make sound conclusions. As seen from

Table 6.1: Regression models

	Models								
	(1-3) Direct			**(4-6) Indirect**			**(7-9) Total**		
	(1) SPEI	(2) NDVI	(3) EM-DAT	(4) SPEI	(5) NDVI	(6) EM-DAT	(7) SPEI	(8) NDVI	(9) EM-DAT
SPEI	1.22 (0.18)			1.47* (0.27)			1.35 (0.23)		
NDVI		1.39* (0.20)			1.92** (0.39)			1.70** (0.31)	
EM-DAT			1.88*** (0.20)			1.99*** (0.23)			1.98*** (0.22)
Population$_{log}$	1.87*** (0.09)	1.65*** (0.08)	1.84*** (0.09)	1.87*** (0.09)	1.65*** (0.08)	1.84*** (0.09)	1.87*** (0.09)	1.65*** (0.08)	1.84*** (0.09)
SHDI	2.63 (1.64)	1.06 (0.94)	2.74 (1.70)	2.72 (1.70)	1.15 (0.87)	2.78 (1.72)	2.69 (1.68)	1.11 (0.84)	2.77 (1.72)
Democracy	0.17*** (0.07)	0.05*** (0.02)	0.15*** (0.06)	0.17*** (0.07)	0.05*** (0.02)	0.15*** (0.06)	0.17*** (0.07)	0.05*** (0.02)	0.15*** (0.06)
Ethnic exclusion	1.19 (0.14)	1.29 (0.17)	1.17 (0.13)	1.19 (0.14)	1.28 (0.17)	1.17 (0.14)	1.19 (0.14)	1.29 (0.17)	1.17 (0.13)
Conflict event$_{t-1}$	5.70*** (0.66)	4.81*** (0.22)	5.90*** (0.69)	5.70*** (0.66)	4.90*** (0.66)	5.92*** (0.69)	5.70*** (0.66)	4.85*** (0.66)	5.91*** (0.69)
Intercept	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Observations	185,865	112,099	186,190	186,140	112,380	186,140	185,840	112,084	186,140
No of groups	7,482	7,487	7,494	7,492	7,492	7,492	7,481	7,486	7,492
LL	-4,114	-3,113	-4,099	-4,114	-3,112	-4,100	-4,114	-3,112	-4,099
AIC	8,246	6,244	8,213	8,244	6,240	8,216	8,245	6,241	8,214

*: p <0.05, **: p <0.01, ***: p <0.001. All models are estimated through multilevel mixed-effects logistic regression with observations nested on cells. Coefficients are shown as odds ratios with standard errors in parentheses. Cells with less than 200 inhabitants or 100 km^2 are omitted.

Figure 6.2, the lower bound of the 95% confidence interval of both the indirect and the total model are pretty close to each other, but where the total model merely crosses the dashed line, the indirect model does not. Consequently, it may be wrong to state that there is a true relationship between indirect drought and conflict and at the same time write off the other two drought models. Hence, there seems to be a weak and inconsistent, albeit positive, relationship between SPEI droughts and the risk of communal conflict.

NDVI, on the other hand, is positive and significant in all three models, with odds ratios varying between 1.39 and 1.92. This is also in accordance with Hypothesis 1: that there is a positive relationship between droughts in general and communal conflict. Moreover, since the effect seems to be stronger than SPEI, this also supports Hypothesis 2: the higher physical impact of the drought, the higher conflict risk. However, since the NDVI data only are available after 2000, the sample size is considerably smaller in this regression model.[26] In turn this results in larger uncertainty. This is particularly prominent in both the indirect and the total model. Moreover, in both the SPEI and NDVI models the indirect effects seem to be strongest. This means that the highest likelihood of conflict, at least from what the regression models are able to estimate, occurs when there are droughts in the neighbouring cells. However, it is important to keep in mind that a one unit increase in the indirect effect equals all neighbouring cells experiencing drought. In many cases this also means that the cell in question experiences drought. Thus, on the one hand this might be due to the coding of the variable. On the other hand, this may be a sign that less precipitation in one area may lead to droughts and conflicts in neighbouring areas. If a cell contains important water sources, such as headwaters for rivers, lack of precipitation in one cell may lead to droughts in neighbouring cells. Similarly, for NDVI, deterioration of vegetation in one area may lead to migration to neighbouring cells, which in turn may spark communal conflicts.

Similarly to NDVI, EM-DAT is also positive and significant in all models. As seen from Figure 6.2, EM-DAT's coefficients are mostly the same in all three models. Since EM-DAT is measured on the first order administrative level, it means that cells and their neighbours often have the same value on the EM-DAT variable. Hence, it

[26]For NDVI N $\approx 112,000$, whereas for SPEI and EM-DAT N $\approx 186,000$.

makes sense that the direct, indirect and total effect reflect the same relationship. In all models the odds ratios are approximately 2 and the coefficients are highly significant. Although this may be a result of the large sample size, the relationship seems to be robust to several robustness checks, as I will elaborate more on in part 6.2.2. The results from the EM-DAT model also support both hypotheses.

When the odds ratio is 2, as in the EM-DAT models, it means that the odds of experiencing a communal conflict incidence double when there is an EM-DAT drought. This should not be interpreted as a doubling of the *probability*. The odds are defined as the probability of an event occurring divided by the probability that it does not occur, and can be written as $\frac{p}{1-p}$, where p is the probability. Thus, the odds may range from 0 to ∞. If the probability is 0.5, the odds are $\frac{0.5}{1-0.5} = 1$. If the odds double from 1 to 2, the probability does not necessarily double. In fact, if the probability is 0.5 (or $\frac{1}{2}$) and the odds double, the new probability is 0.67 (or $\frac{2}{3}$).[27] Similarly for each n-doubling of the odds, when the baseline probability is 0.5, the new probability can be written as $\frac{n}{n+1}$. This illustrates the asymptotic features of the logistic regression model, as the odds may range from 0 to ∞, but the estimated probability will never reach exactly 0 or 1.

In this thesis, however, the probability of experiencing conflict is far from 0.5. In fact, the baseline probability of experiencing a conflict (not taking into account control variables) is 0.0043 (or 0.43%). This means the odds are $\frac{0.0043}{0.9957} = 0.0043$. A doubling of the odds in this case gives an new odds of 0.0086 and hence the probability can be calculated with basic algebra:

$$\begin{aligned}
\frac{x}{1-x} &= 0.0086 \\
x &= 0.0086 - 0.0086x \\
1.0086x &= 0.0086 \\
\frac{1.0086x}{1.0086} &= \frac{0.0086}{1.0086} \\
x &\approx 0.0085
\end{aligned} \tag{6.1}$$

As equation 6.1 shows, if the baseline probability of experiencing conflict is 0.43%

[27] As $\frac{x}{1-x} = 2$ results in $x = \frac{2}{3}$.

and the odds ratio is 2, the new probability of experiencing conflict is 0.85% holding all other variables constant. This illustrates three points. First, when the probability is close to zero, the odds converge towards the probability since $\frac{0}{1} = 0$. Hence, in this case a doubling of the odds actually resembles a doubling of the probability. Second, since the baseline probability (intercept) varies from cell to cell and the values of different control variables vary, changes in the probability are not constant across areas, but need to be calculated for each cell individually. Since 0.43% is the average baseline probability, some areas may have a much higher or much lower baseline probability of experiencing conflict. Thus, while the odds double for all cells, the probability does not double for all cells(!) Lastly, although a doubling of the probability may sound overwhelming, 0.85% is still a relatively low probability of experiencing conflict.

To sum up, the results from the main regression model show that SPEI only weakly and inconsistently supports Hypothesis 1. Since both NDVI and EM-DAT show support for this hypothesis and the effect of EM-DAT is stronger than that of NDVI, the regression model also supports Hypothesis 2. This is interesting as it points to the fact that the higher physical impact of the drought, or the closer we get to measuring the social impacts of the drought, the higher risk of seeing communal conflict. This is scientifically interesting as this shows that using NDVI and EM-DAT to measure drought arguably reveals a larger picture of the relationship between drought and conflict than portrayed by former research. I am not to say that NDVI or EM-DAT measures drought *better* than SPEI, but they arguably capture some aspects that SPEI is not able to capture. Moreover, this also has politically interesting implications as this points to the fact that if we can prevent meteorological droughts becoming socio-economic droughts, we can possibly reduce the risk of experiencing communal conflict. I discuss these implications more in detail in the next chapter.

6.2.1 Control Variables

It is a good sign that the effects of the control variables are consistent across all nine models. The size of the population is significant and positively associated with communal conflict in all models. Suggesting a one unit increase in the log of the population is associated with 84-87% increase in the odds of seeing communal conflict in the SPEI and EM-DAT model. This effect is marginally weaker in the NDVI model, which only contains data between 2000 and 2014, suggesting that the effect of population on communal conflict has decreased in recent years. This is also true when running regressions with SPEI and EM-DAT only including data between 2000-2014, as I show in the next section. Moreover, this is in line with previous findings that higher population increase the likelihood of communal conflict.

Surprisingly, the coefficient of SHDI is positive, suggesting that increased development is associated with a higher risk of conflict. This effect is stronger in the models including the whole time-series, but the effect is not significant in any of the models and should therefore not be given any explanatory power. Nonetheless, this is interesting as it is contrary to former findings. Based on former findings we would expect lower SHDI levels to be associated with conflicts. A possible explanation for this relationship is that SHDI is not measured on the grid cell level, but on the first order administrative level. Thus, if there are large variations in living conditions between cells within these administrative regions, this would not be captured by this variable.

The level of democracy, on the other hand, seems to be a strong predictor of communal conflict. For all models the odds ratio ranges between 0.05 and 0.17. This suggests a decrease in the odds of experiencing conflict with 83%-95% with one unit increase on the democracy score. However, this effect may appear stronger than it is. Since the democracy variable ranges between 0 and 1, where the lowest observed value is 0.012 and the highest is 0.770, it is impossible to have a one unit increase in democracy. Nevertheless, the effect is strong and significant. Previous literature suggests that the relationship between democracy and conflict can be portrayed as an inverted-U shape, with countries neither being strong democracies nor strong autocracies experience more conflicts (Brosché & Elfversson, 2012; Hegre et al., 2001). Since I am not primarily interested in the effect of democracy, but rather include the variable to control for the level of democracy, I have not squared and centred the term to account for this.

Ethnic exclusion is barely positive, but not significant, and hence it does not seem to affect the likelihood of communal conflict to any notable extent. This is also contrary to previous research (see Brosché & Elfversson, 2012; Theisen et al., 2011). As discussed in Chapter 4, this may be a result of the way the variable is operationalised. Communal conflicts are often tied to pastoral and rural marginalisation (Benjaminsen et al., 2009). This is not captured by this variable as it only accounts for political exclusion on the state level.

Lastly, experiencing a conflict the previous year has a very strong effect on the likelihood of experiencing conflict the given year. This is not surprising as the dependent variable is conflict incidence and not conflict onset. Across all models the odds of experiencing a conflict in year t increase between four and six times if the cell experienced a conflict in year *t-1*.

6.2.2 Robustness Tests

To test whether the results are prone to the choice of estimator or data availability, I run several robustness tests.

Since the NDVI data only are available from 2000, using SPEI and EM-DAT data from 1989 may not be feasible when comparing the effects to NDVI. To examine this, I re-run the regressions only using data between 2000 and 2014. The coefficients are shown in Figure 6.3a.[28] When using data between 2000 and 2014, both SPEI's and EM-DAT's coefficients increase. Now SPEI's coefficients are significant in all models. Similarly, EM-DAT's coefficients suggest an increase in the odds by 135%. Interestingly, this is contrary to former findings, which find that the relationship between drought and conflict has diminished in recent years (Ciccone, 2011). However, these results are most likely driven by the increase in the number of communal conflicts in later years (recall the conflict trend graph from Figure 4.1b).

As a second robustness test, I use other operationalisations of the drought variables. Since the cutoff imposed on SPEI and NDVI may either be too strict or too weak, I run models where SPEI and NDVI are measured as continuous deviations from the cell mean. I have reversed the scale, meaning negative values suggest more precipitation and higher vegetation, whereas positive values are associated with droughts. This is done in order to get the same interpretation that positive coefficients suggest that droughts are associated with more conflict incidences. Both

[28]Regression table can be found in Table A.2 in appendix.

the SPEI and NDVI models are positive. SPEI is barely significant on the 0.05-level, whereas NDVI is not. Additionally, I construct a cumulative model representing the number of years in a row with drought. Coefficients are positive for all three indicators, but only the EM-DAT coefficient is significant. Lastly, I run a model where drought is lagged one year. Similar to the cumulative model, all coefficients are positive, but only EM-DAT obtain significance on the 95%-level. I plot the coefficients in Figure 6.3b.[29]

As discussed in Chapter 4, the geo-precision in the conflict data varies. To account for this, I run a regression model excluding all conflict observations not obtaining the highest level of geo-precision. This means that we know that all conflicts happened in the cell they are coded in. In this model SPEI is not significant, whereas NDVI and EM-DAT are positive and significant. Moreover, since the conflict between Christians and Muslims in Nigeria make up 71 of the 995 conflict observations, I also run a regression model excluding this conflict. Similarly, in this model SPEI is not significant, but NDVI and EM-DAT are positive and significant. Additionally, to account for the MAUP, where the size of the cells may affect the regression results, I run a model examining the direct effect of the drought measures with cells being 2° × 2°. In this model only EM-DAT is significant. On the one hand, this may support the ar-

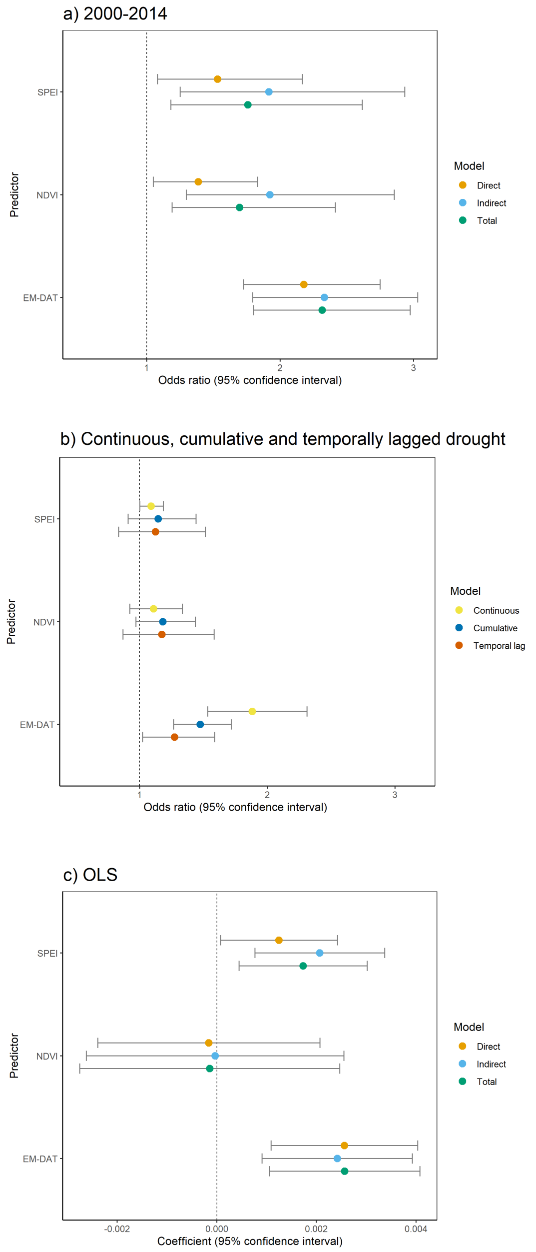


Figure 6.3: Robustness checks

[29]Regression table can be found in Table A.3 in appendix.

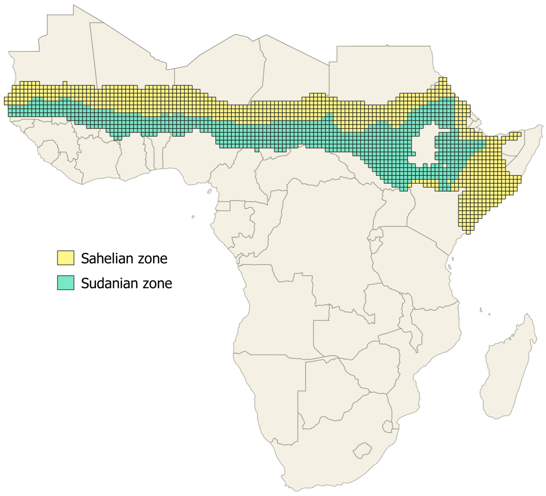
gument that there might be a diffusion effect from other cells when using SPEI and NDVI. On the other hand, this may also reflect the "large sample size fallacy" as the number of observations is drastically reduced in this model.[30]

So far, all regressions have been estimated using multilevel logistic models. Although there are strong theoretical justifications for using a multilevel model, the estimator itself may be prone to various specifications. Hence, if the results are robust, other estimators should yield similar results. To test this, I run an OLS model with cell-fixed effects. As discussed in the previous chapter, a caveat with the fixed effects OLS is that it only utilises within-variation and not between-variation. However, in this case, this is an important addition as it shows in Figure 6.3c that the NDVI effect is close to zero in the within-models.[31] This suggest the effect of NDVI on communal conflict is a result of the between-variation. In other words, the effect is a result of NDVI droughts taking place in the same areas as communal conflicts.

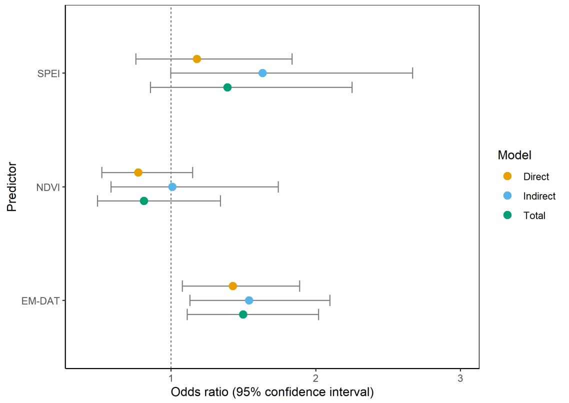
This has several implications for the theoretical discussion. On the one hand, it might be the case that some areas are more prone to experience NDVI droughts and these same areas see a higher frequency of communal conflicts. This effect would not be captured by SPEI as a SPEI drought may occur everywhere, whilst it would to a large extent be captured by NDVI as this is dependent on the land type. However, this may be problematic when running a statistical analysis of the whole Sub-Saharan Africa. Thus, it might be that the NDVI results actually show that conflicts between farmers and herders tend to take place in areas where vegetation varies. Areas with variation in NDVI correlate strongly with areas containing cropland, grassland and shrubland, which is in fact where farmers and herders live. This is especially problematic when large parts of the sample (Sub-Saharan Africa) never experience NDVI droughts, nor communal conflicts. To further investigate this, I identify the areas where both NDVI droughts and communal conflicts tend to take place, namely in the Sudano-Sahelian Zone.

[30]I have not made graphs from these regression models, but regression tables can be found in Table A.4, Table A.5 and Table A.6 in appendix.

[31]Regression table can be found in Table A.7 in appendix.



(a) Map of the Sudano-Sahelian Zone



(b) Odds ratios for the Sudano-Sahelian Zone

Figure 6.4: The Sudano-Sahelian Zone

6.3 The Sudano-Sahelian Zone

The Sahel region is located on the fringes of the Saharan desert (Benjaminsen, 2008). In recent years the area has received a great amount of scientific focus, both due to the high number of conflicts occurring and due to the assumed climatic vulnerability in the area. The Sahel region is characterised as one of the poorest, least developed regions of the world, and approximately 80%-90% of the workforce is active in the agricultural sector (Vågen & Gumbricht, 2012).

There is no exact definition of what constitutes the Sahel. The region is typically defined as the area receiving between 200mm and 600mm mean annual rainfall. This is marked with yellow in the map in Figure 6.4a. However, since both communal conflicts and NDVI droughts also tend to take place in what is often referred to as the Sudanian zone, the area receiving 600mm-1000mm mean annual rainfall, I extend the notion of 'Sahel' to include both the isohyet[32] containing the Sahelian zone and the Sudanian zone.[33] The Sudanian zone is with green cells in Figure 6.4a.

The new regression results are shown in Figure 6.4b. SPEI behaves more or less similar to the main model, and EM-DAT is positive and significant, but the coefficients are halved. NDVI, however, sees a drastic change. In both the direct and the total model the coefficients of NDVI are negative, suggesting that more

[32]An isohyet is a line on a map connecting points with the same amount of rainfall in a given period.

[33]This has previously done by Karlson and Ostwald (2016) and Bargués Tobella (2016).

vegetation is associated with a higher likelihood of communal conflict. However, none of the coefficients are significant.

These new results can be interpreted as support of previous findings that more vegetation is associated with more communal conflicts, particularly cattle raiding in Eastern Africa (Meier et al., 2007; Theisen, 2012; Witsenburg & Adano, 2009). By further distinguishing between East and West Africa within the Sudano-Sahelian zone, I find that NDVI has a positive effect on conflict in West Africa, but a negative effect in East Africa. This could support these previous findings. However, none of these models are significant as the number of observations is rather scarce.[34]

6.4 Sequential Effects

As discussed previously, SPEI, NDVI and EM-DAT do not necessarily reflect three different theoretical concepts of drought, but may rather represent three different components in the causal chain of a drought. Lack of rainfall (SPEI) leads to less vegetation, loss of crops and pasture (NDVI), which in turn may lead to a socio-economic disaster (EM-DAT). To test this possible causal mechanism, I regress SPEI on NDVI and NDVI on EM-DAT. The results are shown in Figure 6.5.

Unsurprisingly, all coefficients are positive and highly significant. This means if there is lack of precipitation, there is a higher chance of reduction in the quality of vegetation. Similarly, if there is deterioration of vegetation, there is a higher chance of having a drought emergency.

An interesting feature in Figure 6.5 is that the indirect effect seems to have the strongest coefficient in both models. This may suggest that there is a spatial diffusion effect between the different measures. Lack of precipitation may not necessarily affect vegetation in the same cell, but it could lead to a reduction in stream flow water in rivers or water reservoirs. This, in turn, could affect vegetation in neighbouring cells. Similarly, vegetation deterioration in large areas seems to be a better predictor of EM-DAT crisis. This may also be a result of the way “indirect” is measured where the variable only equals 1 if all neighbouring cells experience drought.

This short analysis does not conclude whether SPEI, NDVI and EM-DAT should

[34] 13,866 observations for East Africa and 8,201 for West Africa.

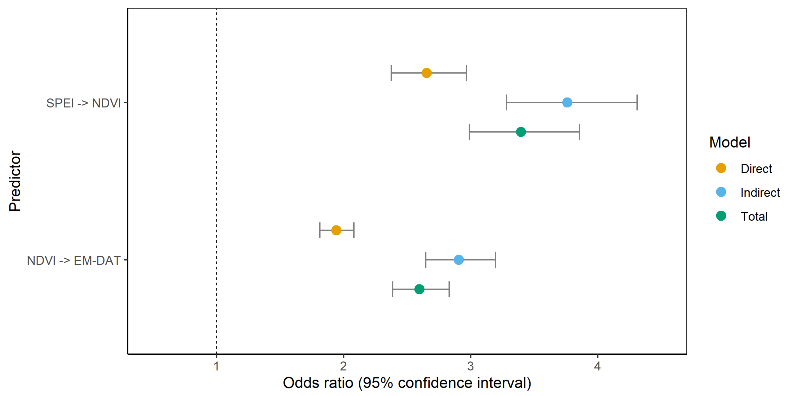


Figure 6.5: Sequential effects of SPEI, NDVI and EM-DAT

be considered three different measures or three different components of the same measure. This could be the basis for discussion in a completely different thesis. Nonetheless, what these results show is that whether or not we consider these measures as similar or different droughts, EM-DAT is arguably a closer to measuring the deterioration of living conditions (be it cropland, grazing land or water sources) than SPEI.

6.5 Explicit Mention of "Drought"

Before summarising the thesis and discussing future implications, I examine whether drought has been argued to be a possible contributing factor in the different conflicts.

First, I identify all conflict observations coinciding with a drought in the same cell (the direct effect). This amounts to 32 SPEI conflicts, 46 NDVI conflicts and 64 EM-DAT conflicts. Moreover, I examine whether "drought" has been explicitly mentioned as a plausible contributing factor to the conflict according to the UCDP Non-State Conflict Dataset (Sundberg et al., 2012; Pettersson, et al., 2019). According to the dataset $\frac{5}{32}$ of SPEI conflicts contained drought as a possible explanatory factor, $\frac{21}{46}$ of NDVI conflicts contained drought and $\frac{19}{64}$ EM-DAT of conflicts contained drought. This is shown as a pie chart in Figure 6.6 where the darker areas represents explicit mentions of drought.

There are two main findings from this quick investigation. On the one hand, this suggests that both NDVI and EM-DAT are better than SPEI at identifying conflicts where drought has played a contributing factor. On the other hand, since

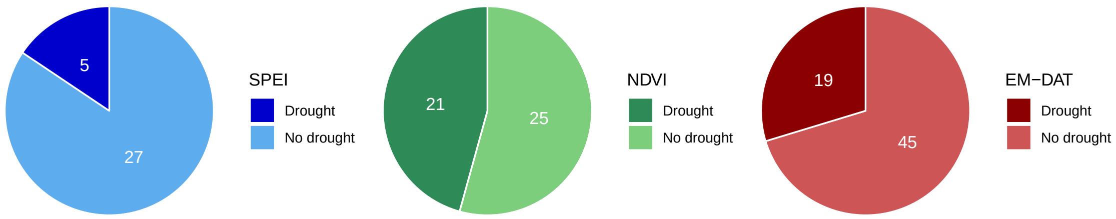


Figure 6.6: Drought explicitly mentioned in UCDP

NDVI has the largest share of "drought-related conflicts", it might denote that these conflicts tend to take part in areas with more cropland and grassland, such as the Sudano-Sahelian Zone.

To sum up the analysis in general, results from the main regression model are both in line with Hypothesis 1: *Droughts are associated with a higher likelihood of communal conflict*, and Hypothesis 2: *The higher physical impact of the drought, the higher the likelihood of communal conflict.*. As shown through robustness tests, this relationship does not seem to be dependent on minor changes in the data. However, when only considering the within-variation through the fixed effects OLS model, it becomes apparent that the correlation between NDVI and communal conflict is based on spatial correlation rather than temporal correlation. NDVI droughts and communal conflicts tend to take place in the same areas, but not necessarily within the same years.

Chapter 7

Conclusion

7.1 Summary

The aim of this thesis has been to answer the research question: *Do different conceptualisations of drought affect the likelihood of communal conflict?* I have argued that a problematic aspect with former literature has been the lack of discussion about what constitutes a drought. The operational validity of SPEI as a drought measure has mostly been taken for granted, although it contains several limitations. My argument has been that the largest caveat with using SPEI as a drought measure is that it primarily captures *meteorological* drought. Moreover, the low number of in situ climate stations used to measure precipitation and temperature makes the measurement less reliable, and recent research suggests that the estimates may not be as exogenous to conflict as previously assumed. Hence, to get a more nuanced answer to the research question on how drought affects communal conflict I have tested three different drought measures: SPEI, NDVI and EM-DAT.

In line with Wilhite and Glantz' (1985) distinctions between different drought definitions, I have argued that SPEI can be perceived as a measure of meteorological drought, NDVI as a measure of agricultural drought and EM-DAT as a measure of socio-economic drought. Combining the implications of these different drought definitions with previous literature on how drought and scarcity may deprive livelihood (see e.g. Homer-Dixon, 1999) I have formulated two hypotheses.

Hypothesis 1 (H1): *Droughts are associated with a higher likelihood of communal conflict.*

Hypothesis 2 (H2): *The higher physical impact of the drought, the higher the likelihood of communal conflict.*

My main model shows conditional support for both hypotheses. SPEI is positive in both the direct, indirect and total model, but only significant in the indirect model. On the one hand, this may suggest that using $0.5° \times 0.5°$ cells are too spatially disaggregated to capture to true effect. However, the direct effect is neither significant when using cells with $2° \times 2°$. On the other hand, this may indicate that there is a diffusion effect, where the lack of precipitation in one area may lead to a drought in surrounding areas. This is also supported by the analysis examining the possibility of sequential effects, where I regress SPEI on NDVI and NDVI on EM-DAT.

Contrary to SPEI, NDVI is positive and significant in all three models in the main regression analysis. However, as I have discussed extensively in this thesis, the estimated relationship between NDVI drought and communal conflict seems to be a result of spatial correlation rather than temporal correlation. Put differently, NDVI droughts and communal conflicts tend to take place in the same areas, but not necessarily within the same years. This is evident both from the OLS regression with fixed effects, and when limiting the sample to only those areas where both NDVI droughts and communal conflicts tend to take place — the Sudano-Sahelian Zone. In both these robustness tests the significance of NDVI disappears.

EM-DAT obtains by far the most robust relationship with conflict in this study. EM-DAT's coefficients are positive and significant in all models and robustness tests, suggesting there is a robust relationship between EM-DAT droughts and communal conflict incidences. Nevertheless, the high possibility of endogeneity in this study makes it not suitable to draw a causal inference that there is a causal relationship between EM-DAT and conflict as the estimated relationship may in part be caused by reverse causality.

7.2 Strengths and Limitations

The main strength of this thesis is the questioning and investigation of a part largely neglected in former studies on the climate-conflict nexus, namely how to measure drought and how it is the implications of the drought we are interested in. The application of widely different conceptualisations and operationalisations of drought widens the scope regarding how researchers on the climate-conflict nexus should understand the concept of drought.

Moreover, the use of spatially disaggregated data by using grid-cells as units offers a much more precise understanding of the local relationship between drought and conflict, than using country or sub-national units. Similarly, benefiting from a multilevel logistic regression model reveals a more accurate picture as it utilises more information than a regular pooled or fixed-effects model.

Another strength with this thesis is the thorough discussion of the limitations and particularly the investigation of NDVI droughts in the Sudano-Sahelian Zone. This provides nuances to the relationship estimated by the main multilevel regression model, as it takes into consideration that vegetation levels may be dependent on type of vegetation in each cell.

Nevertheless, there are some obvious limitations with this research design. The main caveat in this thesis is the problem of endogenous regressors. This is also the reason why researchers have relied on precipitation-based measures in the first place, as SPEI and other precipitation measures are assumed to be exogenous to conflict: Precipitation may affect conflict, but conflict does not affect precipitation.[35] NDVI and EM-DAT, however, may in part be caused by conflict and hence they may be endogenous: Conflict may lead to loss of vegetation or increase the number of people affected by an ongoing drought. One way to account for this is to use instrumental variables regression by using proxies for the various drought variables. However, this is exactly what the use of SPEI is assumed to capture as rainfall is assumed to be an exogenous proxy for drought. A second possibility is to use lagged variables, which I will touch upon in the next section.

A second short-coming relates to the spatial resolution of the EM-DAT data. Contrary to SPEI and NDVI, EM-DAT is not measured on the grid cell level, but rather on the first order administrative level. This may be problematic when considering the local effects as the variable is not distributed independently between cells within the same unit. One way to account for this is using sub-national levels as units in the analysis rather than grid cells.

[35] Unless the use of very advanced chemical and biological weapons, which (obviously) is not the case in communal conflicts.

The third short-coming relates to the interpretation of NDVI deviations. Since these deviations are dependent on the land type, some areas are able to have much larger deviations than other areas. On the one hand, this seems logical. It might be the case that areas experiencing a loss of vegetation should be dependent on, and not relative to, the type of vegetation. However, on the other hand, this leads to only some areas being able to experience drought, which may be problematic.

Finally, the representativeness of the time-series in this study is an additional limitation. Although I include 26 years, the findings in this study are neither representative for years before 1989, nor for years after 2014. This is not a specific caveat limited to this thesis, but rather a short-coming that applies to all studies in social science. However, this is particularly important to acknowledge on the climate-conflict nexus as we often are interested in the future, with the coming climate changes. On a similar note, examining climate variability and conflict on the yearly level may not be ideal. Both droughts and conflict may be short-lived. Hence, measuring these variables on the monthly level may provide a more accurate picture than using years. However, using monthly data instead of yearly data has not been possible within the scope of this thesis.

7.3 Implications and Further Research

The main interest of this thesis has been to investigate whether different operationalisations of drought yield different estimates on the conflict risk. My aim is not to say that any of these measures are *better* measures of drought than SPEI. My argument is that these other concepts arguably capture some of the aspects we are interested in, which SPEI does not capture. As discussed above there are certain problems with all these three measures. However, my aim and hope is to shed light on this knowledge gap and provide valuable information on how we best can operationalise the theoretical concept of drought we are interested in, in future studies.

The main finding in this thesis is that the drought measure with highest social impact is the best predictor of communal conflicts. The question then arising is how can we prevent meteorological droughts becoming socio-economic droughts? Since my finding is that socio-economic droughts are more associated with conflict than meteorological droughts, if we can prevent meteorological droughts becoming socio-economic droughts in the future, then we can possibly reduce the risk of communal conflict. However, the simple statement that if droughts are associated with more conflicts, and we will see more droughts in the future, we will see more conflicts, is not necessarily true. Vast improvements in living standards, increased drought resilience and local adaptations to weather patterns make people more resilient to the physical impacts of meteorological drought and may reduce the chances of lack of precipitation becoming a socio-economic disaster. Paradoxically, the major driver of climate change, economic development, is also one of the major drivers of reduced conflict risk and climate vulnerability (Gartzke, 2012).

Even though this thesis finds a robust relationship between EM-DAT and communal conflicts, this should not be interpreted as evidence of climate playing a large role in conflicts. In many situations politics are still to blame. Blaming conflicts on climate may be an easy way out instead of holding politicians accountable. For example, in 1994 the Washington Post argued that a key underlying cause of the Rwandan genocide was a struggle over land (Homer-Dixon, 1999, p. 17). However, as Homer-Dixon (1999, p. 17) argues, this is taking the focus away from the main cause — a struggle among ethnic groups for control over the Rwandan state. Similarly, some researchers have argued that the ongoing civil war in Syria starting in 2011 was partly sparked by a drought (Gleick, 2014). Yet researchers should be careful in disclaiming the liabilities of politicians.

These examples also point to a caveat with quantitative studies. A statistically significant relationship does not imply causation. Most likely would a statistical analysis including Syria show that a drought occurred right before the conflict erupted. Hence, we are still in need of proper qualitative in-depth studies in order to best reveal the causal mechanisms at play.

In further research it would be interesting to study the relationship between these three drought measures and other types of conflict. Since I focus on communal conflict, I expect the conflict potential of drought to occur within the same year. This imposes the challenge of endogeneity. Hence, one way to reduce this problem is to use lagged versions of the drought variables. Examining the relationship between drought and popular riots by using a lagged drought variable, would be interesting and reduce some of these endogeneity problems.

Another point of departure for further research is to use the NDVI and EM-DAT data to calibrate the SPEI variable. One way to do this is by testing different cutoffs. Which SPEI cutoff provides the best predictions of socio-economic drought? SPEI values $< -1, < -1.5$ or < -2? Moreover, this could also be used to calibrate the temporal way of using SPEI in future research by testing which SPEI version best predicts a socio-economic drought: SPEI-1, SPEI-3, SPEI-6 or SPEI-12? Or to what extent should the SPEI variable be lagged in order to best capture the socio-economic drought? By one year? Or rather by a few months? Additionally, do these decisions vary depending on different land types? These are all highly interesting questions this thesis has shed light on and made possible to further investigate.

Finally, although this thesis, along with former studies, has unveiled that there seems to be a relationship between drought and conflict, there is now mounting pressure to acquire knowledge on how climate change *mitigation* may affect conflict. Contrary to drought, which is only assumed to carry conflict potential in rural and less developed areas, we see evidence that climate change mitigation may spark tensions all across the globe. In France, the Yellow vests movement has been rioting since 2018, and leading up to the 2019 Norwegian local elections there was a massive uproar against road tolls. In Asia, Reducing Emissions from Deforestation and Forest Degradation (REDD+) has sparked conflicts over management of forests and natural resources (Patel et al., 2013), and in Sub-Saharan Africa climate mitigation projects have been accused of “green grabbing” — expropriating agricultural areas and grazing land in order to conserve nature or plant trees (Cavanagh & Benjaminsen, 2014). Hence, one may ask whether it is climate change or climate change mitigation that carry the largest conflict potential.

References

Abbass, I. M. (2012). No retreat no surrender: Conflict for survival between Fulani pastoralists and farmers in northern Nigeria. *European Scientific Journal*, *8*(1), 331–346.

Acock, A. (2014). *A Gentle Introduction to Stata* (4th ed.). Texas: Stata Press.

Afrobarometer. (2018). Online Data Analysis Tool: Benin, Botswana, Burkina Faso, Cameroon, Cape Verde, Côte d'Ivoire, Gabon, Gambia, Ghana, Guinea, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritius, Mozambique, Namibia, Niger, Nigeria, Sao Tome and Principe, Senegal, Sierra Leone, South Africa, Sudan, Swaziland, Tanzania, Togo, Uganda, Zambia, Zimbabwe, (R7 2016/2018). Retrieved 2020-06-09, from `http://afrobarometer.org/online-data-analysis/analyse-online`

Amankwaa, O. (2019). *Climate change and farmer-herder conflicts in the Asante Akim North District of Ghana* (Unpublished master's thesis). University of Oslo, Oslo.

Bargués Tobella, A. (2016). *The importance of tree cover for water resources in semiarid West Africa* (Doctoral thesis, Acta Universitatis Agriculturae Sueciae, Umeå). Retrieved 2020-05-26, from `https://pub.epsilon.slu.se/13553/` (ISBN: 9789157686480 ISSN: 1652-6880 Library Catalog: pub.epsilon.slu.se Num Pages: 81 Place: Umeå Volume: 2016:73)

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: doi:10.18637/jss.v067.i01.

Benjaminsen, T. A. (2008). Does Supply-Induced Scarcity Drive Violent Conflicts in the African Sahel? The Case of the Tuareg Rebellion in Northern Mali. *Journal of Peace Research*, *45*(6), 819–836. doi: 10.1177/0022343308096158

Benjaminsen, T. A., Alinon, K., Buhaug, H., & Buseth, J. T. (2012). Does climate change drive land-use conflicts in the Sahel? *Journal of Peace Research*, *49*(1), 97–111. doi: 10.1177/0022343311427343

Benjaminsen, T. A., & Ba, B. (2009). Farmer–herder conflicts, pastoral marginalisation and corruption: a case study from the inland Niger delta of Mali. *Geographical Journal*, *175*(1), 71–81. doi: 10.1111/j.1475-4959.2008.00312.x

Benjaminsen, T. A., Maganga, F. P., & Abdallah, J. M. (2009). The Kilosa Killings: Political Ecology of a Farmer–Herder Conflict in Tanzania. *Development and Change*, *40*(3), 423–445. doi: 10.1111/j.1467-7660.2009.01558.x

Berner Jr, E. (2009). *Evapotranspirasjon.* Retrieved 2020-06-12, from `https://snl.no/evapotranspirasjon`

Blomfield, A. (2018). The bloody cattle conflict pushing Nigeria to the edge of civil war. *The Telegraph.* Retrieved 2020-03-22, from `https://www.telegraph.co.uk/global-health/terror-and-security/battle-scarce-resources-muslim-cattle-herders-christian-farmers/`

Bronkhorst, S. (2011). *Climate change and conflict: lessons for conflict resolution from the southern Sahel of Sudan* (Tech. Rep.). Durban: Accord.

Brosché, J., & Elfversson, E. (2012). Communal conflict, civil war, and the state: Complexities, connections, and the case of Sudan. *African Journal on Conflict Resolution*, *12*(1), 9–32.

Buhaug, H. (2010). Climate not to blame for African civil wars. *Proceedings of the National Academy of Sciences*, *107*(38), 16477–16482. doi: 10.1073/pnas.1005739107

Buhaug, H., Benjaminsen, T. A., Sjaastad, E., & Theisen, O. M. (2015). Climate variability, food production shocks, and violent conflict in Sub-Saharan Africa. *Environmental Research Letters*, *10*(125015). doi: 10.1088/1748-9326/10/12/125015

Buhaug, H., & Rød, J. K. (2006). Local determinants of African civil wars, 1970–2001. *Political Geography*, *25*(3), 315–335. doi: 10.1016/j.polgeo.2006.02.005

Burke, M. B., Miguel, E., Satyanath, S., Dykema, J. A., & Lobell, D. B. (2009). Warming increases the risk of civil war in Africa. *Proceedings of the National Academy of Sciences*, *106*(49), 20670–20674. doi: 10.1073/pnas.0907998106

Busby, J. (2018). Taking Stock: the Field of Climate and Security. *Current Climate Change Reports*, *4*(4), 338–346. doi: 10.1007/s40641-018-0116-z

Butler, C. K., & Gates, S. (2012). African range wars: Climate, conflict, and property rights. *Journal of Peace Research*, *49*(1), 23–34.

Cavanagh, C., & Benjaminsen, T. A. (2014). Virtual nature, violent accumulation: The 'spectacular failure' of carbon offsetting at a Ugandan National Park. *Geoforum*, *56*, 55–65.

Cederman, L.-E., & Gleditsch, K. S. (2009). Introduction to Special Issue on "Disaggregating Civil War". *Journal of Conflict Resolution*, *53*(4), 487–495. doi: 10.1177/0022002709336454

Cederman, L.-E., Gleditsch, K. S., & Buhaug, H. (2013). *Inequality, Grievances, and Civil War.* Cambridge University Press.

Checchi, F., & Robinson, W. C. (2013). Mortality among populations of southern and central Somalia affected by severe food insecurity and famine during 2010–2012. *FAO/FSNAU and FEWS NET*. Retrieved 2020-05-19, from `https://www.fsnau.org/in-focus/study-report-mortality-among-populations-southern-and-central-somalia-affected-severe-food-`

Christophersen, K.-A. (2018). *Introduksjon til statistisk analyse: regresjonsbaserte metoder og anvendelse.* Oslo: Gyldendal akademisk.

Ciccone, A. (2011). Economic Shocks and Civil Conflict: A Comment. *American Economic Journal: Applied Economics*, *3*(4), 215–227. Retrieved 2020-05-26, from http://pubs.aeaweb.org/doi/10.1257/app.3.4.215 doi: 10.1257/app.3.4.215

CISEIN, & CIAT. (2005). *Gridded Population of the World, Version 3 (GPWv3): Population Count Grid.* Palisades, New York. Retrieved 2019-11-28, from grid.prio.org

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., ... Booth, B. B. B. (2013). Long-term Climate Change: Projections, Commitments and Irreversibility. *Climate Change 2013 - The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1029–1136. (Publisher: Cambridge University Press)

Cooper, P. J. M., Dimes, J., Rao, K. P. C., Shapiro, B., Shiferaw, B., & Twomlow, S. (2008). Coping better with current climatic variability in the rain-fed farming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? *Agriculture, Ecosystems & Environment*, *126*(1), 24–35. doi: 10.1016/j.agee.2008.01.007

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., ... Ziblatt, D. (2019). *"V-Dem Dataset v9." Varieties of Democracy (V-Dem) Project.* Retrieved 2019-11-03, from https://www.v-dem.net/en/data/archive/previous-data/data-version-9/

Corradin, C. (2016). Israel: Water as a tool to dominate Palestinians. *Al Jazeera*. Retrieved 2020-05-02, from https://www.aljazeera.com/news/2016/06/israel-water-tool-dominate-palestinians-160619062531348.html

Cubash, U., Wuebbles, D., Chen, D., Facchini, M. C., Frame, D., Mahowald, N., & Winther, J.-G. (2013). Introduction. In T. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, United Kingdom and New York, USA.: Cambridge University Press.

Dercon, S., & Gollin, D. (2014). Agriculture in African Development: Theories and Strategies. *Annual Review of Resource Economics*, *6*(1), 471–492. doi: 10.1146/annurev-resource-100913-012706

Detges, A. (2014). Close-up on renewable resources and armed conflict. *Political Geography*, *42*, 57–65. doi: 10.1016/j.polgeo.2014.06.003

Dick, B. (2009). *Vegetasjonsindeks.* Retrieved 2019-11-21, from https://snl.no/vegetasjonsindeks

Didan, K., Munoz, A. B., Solano, R., & Huete, A. (2015). *MODIS Vegetation Index User's Guide (MOD13 Series).* University of Arizona: Tucson, AZ.

Dube, T. (2019). Victoria Falls 'bone dry'? How drought and dry season are hitting the world's largest waterfall. *AFP Fact Check*. Retrieved 2020-02-26, from https://factcheck.afp.com/victoria-falls-bone-dry-how-drought-and-dry-season-are-hitting-worlds-largest-waterfall

Döring, S. (2020). Come rain, or come wells: How access to groundwater affects communal violence. *Political Geography*, *76*, 102073. doi: 10.1016/j.polgeo.2019.102073

Epstein, L., & King, G. (2002). The Rules of Inference. *The University of Chicago Law Review*, *69*(1), 1–133. doi: 10.2307/1600349

European Commission. (2018). *Uppsala Conflict Data Program.* Retrieved 2020-02-12, from https://ec.europa.eu/knowledge4policy/dataset/ds00121_en

FAO. (n.d.). *Crisis in the Horn of Africa.* Retrieved 2020-02-20, from http://www.fao.org/emergencies/crisis/horn-of-africa/en/

FAO. (2016). *AQUASTAT Main Database, Food and Agriculture Organization of the United Nations (FAO).* Retrieved 2020-04-12, from http://www.fao.org/aquastat/en/

FAO. (2017). The impact of disasters on agriculture. Retrieved 2019-11-09, from http://www.fao.org/3/a-i7279e.pdf

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

Fjelde, H. (2015). Farming or Fighting? Agricultural Price Shocks and Civil War in Africa. *World Development*, *67*, 525–534. doi: 10.1016/j.worlddev.2014.10.032

Fjelde, H., & von Uexkull, N. (2012). Climate triggers: Rainfall anomalies, vulnerability and communal conflict in Sub-Saharan Africa. *Political Geography*, *31*(7), 444–453. doi: 10.1016/j.polgeo.2012.08.004

Fjelde, H., & Østby, G. (2014). Socioeconomic Inequality and Communal Conflict: A Disaggregated Analysis of Sub-Saharan Africa, 1990–2008. *International Interactions*, *40*(5), 737–762. doi: 10.1080/03050629.2014.917373

Fotheringham, A. S., & Wong, W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A: Economy and Space*, *23*(7), 1025–1044.

Gall, M., Borden, K. A., & Cutter, S. L. (2009). When Do Losses Count?: Six Fallacies of Natural Hazards Loss Data. *Bulletin of the American Meteorological Society*, *90*(6), 799–810. doi: 10.1175/2008BAMS2721.1

Gamon, J. A., Field, C. B., Goulden, M. L., Griffin, K. L., Hartley, A. E., Joel, G., ... Valentini, R. (1995). Relationships Between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. *Ecological Applications*, *5*(1), 28–41. Retrieved 2020-05-27, from https://www.jstor.org/stable/1942049 (Publisher: Ecological Society of America) doi: 10.2307/1942049

Gartzke, E. (2012). Could climate change precipitate peace? *Journal of Peace Research*, *49*(1), 177–192.

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York: Cambridge University Press.

Gerring, J. (2004). What is a case study and what is it good for? *American political science review*, *98*(2), 341–354.

Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M., & Strand, H. (2002). Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research*, *39*(5), 615–637. doi: 10.1177/0022343302039005007

Gleick, P. H. (2014). Water, drought, climate change, and conflict in Syria. *Climate, and Society*, *6*(3), 331–340.

Grieco, J., Ikenberry, J. G., & Mastanduno, M. (2019). *Introduction to International Relations: Enduring Questions and Contemporary Perspectives.* London: Palgrave.

Guha-Sapir, D., Below, R., & Hoyois, P. (2016). *EM-DAT: International Disasters Database.* Brussels: Centre for Research on the Epidemiology of Disasters (CRED), Université Catholique de Louvain.

Harari, M., & Ferrara, E. L. (2018). Conflict, Climate, and Cells: A Disaggregated Analysis. *The Review of Economics and Statistics*, *100*(4), 594–608. doi: 10.1162/rest_a_00730

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., ... Zhai, P. (2013). Observations: Atmosphere and Surface. In T. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, United Kingdom and New York, USA.: Cambridge University Press.

Hegre, H., Ellingsen, T., Gates, S., & Gleditsch, N. P. (2001). Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816–1992. *American Political Science Review*, *95*(1), 33–48. doi: 10.1017/S0003055401000119

Hendrix, C. S., & Salehyan, I. (2012). Climate change, rainfall, and social conflict in Africa. *Journal of Peace Research*, *49*(1), 35–50. doi: 10.1177/0022343311426165

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, *81*(396), 945–960.

Homdee, T., Pongput, K., & Kanae, S. (2016). A comparative performance analysis of three standardized climatic drought indices in the Chi River basin, Thailand. *Agriculture and Natural Resources*, *50*(3), 211–219. doi: 10.1016/j.anres.2016.02.002

Homer-Dixon, T. F. (1999). *Environment, Scarcity, and Violence.* Princeton, New Jersey: Princeton University Press.

Hsiang, S. M., Burke, M., & Miguel, E. (2013). Quantifying the Influence of Climate on Human Conflict. *Science*, *341*(6151). doi: 10.1126/science.1235367

Hundie, B. (2010). Conflicts between Afar Pastoralists and their Neighbors: Triggers and Motivations. *International Journal of Conflict and Violence (IJCV)*, *4*(1), 134–148. doi: 10.4119/ijcv-2820

Högbladh, S. (2020). *UCDP GED Codebook version 20.1.* Department of Peace and Conflict Research, Uppsala University.

International Water Management Institute. (2010). *Managing water for rainfed agriculture* (Tech. Rep. No. 10). Colombo, Sri Lanka: International Water Management Institute (IWMI). (doi: 10.5337/2010.223)

Karlson, M., & Ostwald, M. (2016). Remote sensing of vegetation in the Sudano-Sahelian zone: A literature review from 1975 to 2014. *Journal of Arid Environments*, *124*, 257–269. doi: 10.1016/j.jaridenv.2015.08.022

Kellstedt, P. M., & Whitten, G. D. (2018). *The fundamentals of political science research* (3rd ed.). New York: Cambridge University Press.

Kennedy, P. (2008). *A guide to econometrics* (6th ed.). Malden, MA: Blackwell Pub.

Keyantash, J. & National Center for Atmospheric Research Staff. (2018). *The Climate Data Guide: Standardized Precipitation Index (SPI).* Retrieved 2020-01-11, from `https://climatedataguide.ucar.edu/climate-data/standardized-precipitation-index-spi`

Kim, J.-S., & Frees, E. W. (2006). Omitted Variables in Multilevel Models. *Psychometrika*, *71*(4), 659. doi: 10.1007/s11336-005-1283-0

King, G., Keohane, R. O., & Verba, S. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

Klein Goldewijk, K., Beusen, A., de Vos, M., & van Drecht, G. (2011). The HYDE 3.1 spatially explicit database of human induced land use change over the past 12,000 years. *Global Ecology and Biogeography*, *20*(1), 73–86. doi: 10.1111/j.1466-8238.2010.00587.x.

Klein Goldewijk, K., Beusen, A., & Janssen, P. (2010). Long term dynamic modeling of global population and built-up area in a spatially explicit way. *The Holocene*, *20*(4), 565–573. doi: 10.1177/0959683609356587

Koch, G. G. (2014). Intraclass Correlation Coefficient. In *Wiley StatsRef: Statistics Reference Online.* American Cancer Society.

Koren, O. (2018). Food Abundance and Violent Conflict in Africa. *American Journal of Agricultural Economics*, *100*(4), 981–1006. doi: 10.1093/ajae/aax106

Lantz, B. (2013). The large sample size fallacy. *Scandinavian Journal of Caring Sciences*, *27*(2), 487–492.

Lillesand, T., Kiefer, R., & Chipman, J. (2004). *Remote Sensing and Image Interpretation* (5th ed.). New York: John Wiley.

Linke, A. M., O'Loughlin, J., McCabe, J. T., Tir, J., & Witmer, F. D. W. (2015). Rainfall variability and violence in rural Kenya: Investigating the effects of drought and the role of local institutions with survey data. *Global Environmental Change*, *34*, 35–47. doi: 10.1016/j.gloenvcha.2015.04.007

Loebenstein, G., & Thottappilly, G. (2009). *The Sweetpotato.* Springer Science & Business Media.

Mach, K. J., Kraan, C. M., Adger, W. N., Buhaug, H., Burke, M., Fearon, J. D., ... von Uexkull, N. (2019). Climate as a risk factor for armed conflict. *Nature*, *571*(7764), 193–197. doi: 10.1038/s41586-019-1300-6

Maxwell, D., & Fitzpatrick, M. (2012). The 2011 Somalia famine: Context, causes, and complications. *Global Food Security*, *1*(1), 5–12. doi: 10.1016/j.gfs.2012.07.002

McDougal, T., Hagerty, T., Inks, L., Dowd, C., & Conroy, S. (2015). Macroeconomic benefits of farmer-pastoralist peace in Nigeria's Middle Belt: An input-output analysis approach. *The Economics of Peace and Security Journal*, *10*(1). doi: 10.15355/epsj.10.1.66

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and dureation to time scales. *Proceedings of the 8th Conference on Applied Climatology*, *17*(22), 179–183.

Meier, P., Bond, D., & Bond, J. (2007). Environmental influences on pastoral conflict in the Horn of Africa. *Political Geography*, *26*(6), 716–735. doi: 10.1016/j.polgeo.2007.06.001

Miguel, E., & Satyanath, S. (2011). Re-examining Economic Shocks and Civil Conflict. *American Economic Journal: Applied Economics*, *3*(4), 228–232. doi: 10.1257/app.3.4.228

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy*, *112*(4), 725–753. doi: 10.1086/421174

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1-2), 202–216. doi: 10.1016/j.jhydrol.2010.07.012

Mjøs, O. D. (2007). *The Nobel Peace Prize 2009: Award ceremony speech.* Retrieved 2019-08-22, from `https://www.nobelprize.org/prizes/peace/2007/ceremony-speech/`

Naumann, G., Alfieri, L., Wyser, K., Mentaschi, L., Betts, R. A., Carrao, H., ... Feyen, L. (2018). Global Changes in Drought Conditions Under Different Levels of Warming. *Geophysical Research Letters*, *45*(7), 3285–3296. doi: 10.1002/2017GL076521

Nordkvelle, J., Rustad, S. A., & Salmivalli, M. (2017). Identifying the effect of climate variability on communal conflict through randomization. *Climatic Change*, *141*(4), 627–639. doi: 10.1007/s10584-017-1914-3

Obama, B. H. (2009). *Nobel Lecture: A Just and Lasting Peace.* Retrieved 2019-08-22, from `https://www.nobelprize.org/nobel_prizes/peace/laureates/2009/obama-lecture_en.html`

O'Loughlin, J., Linke, A. M., & Witmer, F. D. (2014). Effects of temperature and precipitation variability on the risk of violence in sub-Saharan Africa, 1980–2012. *Proceedings of the National Academy of Sciences*, *111*(47), 16712–16717.

O'Loughlin, J., Witmer, F. D. W., Linke, A. M., Laing, A., Gettelman, A., & Dudhia, J. (2012). Climate variability and conflict risk in East Africa, 1990-2009. *Proceedings of the National Academy of Sciences*, *109*(45), 18344–18349. doi: 10.1073/pnas.1205130109

Onapajo, H., & Usman, A. A. (2015). Fuelling the Flames: Boko Haram and Deteriorating Christian–Muslim Relations in Nigeria. *Journal of Muslim Minority Affairs*, *35*(1), 106–122. doi: 10.1080/13602004.2015.1007667

Patel, T., Dhiaulhaq, A., Gritten, D., Yasmi, Y., De Bruyn, T., Paudel, N. S., ... Suzuki, R. (2013). Predicting future conflict under REDD+ implementation. *Forests*, *4*(2), 343–363.

Pedro-Monzonís, M., Solera, A., Ferrer, J., Estrela, T., & Paredes-Arquiola, J. (2015). A review of water scarcity and drought indexes in water resources planning and management. *Journal of Hydrology*, *527*, 482–493.

Peterson, B. D. (2016). *The Ethnic Power Relations Data: A Critique* (Tech. Rep.). Charlottesville, VA: University of Virginia.

Pettersson, T. (2010). Non-state conflicts 1989-2008: global and regional patterns. In T. Pettersson & L. Themnér (Eds.), *States in armed conflict 2009* (pp. 183–202). Uppsala, Department of Peace and Conflict Research.

Pettersson, T., Högbladh, S., & Öberg, M. (2019). Organized violence, 1989–2018 and peace agreements. *Journal of Peace Research*, *56*(4), 589–603. doi: 10.1177/0022343319856046

Raleigh, C., & Kniveton, D. (2012). Come rain or shine: An analysis of conflict and climate variability in East Africa. *Journal of Peace Research*, *49*(1), 51–64.

Rembold, F., Oduori, S. M., Gadain, H., & Toselli, P. (2013). Mapping charcoal driven forest degradation during the main period of Al Shabaab control in Southern Somalia. *Energy for Sustainable Development*, *17*(5), 510–514. doi: 10.1016/j.esd.2013.07.001

Rosvold, E., & Buhaug, H. (2020). *Geocoded Disasters: A New Dataset.* Oslo: Peace Research Institute Oslo (PRIO).

Rouse, J. W., Haas, R. H., & Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. *3rd ERTS Symposium, NASA SP-351 I*, 309–317.

Salehyan, I. (2014). Climate change and conflict: Making sense of disparate findings. *Political Geography*, *43*, 1–5. doi: 10.1016/j.polgeo.2014.10.004

Salehyan, I., & Hendrix, C. S. (2014). Climate shocks and political violence. *Global Environmental Change*, *28*, 239–250. doi: 10.1016/j.gloenvcha.2014.07.007

Scheff, J., & Frierson, D. M. W. (2014). Scaling Potential Evapotranspiration with Greenhouse Warming. *Journal of Climate*, *27*(4), 1539–1558. doi: 10.1175/JCLI-D-13-00233.1

Schultz, K. A., & Mankin, J. S. (2019). Is Temperature Exogenous? The Impact of Civil Conflict on the Instrumental Climate Record in Sub-Saharan Africa. *American Journal of Political Science*, *63*(4), 723–739. doi: 10.1111/ajps.12425

Seter, H. (2016). Connecting climate variability and conflict: Implications for empirical testing. *Political Geography*, *53*, 1–9. doi: 10.1016/j.polgeo.2016.01.002

Shettima, A. G., & Tar, U. A. (2008). Farmer-pastoralist conflict in West Africa: exploring the causes and consequences. *Information, society and justice journal*, *1*(2), 163–184.

Shiru, M. S., Shahid, S., Alias, N., & Chung, E.-S. (2018). Trend Analysis of Droughts during Crop Growing Seasons of Nigeria. *Sustainability*, *10*(3), 871. doi: 10.3390/su10030871

Smits, J., & Permanyer, I. (2019). The Subnational Human Development Database. *Scientific Data*, *6*(1). Retrieved 2020-03-01, from `https://globaldatalab.org/shdi/shdi/` doi: 10.1038/sdata.2019.38

Solow, A. R. (2013). A call for peace on climate and conflict. *Nature*, *497*(7448), 179–180. doi: 10.1038/497179a

Sommet, N., & Morselli, D. (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, *30*, 203–218. doi: 10.5334/irsp.90

Stock, J. H., & Watson, M. W. (2011). *Introduction to econometrics* (3rd ed.). Boston: Pearson Education Inc.

Sulieman, H., & Young, H. (2019). Transforming pastoralist mobility in West Darfur: Understanding Continuity and Change. *Policy and Practice*, *2*, 26.

Sundberg, R., Eck, K., & Kreutz, J. (2012). Introducing the UCDP Non-State Conflict Dataset. *Journal of Peace Research*, *49*(2), 351–362. Retrieved 2019-09-10, from `https://ucdp.uu.se/` doi: 10.1177/0022343311431598

Sundberg, R., & Melander, E. (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, *50*(4), 523–532. Retrieved 2019-09-10, from `https://ucdp.uu.se/` doi: 10.1177/0022343313484347

Tannehill, I. R. (1947). *Drought, its causes and effects* (Vol. 64) (No. 1). Princeton University Press.

Theisen, O. M. (2012). Climate clashes? Weather variability, land pressure, and organized violence in Kenya, 1989–2004. *Journal of Peace Research*, *49*(1), 81–96. doi: 10.1177/0022343311425842

Theisen, O. M., Gleditsch, N. P., & Buhaug, H. (2013). Is climate change a driver of armed conflict? *Climatic Change*, *117*(3), 613–625. doi: 10.1007/s10584-012-0649-4

Theisen, O. M., Holtermann, H., & Buhaug, H. (2011). Climate wars? Assessing the claim that drought breeds conflict. *International Security*, *36*(3), 79–106. doi: https://doi.org/10.1162/ISEC_a_00065

Tollefsen, A. F., & Buhaug, H. (2015). Insurgency and inaccessibility. *International Studies Review*, *17*(1), 6–25.

Tollefsen, A. F., Strand, H., & Buhaug, H. (2012). PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, *49*(2), 363–374. doi: 10.1177/0022343311431287

Tollefsen, A. F., Vestby, J., Landsverk, P. G., Larsen, V. B., & Bahgat, K. (2020). PRIO-GRID v3.0 dataset. *GitHub repository*. Retrieved 2020-01-09, from `https://github.com/prio-data/priogrid`

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, *4*(1), 17–22. doi: 10.1038/nclimate2067

Turner, M. D. (2004). Political ecology and the moral dimensions of “resource conflicts”: the case of farmer–herder conflicts in the Sahel. *Political Geography*, *23*(7), 863–889. doi: 10.1016/j.polgeo.2004.05.009

Turner, M. D., Ayantunde, A. A., Patterson, K. P., & Patterson, E. D. (2011). Livelihood Transitions and the Changing Nature of Farmer–Herder Conflict in Sahelian West Africa. *The Journal of Development Studies*, *47*(2), 183–206. doi: 10.1080/00220381003599352

van Baalen, S., & Mobjörk, M. (2016). *A coming anarchy?: Pathways from climate change to violent conflict in East Africa.* Stockholm: SIPRI & Stockholm University.

van Weezel, S. (2015). Economic shocks & civil conflict onset in Sub-Saharan Africa, 1981–2010. *Defence and Peace Economics*, *26*(2), 153–177. doi: 10.1080/10242694.2014.887489

Vestby, J. (2018). *Climate, development, and conflict: Learning from the past and mapping uncertainties of the future* (Phd). University of Oslo, Oslo.

Vestby, J. (2019). Climate variability and individual motivations for participating in political violence. *Global Environmental Change*, *56*, 114–123. doi: 10.1016/j.gloenvcha.2019.04.001

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, *23*(7), 1696–1718. doi: 10.1175/2009JCLI2909.1

Vogt, M., Bormann, N.-C., Rüegger, S., Cederman, L.-E., Hunziker, P., & Girardin, L. (2015). Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Dataset Family. *Journal of Conflict Resolution*, *59*(7), 1327–1342. doi: doi:10.1177/0022002715591215

von Uexkull, N. (2014). Sustained drought, vulnerability and civil conflict in Sub-Saharan Africa. *Political Geography*, *43*, 16–26. doi: 10.1016/j.polgeo.2014.10.003

von Uexkull, N., Croicu, M., Fjelde, H., & Buhaug, H. (2016). Civil conflict sensitivity to growing-season drought. *Proceedings of the National Academy of Sciences*, *113*(44), 12391–12396. doi: 10.1073/pnas.1607542113

von Uexkull, N., d'Errico, M., & Jackson, J. (2020). Drought, Resilience, and Support for Violence: Household Survey Evidence from DR Congo. *Journal of Conflict Resolution*, 28. doi: 10.1177/0022002720923400

von Uexkull, N., & Pettersson, T. (2018). Issues and Actors in African Nonstate Conflicts: A New Data Set. *International Interactions*, *44*(5), 953–968. doi: 10.1080/03050629.2018.1493478

Vågen, T.-G., & Gumbricht, T. (2012). *Sahel atlas of changing landscapes: tracing trends and variations in vegetation cover and soil condition*. UNEP. Retrieved 2020-05-20, from `https://wedocs.unep.org/handle/20.500.11822/8599`

Wan, Z., Wang, P., & Li, X. (2004). Using MODIS Land Surface Temperature and Normalized Difference Vegetation Index products for monitoring drought in the southern Great Plains, USA. *International Journal of Remote Sensing*, *25*(1), 61–72. doi: 10.1080/0143116031000115328

Wig, T., & Kromrey, D. (2018). Which groups fight? Customary institutions and communal conflicts in Africa. *Journal of Peace Research*, *55*(4), 415–429. doi: 10.1177/0022343317740416

Wilhite, D. A. (2000). Chapter 1: Drought as a Natural Hazard: Concepts and Definitions. In D. A. Wilhite (Ed.), *Drought: A Global Assessment* (Vol. 1, pp. 3–18). London: Routledge.

Wilhite, D. A., & Glantz, M. H. (1985). Understanding the Drought Phenomenon: The Role of Definitions. *Water International*, *10*(3), 111–120.

Witsenburg, K. M., & Adano, W. R. (2009). Of Rain and Raids: Violent Livestock Raiding in Northern Kenya. *Civil Wars*, *11*(4), 514–538. doi: 10.1080/13698240903403915

Young, O. R. (2016). The Paris Agreement: Destined to Succeed or Doomed to Fail? *Politics and Governance*, *4*(3), 124–132. doi: https://doi.org/10.17645/pag.v4i3.635

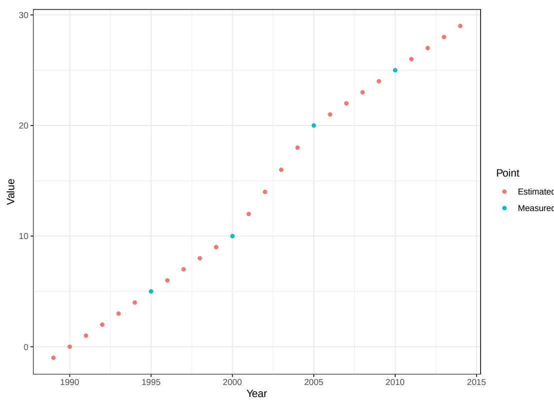
Appendix A

Appendix

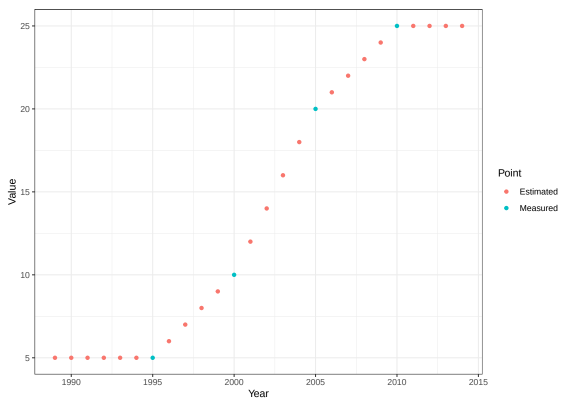
Table A.1: Geo-precision from UCDP-GED

Precision level	Number of conflicts	Known point
1	422 (42.4%)	Exact coordinates
2	187 (18.8%)	Within 25km from exact location
3	216 (21.7%)	Second order administrative unit
4	121 (12.2%)	First order administrative unit
5	42 (4.2%)	Section of country (e.g. Northern Uganda), or linear feature (e.g. long river, border, longer road)
6	7 (0.7%)	Country

Source: Högbladh (2020)



(a) Linear extrapolation



(b) “Next observation carried backward”

Figure A.1: Different extrapolation techniques

Figure A.2: NDVI variation

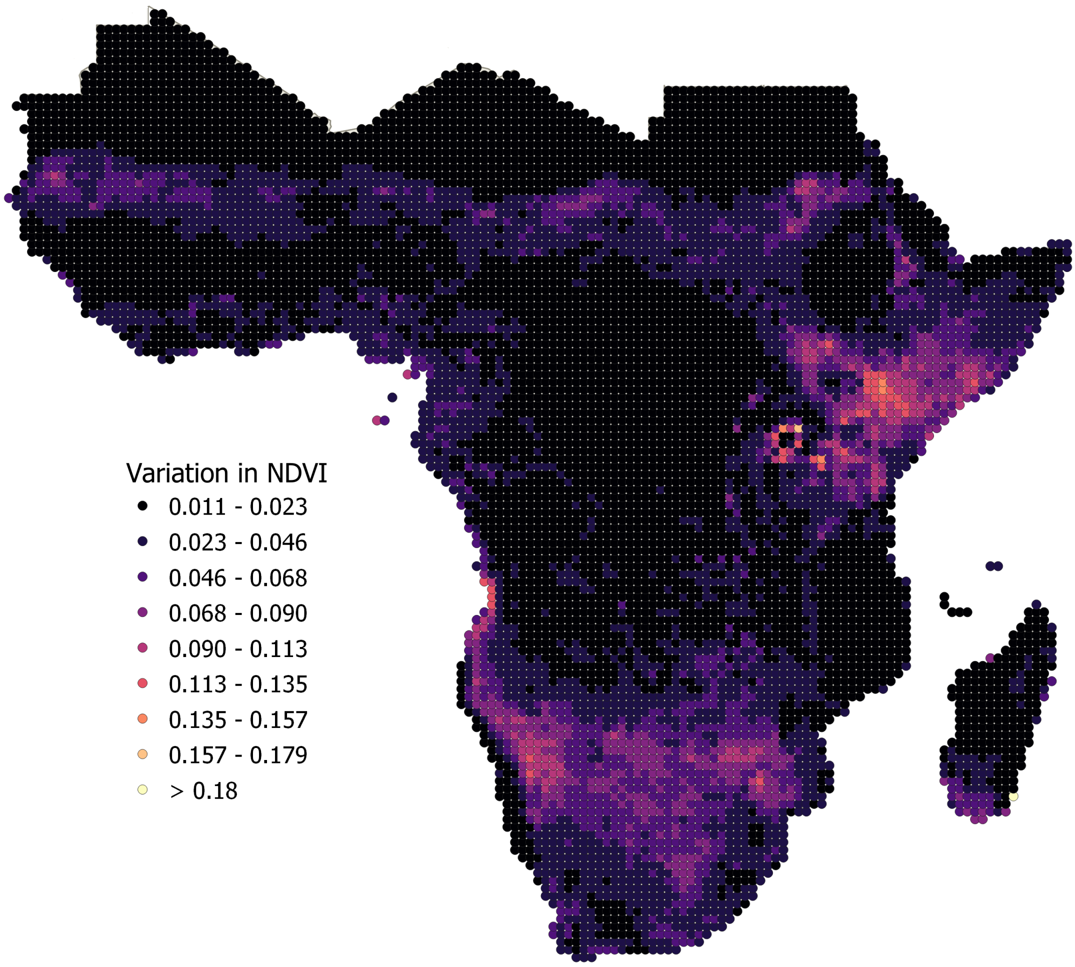
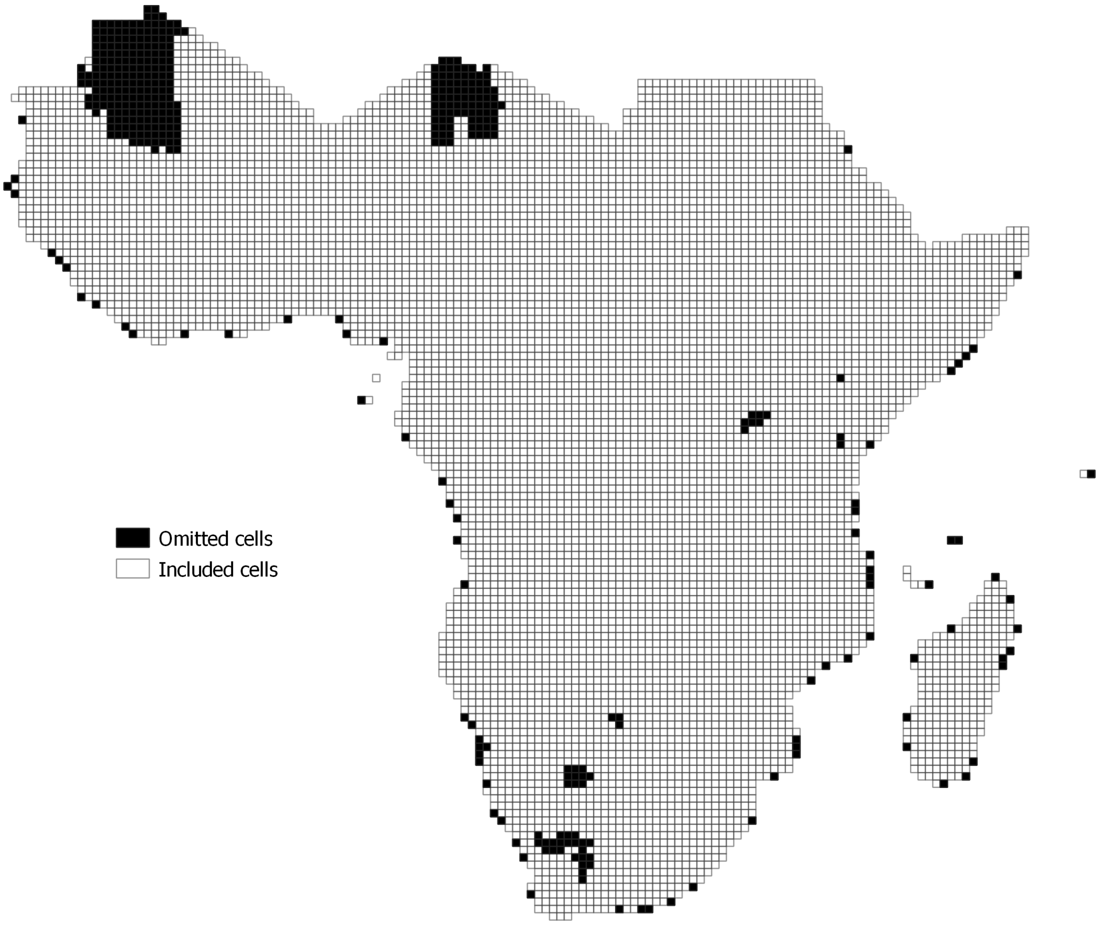
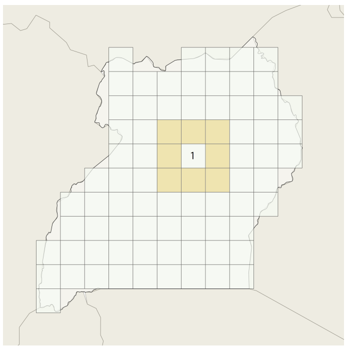


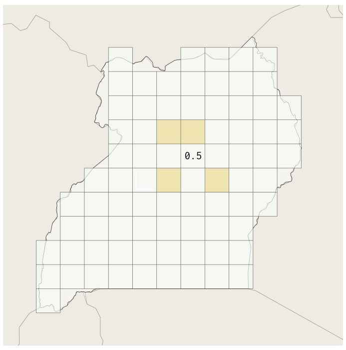
Figure A.3: Cells included in the analysis



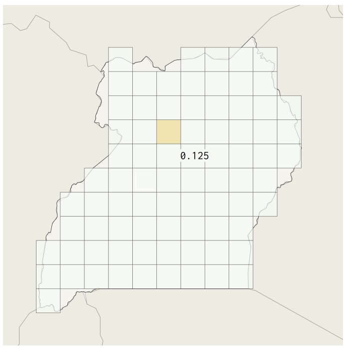
Cells with less than 200 inhabitants or 100km^2 are omitted. In total 424 cells are omitted and the sample contains $8,410 - 424 = 7,986 cells$.



(a) $\frac{8}{8}$ neighbours have drought



(b) $\frac{4}{8}$ neighbours have drought



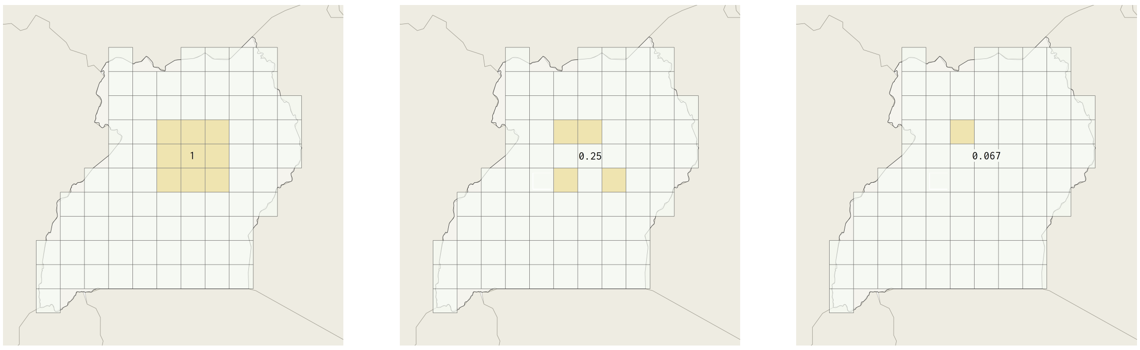
(c) $\frac{1}{8}$ neighbours have drought

Figure A.4: Indirect effect of drought

Table A.2: Multilevel regression 2000-2014

	Models								
	(1-3) Direct			**(4-6) Indirect**			**(7-9) Total**		
	(1) SPEI	(2) NDVI	(3) EM-DAT	(4) SPEI	(5) NDVI	(6) EM-DAT	(7) SPEI	(8) NDVI	(9) EM-DAT
SPEI	1.53*			1.92**			1.76**		
	(0.27)			(0.42)			(0.36)		
NDVI		1.39*			1.92**			1.70**	
		(0.20)			(0.39)			(0.31)	
EM-DAT			2.18***			2.33***			2.31***
			(0.26)			(0.31)			(0.30)
Population_{log}	1.65***	1.65***	1.62***	1.66***	1.65***	1.62***	1.65***	1.65***	1.62***
	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)
SHDI	1.14	1.06	1.05	1.19	1.15	1.05	1.18	1.11	1.05
	(0.87)	(0.81)	(0.80)	(0.91)	(0.87)	(0.80)	(0.90)	(0.84)	(0.80)
Democracy	0.05***	0.05***	0.04***	0.04***	0.05***	0.04***	0.05***	0.05***	0.04***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Ethnic exclusion	1.28	1.29	1.29	1.27	1.28	1.29	1.28	1.28	1.29
	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
$\text{Conflict event}_{t-1}$	4.75***	4.81***	5.07***	4.74***	4.90***	5.11***	4.74***	4.85***	5.10***
	(0.64)	(0.65)	(0.69)	(0.64)	(0.66)	(0.70)	(0.64)	(0.66)	(0.70)
Intercept	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	112,207	112,099	112,410	112,380	112,380	112,380	112,192	1120,84	112,380
No of groups	7,482	7,487	7,494	7,492	7,492	7,492	7,481	7,486	7,492
LL	-3,114	-3,114	-3,097	-3,113	-3,112	-3,098	-3,113	-3,112	-3,097
AIC	6,244	6,244	6,210	6,242	6,240	6,213	6,242	6,241	6,210

*: p <0.05, **: p <0.01, ***: p <0.001. All models are estimated through multilevel mixed-effects logistic regression with observations nested on cells. Coefficients are shown as odds ratios with standard errors in parentheses. Cells with less than 200 inhabitants or 100 km^2 are omitted.



(a) $\frac{8}{8}$ neighbours have drought

(b) $\frac{4}{8}$ neighbours have drought

(c) $\frac{1}{8}$ neighbours have drought

Figure A.5: Total effect of drought

Table A.3: Multilevel regression with continuous, cumulative and lagged drought

Continuous SPEI and NDVI values are inverted.

	Models								
	(1-3) Continuous			**(4-6) Cumulative**			**(7-9) Lagged t-1**		
	(1) SPEI	(2) NDVI†	(3) EM-DAT‡	(4) SPEI	(5) NDVI	(6) EM-DAT	(7) SPEI	(8) NDVI	(9) EM-DAT
SPEI	1.09*			1.14			1.13		
	(0.05)			(0.13)			(0.17)		
NDVI		1.11			1.18			1.17	
		(0.10)			(0.12)			(0.18)	
EM-DAT			1.88***			1.47***			1.27*
			(0.20)			(0.11)			(0.14)
Population$_{log}$	1.87***	1.64***	1.84***	1.87***	1.65***	1.84***	1.86***	1.62***	1.85***
	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)	(0.09)
SHDI	2.74	1.07	2.74	2.62	1.04	2.78	2.59	1.80	2.49
	(1.71)	(0.82)	(1.70)	(1.64)	(0.79)	(1.72)	(1.62)	(1.41)	(1.55)
Democracy	0.17***	0.05***	0.15***	0.17***	0.05***	0.15***	0.16***	0.04***	0.16***
	(0.07)	(0.02)	(0.06)	(0.07)	(0.02)	(0.06)	(0.07)	(0.02)	(0.07)
Ethnic exclusion	1.19	1.29	1.17	1.19	1.29	1.17	1.19	1.23	1.18
	(0.14)	(0.17)	(0.13)	(0.14)	(0.17)	(0.13)	(0.14)	(0.17)	(0.13)
Conflict event$_{t-1}$	5.69***	4.79***	5.90***	5.70***	4.77***	5.87***	5.70***	5.21***	5.67***
	(0.66)	(0.65)	(0.69)	(0.66)	(0.64)	(0.68)	(0.66)	(0.74)	(0.66)
Intercept	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	185,865	112,099	186,190	185,865	112,099	186,190	185,865	104,612	186,190
No of groups	7,482	7,487	7,494	7,482	7,487	7,494	7,482	7,487	7,494
LL	-4,114	-3,116	-4,099	-4,115	-3,115	-4,105	-4,115	-2,920	-4,114
AIC									

*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$. All models are estimated through multilevel mixed-effects logistic regression with observations nested on cells. Coefficients are shown as odds ratios with standard errors in parentheses. Cells with less than 200 inhabitants or 100 km^2 are omitted.

† Original NDVI values are multiplied by 10 in order to be similar to the SPEI values.

‡ EM-DAT values are not continuous, but similar to the ones used in the main model (binary).

Table A.4: Multilevel regression with geo-precision level 1

	Models								
	(1-3) Direct			**(4-6) Indirect**			**(7-9) Total**		
	(1) SPEI	(2) NDVI	(3) EM-DAT	(4) SPEI	(5) NDVI	(6) EM-DAT	(7) SPEI	(8) NDVI	(9) EM-DAT
SPEI	1.29			1.50			1.42		
	(0.28)			(0.40)			(0.36)		
NDVI		1.72*			2.54**			2.23**	
		(0.36)			(0.75)			(0.59)	
EM-DAT			2.34***			2.45***			2.46***
			(0.36)			(0.43)			(0.41)
Population$_{log}$	1.76***	1.60***	1.76***	1.77***	1.61***	1.76***	1.76***	1.61***	1.76***
	(0.10)	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)
SHDI	1.87	0.61	2.26	1.89	0.63	2.25	1.90	0.62	2.28
	(1.48)	(0.59)	(1.79)	(1.50)	(0.60)	(1.79)	(1.50)	(0.59)	(1.81)
Democracy	0.12***	0.08***	0.11***	0.12***	0.08***	0.11***	0.12***	0.08***	0.11***
	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)
Ethnic exclusion	1.08	1.02	1.06	1.08	1.01	1.06	1.08	1.02	1.06
	(0.17)	(0.18)	(0.16)	(0.17)	(0.18)	(0.16)	(0.17)	(0.18)	(0.16)
Conflict event$_{t-1}$	12.66***	15.25***	12.90***	12.61***	15.36***	12.88***	12.64***	15.33***	12.89***
	(2.43)	(3.72)	(2.48)	(2.43)	(3.72)	(2.47)	(2.43)	(3.73)	(2.48)
Intercept	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	193,815	117,660	195,689	195,639	118,003	195,639	193,790	117,645	195,639
No of groups	7,821	7,885	7,896	7,894	7,894	7,894	7,820	7,884	7,894
LL	-2,014	-1,484	-2,002	-2,014	-1,482	-2,004	-2,014	-1,482	-2,003
AIC	4,044	2,983	4,020	4,045	2,981	4,024	4,044	2,981	4,021

*: p <0.05, **: p <0.01, ***: p <0.001. All models are estimated through multilevel mixed-effects logistic regression with observations nested on cells. Coefficients are shown as odds ratios with standard errors in parentheses. Cells with less than 200 inhabitants or 100 km^2 are omitted.

Table A.5: Multilevel regression without conflict id 4895

	Models								
	(1-3) Direct			**(4-6) Indirect**			**(7-9) Total**		
	(1) SPEI	(2) NDVI	(3) EM-DAT	(4) SPEI	(5) NDVI	(6) EM-DAT	(7) SPEI	(8) NDVI	(9) EM-DAT
SPEI	1.09			1.28			1.18		
	(0.17)			(0.25)			(0.22)		
NDVI		1.39*			1.99***			1.72**	
		(0.20)			(0.40)			(0.31)	
EM-DAT			1.93***			2.05***			2.04***
			(0.20)			(0.24)			(0.23)
Population$_{log}$	1.78***	1.56***	1.75***	1.78***	1.56***	1.75***	1.78***	1.56***	1.75***
	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)
SHDI	2.90	1.21	3.11	2.98	1.32	3.15	2.95	1.26	3.15
	(1.82)	(0.93)	(1.94)	(1.87)	(1.01)	(1.96)	(1.85)	(0.97)	(1.96)
Democracy	0.12***	0.04***	0.11***	0.12***	0.04***	0.11***	0.12***	0.04***	0.11***
	(0.05)	(0.02)	(0.05)	(0.05)	(0.02)	(0.05)	(0.05)	(0.02)	(0.05)
Ethnic exclusion	1.22	1.35*	1.19	1.21	1.35*	1.19	1.21	1.35*	1.19
	(0.14)	(0.18)	(0.14)	(0.14)	(0.18)	(0.14)	(0.14)	(0.18)	(0.14)
Conflict event$_{t-1}$	6.24***	5.36***	6.48***	6.24***	5.48***	6.51***	6.24***	5.42***	6.50***
	(0.74)	(0.75)	(0.77)	(0.74)	(0.77)	(0.78)	(0.74)	(0.76)	(0.78)
Intercept	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	185,825	112,064	186,150	186,100	112,345	186,100	185,800	112,049	186,100
No of groups	7,482	7,487	7,494	7,492	7,492	7,492	7,481	7,486	7,492
LL	-3,962	-2,979	-3,944	-3,961	-2,977	-3,945	-3,961	-2,977	-3,943
AIC	7,939	5,974	7,903	7,938	5,969	7,905	7,939	5,971	7,903

*: p <0.05, **: p <0.01, ***: p <0.001. All models are estimated through multilevel mixed-effects logistic regression with observations nested on cells. Coefficients are shown as odds ratios with standard errors in parentheses. Cells with less than 200 inhabitants or 100 km^2 are omitted.

Table A.6: Multilevel regression with 2° × 2° cells

	Models		
	(1) SPEI	(2) NDVI	(3) EM-DAT
SPEI	1.40 (0.37)		
NDVI		1.11 (0.14)	
EM-DAT			1.42* (0.20)
Population_{log}	2.48*** (0.26)	2.26*** (0.26)	2.42*** (0.25)
SHDI	1.11 (1.19)	0.14 (0.21)	1.12 (1.19)
Democracy	0.11*** (0.08)	0.01*** (0.01)	0.10*** (0.07)
Ethnic exclusion	0.093 (0.15)	1.00 (0.20)	0.92 (0.15)
Conflict event_{t-1}	4.87*** (0.64)	3.86*** (0.62)	4.95*** (0.65)
Intercept	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Observations	15,200	8,734	15,200
No of groups	588	584	588
LL	-1,597	-1,080	-1,594
AIC	3,210	2,176	3,205

*: p <0.05, **: p <0.01, ***: p <0.001. All models are estimated through multilevel mixed-effects logistic regression with observations nested on cells. Coefficients are shown as odds ratios with standard errors in parentheses. Cells with less than 200 inhabitants or 100 km^2 are omitted.

Table A.7: OLS regression with fixed effects

To make the table readable all coefficients and standard errors are multiplied by 100

	Models								
	(1-3) Direct			**(4-6) Indirect**			**(7-9) Total**		
	(1) SPEI	(2) NDVI	(3) EM-DAT	(4) SPEI	(5) NDVI	(6) EM-DAT	(7) SPEI	(8) NDVI	(9) EM-DAT
SPEI	0.12*			0.21**			0.17**		
	(0.06)			(0.07)			(0.07)		
NDVI		-0.02			-0.00			-0.01	
		(0.11)			(0.13)			(0.13)	
EM-DAT			0.26***			0.24**			0.26***
			(0.08)			(0.08)			(0.08)
Population_{log}	0.75***	0.44	0.75***	0.75***	0.44	0.75***	0.76***	0.44	0.75***
	(0.11)	(0.28)	(0.11)	(0.11)	(0.28)	(0.11)	(0.11)	(0.28)	(0.11)
SHDI	-0.58	-0.76	-0.60	-0.56	-0.76	-0.59	-0.57	-0.76	-0.60
	(0.62)	(0.10)	(0.62)	(0.61)	(0.10)	(0.0062)	(0.62)	(0.10)	(0.62)
Democracy	0.31	1.02*	0.31	0.31	1.02*	0.31	0.31	1.02*	0.31
	(0.24)	(0.44)	(0.24)	(0.24)	(0.43)	(0.24)	(0.24)	(0.44)	(0.24)
Ethnic exclusion	-0.00	-0.03	-0.01	-0.01	-0.03	-0.01	-0.00	-0.03	-0.01
	(0.08)	(0.12)	(0.08)	(0.08)	(0.12)	(0.08)	(0.08)	(0.12)	(0.08)
Conflict event_{t-1}	14.53***	10.55***	14.53***	14.53***	10.55***	14.53***	14.53***	10.55***	14.53***
	(1.76)	(1.93)	(1.76)	(1.76)	(1.93)	(1.76)	(1.76)	(1.93)	(1.76)
Intercept	-6.96***	-3.85	-6.9***	-6.98***	-3.83	-6.88***	-6.98***	-3.85	-6.88***
	(0.97)	(2.60)	(0.97)	(0.97)	(2.59)	(0.96)	(0.97)	(2.60)	(0.96)
Observations	185,865	112,099	186,190	186,140	112,380	186,140	185,840	112,084	186,140
No of groups	7,482	7,487	7,494	7,492	7,492	7,492	7,481	7,486	7,492
R^2 within	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.01	0.02
R^2 between	0.12	0.21	0.12	0.12	0.21	0.12	0.12	0.21	0.12
R^2 overall	0.03	0.05	0.03	0.03	0.05	0.03	-0.03	0.05	0.03

*: p <0.05, **: p <0.01, ***: p <0.001. All models are estimated with OLS with fixed effects on the cell level. Robust standard errors clustered on the cell level is provided in parentheses. Cells with less than 200 inhabitants or 100 km^2 are omitted.