Nijmegen School of Management
Department of Economics and Business Economics
Master's Thesis Economics (MAN-MTHEC)

Climate change as a driver of migration?

A sub-national empirical study on the extent to which climate change is a driver of migration in Africa.

By Dirk Arts (s1028419)

Nijmegen, 26 June 2023

Program: Master's Program in Economics
Specialisation: International Economics & Development
Supervisor: dr. G. Solano

Radboud Universiteit

Abstract

By being the first multi-country research that explores the extent to which climate change is a migratory driver at the sub-national level, this research provides new and distinctive evidence on the relationship between climate change, in particular gradual changes in temperature and precipitation, and migration. The empirical analysis of a newly-created dataset on 529 African sub-national regions observed from 2000 through 2020 reveals that temperature increases and precipitation changes are significantly and positively related to increased migration. Namely, a one percent increase in regional temperature over 5 years is related to an increase of 0.31 percent in migration from that region. Next to that, a change in precipitation over 5 years is linked to migration through a non-linear relationship, in which decreases as well as precipitation increases up to 55 percent lead to increased migration, while decreases in precipitation sort stronger effects on migration than increases in precipitation. An investigation into differences between regions finds that found effects are stronger for relatively unwealthy regions, and only hold for relatively agriculture-dependent regions, which suggests agriculture to be a channel through which climate change affects migration and resilience to be a moderator of the relationship. Finally, for relatively wet regions, only precipitation increases are related to increased migration, while for relatively dry regions, both decreases and increases are related to increased migration.

Table of contents

1. Introduction

Migration and climate change are two of the largest global challenges humanity is currently facing (United Nations, 2023). In the short- and long-term, climate change could result in certain regions on Earth becoming less inhabited due to changing regional climates, in particular changing temperatures and precipitation. In turn, this could lead to major migration. Already in 2011, about 500 million people thought that they would need to move within the next five years because of environmental problems (Esipova et al, 2011). With climate change only intensifying since then, this number is projected to have only increased (Lee et al., 2023; IEP, 2022). To assess whether this is indeed going to happen, one could investigate whether this already has been happening, as climate change-induced temperature and precipitation changes have already been notably present since the 1980s (IPCC, 2021; Yavçan, 2021). Therefore, this thesis aims to answer the following research question: To what extent is climate change a migratory driver in low- and middle-income countries? I will seek to answer this question by investigating the effect of sub-national regional temperature and precipitation changes on migration from that region in Africa.

To this end, I focus on migration due to gradual changes in climate, in particular temperature and precipitation changes. Next to that, the focus of this research is on Africa because only one continent could be covered due to limited resources, and Africa was chosen because of the unique characteristics that make it especially interesting for this research, which is further explained in the methodology.

The research question is a particularly relevant and valuable question to answer because policymakers can use the results in combination with climate models to predict future population distributions and which regions will experience immigration or emigration and act accordingly to mitigate adverse consequences for affected populations. For instance, if the results would suggest that regional temperature and precipitation changes do not result in migration from that region, policymakers could consider to investing additionally in adaptation and resilience to these and future climatic changes. On the other hand, if the results would suggest that regional temperature and precipitation changes do indeed result in migration, policymakers could consider supporting regions that are suggested to be relatively more attractive in welcoming migrants after temperature and precipitation changes.

The contribution of this research is twofold. First, this is the first multi-country research that explores the extent to which climate change is a migratory driver at the sub-national level. This is a key contribution because people migrate more within borders than across them, and temperature and precipitation levels can differ significantly across regions in a country (UNDP, 2009; IOM, 2020). Therefore, significantly more migratory movements are detected by measuring migration at the sub-national level. This is made possible by the novel sub-national dataset on temperature, precipitation, and population distribution in Africa used in this research. By using gridded reanalysis satellite data

(MERRA-2) on temperature and precipitation, measurements are standardized and continuous, which results in a balanced panel across time and space. This allows the investigation of the effects of climate change on migration over time at the regional level, while still comparable across regions and countries. Secondly, while gradual climate change is increasingly recognised as an important driver of migration, no clear consensus on the impact of temperature and precipitation changes on migration has been reached due to great heterogeneity in past results (Black et al., 2011; Helbling et al., 2023). As a result, this research attempts to contribute to the academic field by providing new and distinctive evidence on the relationship between climate change and migration.

To this end, this research deploys a large dataset of 529 African regions observed from 2000 through 2020 to investigate the extent to which temperature and precipitation changes are migratory drivers. The results of a panel regression analysis including region- and time-fixed effects model and control variables on other migratory drivers indicate that indeed a changing climate is related to migration. Namely, an increase in temperature is related to increased migration, while precipitation change is related to migration through a non-linear quadratic relationship, in which both increases and decreases in precipitation are predominantly related to increased migration. An investigation into possible heterogeneity between regions finds that a changing climate is particularly related to migration for relatively unwealthy and agriculture-dependent regions. This suggests agriculture to be a channel through which climate change affects migration and wealth to be a moderator of the relationship.. Next to that, while for relatively dry regions both decreases as well as increases in precipitation are related to increased migration, for relatively wet regions only increases in precipitation are related to increased migration. This finding suggests that populations in relatively dry regions are more vulnerable to precipitation decreases, while equally vulnerable to precipitation increases.

The structure of the research is as follows: The antecedent introduction has set out the research problem and its relevance, and provided a summary of the research; the theoretical framework situates the research problem in the academic field of research and give an overview of past results to be able to formulate hypotheses; the methodology will describe the data and methodological approach to obtain results to allow testing the hypotheses and answering the research question; the results will report the acquired outcomes of the methodological approach and reflect on the hypotheses; the conclusions and discussion will summarize preceding chapters, answer the research question, reflect on limitations of the research, formulate policy implications of the results, and give suggestions for the way forward for future research.

2. Theoretical framework

2.1 Migration

Migration is the act of moving from one place of residence to settle in another either temporarily or permanently (McAuliffe & Triandafyllidou, 2022). This research explores whether migration is driven by climate change. First of all, however, it is important to understand why migration happens and why people would want to migrate in the first place.

At the macro-level, structural spatial disparities in livelihood conditions shape the right circumstances in which migration is more likely. At the meso- and micro-level, people can be driven to migrate by anything that affects the perception of the conditions and prospects of their livelihoods at the place of residence compared to those of potential destinations (Czaika & Reinprecht, 2022). Several factors have been identified to be shaping spatial disparities and people's perception and therefore driving migration. The main contextual conditions that either play a direct or indirect role in driving migration are hereafter divided into macro-level and micro- and meso-level conditions (Black et al., 2011; Hunter et al., 2015; Czaika & Reinprecht, 2022):

- Macro-level drivers are created by spatial disparities in contextual conditions. Examples of spatial disparities are: economic conditions related to employment, income, and producer- and consumer prices; political conditions related to freedom, conflict, security, discrimination, and persecution; social conditions related to education, family obligations, gender norms, and well-being; demographic conditions related to health, the prevalence of diseases, and the size and structure of the population; and environmental conditions related to the exposure to hazards, land productivity, habitability, and food, energy and water security.
- Meso-level conditions include the costs of moving, social networks, legal framework, and migrant organisations matter.
- Micro-level drivers refer to the individual's characteristics, aspirations, and preferences that affect the decision to migrate or not.

The described macro-, meso-, and micro-level conditions are the backdrop to the decision-making process of potential migrants. Early theorisations on the decision for people to migrate or stay are derived from economics, in which migration can be described by a cost-benefit model or a push-pull model (Sjaastad, 1962; Lee, 1966). Namely, the neoclassical migration theory argues that people tend to move if the expected returns to migration are positive. As a result, the relative lack of economic opportunities and disappointing livelihood conditions in combination with the prospect of non-improvement should push people to migrate from their place of residence. Similarly, the availability of economic opportunities or better livelihood conditions at a potential destination should pull these people towards it. At the same time, however, it could also be economically rational to not move due to the costs of migrating and place-dependent losses in knowledge, skills, or economic- and social networks (Sjaastad, 1962; Fischer et al., 1997; Bircan et al., 2020). Migration between and within countries, for instance from rural to urban areas, should continue until differences in economic- and livelihood conditions cease to exist (Harris & Todaro, 1968; Bircan et al., 2020).

While models based on neoclassical migration theories explain migration flows relatively well, they fail to explain why the majority of people choose to not migrate (Hagen-Zanker, 2008). Billions of people should have arguably valid incentives to migrate due to large disparities in economic, political, or social opportunities across and even within countries. However, a recent poll showed that only around 750 million adults would like to migrate if they had the possibility, which corresponds to one in eight adults (Esipova et al., 2018). Therefore many others do not consider migrating, even though they would have good reasons for moving elsewhere.

The obvious reason for immobility is potential migrants are not simply pushed and pulled between places to achieve the highest expected returns (Czaika & Reinprecht, 2022; Bircan et al., 2020). In contrast to what complete rationality and perfect access to information predicts, migration comes with a lot of uncertainty and risk as information and expectations on migratory options are always incomplete. As a result, people tend to decide on whether and where to move based on their own but sometimes biased agency and self-determination, and voluntary non-migration despite dire livelihood conditions can be a viable outcome (Bakewell, 2010; Carling & Schewel, 2018; Czaika & Reinprecht, 2022a). Therefore, for people to actively seek to move elsewhere, it is decisive whether potential migrants not only can imagine but also aspire a better life (Appadurai, 2004; Carling & Talleraas, 2016). Often this is not the case because people compare their own well-being and livelihood conditions to others around them instead of to those in other areas or countries (Stark & Taylor, 1991). The aspiration to migrate is therefore argued to be one of the two decisive factors in people becoming migrants or not, according to the aspiration/ability model on migration (Carling & Schewel, 2018).

However, besides individual conditions, many other factors are recently widely acknowledged to influence the decision-making process of potential migrants (Hunter et al., 2015; Hoffmann et al., 2021). For instance, the network theory states that the decision to migrate is an outcome based on the complex social network of a potential migrant combined with contextual political, social, demographic, and environmental conditions (Czaika & Reinprecht, 2022, 2022a). This social network consists of family and friends but also migrant organisations and past migrants, which provides social capital such as information and financial- and practical support necessary for the decision-making process (Boyd, 1989; Ritchey, 1976). As a result, the decision to migrate or stay is complex and evolving over time and is often made not by individuals but within family- or household structures, which is not captured by the neoclassical migration theories (Gubhaju & De Jong, 2009; Bircan et al., 2020; Czaika & Reinprecht, 2022).

Moreover, the decision-making process is still not the final step to migrating. Here comes the non-surprising second decisive factor of the aspiration/ability model into play: the ability to migrate after the decision is made that migration is preferable to non-migration, either by choice or coercion (Carling & Schewel, 2018). Namely, for migration aspirations to be translated into actual migration, potential migrants require financial and social capital, which is often missing. Consequently, only a quarter of all adults that expressed the willingness to migrate permanently actually did between 2011 and 2015 (Esipova et al., 2011; OECD, 2015). So despite, for example, dire political, social, economic, and demographic conditions driving migration aspirations, many are involuntary immobile because of the lack of ability to migrate. This is also in line with the migration hump theory, which predicts that migration from low-income countries increases with socioeconomic development, as it enables more individuals to match their migration aspirations with the ability to migrate (Taylor, 1996; Carling & Schewel, 2018; Hoffmann et al., 2020; Yavçan, 2021).

2.2 Migration and climate

As mentioned, several fundamental drivers of migration such as economic, political, social, and demographic factors have been identified and studied extensively (Black et al., 2011; Hunter et al., 2015). As the effects of climate change become gradually more visible, environmental factors as drivers of migration have been increasingly considered as well (Government Office for Science, 2011; Hunter et al., 2015; Czaika & Reinprecht, 2022a). In particular, the academic debate intensified after the Fourth Assessment Report of the IPCC (2007) included that climate distress has the potential for population migration (Backhaus et al., 2015). However, the surge in the perceived importance of environmental factors in driving migration wrongly suggests that it is a recent phenomenon. Migration is as old as humanity and has been used throughout history by societies as a coping mechanism for climate variability (Massey et al., 1999; Adger et al., 2003; McLeman & Smit, 2006).

Namely, gradual climate change could make habitation in certain places less desirable or even impossible, and people may respond by moving elsewhere (Bohra-Mishra et al., 2013; Hunter et al., 2015). The effect of climate change on livelihood conditions, and therefore on migration aspirations, is complex and manifold. All described drivers of migration are affected by changing temperatures and precipitation in one way or another. Namely, gradual changes in the climate might affect migration through its influence on water- and, through its effect on agricultural productivity and food security (Black et al., 2011; Khavarian-Garmsir et al., 2019; Migali & Natale, 2021). Consequently, economic conditions may alter as well, due to changing incomes and employment in agriculture and changing producer- and consumer prices of food and water. Furthermore, a changing climate might increase health-related risks, such as the prevalence of certain deceases, while it also is linked with increases in the risk of conflict, violence, and corruption (Marchiori et al., 2012; Abel et al., 2019; Khavarian-Garmsir et al., 2019). As a result, a gradually changing climate might induce migration due to the complex and manifold interplay between changing economic, political, social, and demographic conditions and its effect on people's livelihoods and perceptions.

In general, as a response to a changing climate people can stay in place and do nothing, stay in place and mitigate arising problems, or migrate from their habitat. The choice between these options depends on the extent of the problems and mitigation capabilities (Reuveny, 2007). Economic theory would suggest that it could also be economically rational to not move due to place-dependent losses in knowledge, skills, or economic networks (Fischer et al., 1997). Moreover, cultural place attachment, uncertainty, and risks surrounding migration might lead to persisting populations despite severe and undesirable environmental conditions (Adams & Adger, 2013; Czaika & Reinprecht, 2022a). Because of the default option of staying, people are found to only adjust the perception of their living conditions when environmental changes are not only severe and undesirable but also are expected to be persistent (Hsiang, 2016; Falco et al., 2019). For that reason, incidental extreme weather events are suggested to drive migration less than climate change effects that materialize over a longer period of time (Perch-Nielsen et al., 2008; Afifi, 2011; Bohra-Mishra et al., 2013; Beine & Parsons, 2015). Despite location-specific advantages, cultural place attachment, and uncertainty and risks of migration, limits in capabilities to adapt to a changing climate could still result in people choosing to move elsewhere (Adger et al., 2003). Therefore, the question would be whether the recent climate change has been detrimental enough to livelihoods to overcome the preference for staying and induce widespread migration.

2.3 Direct- and indirect effects

Historically, there has been a paucity of empirical research on the effects of climate change on migration, partly due to sparse data (Henry et al., 2004a; Bohra-Mishra et al., 2013; Czaika & Reinprecht, 2022a). However, as climate change is becoming more and more visible, a steadily growing body of quantitative empirical research has investigated the effects of gradual climate change - in particular temperature and precipitation change - on migration in low- and middle-income countries (Hoffmann et al., 2021). As a result, gradual climate change is increasingly recognised as an important driver of migration (Berlemann & Steinhardt, 2017). Nonetheless, no clear consensus on the adverse impact of temperature and precipitation changes on migration has been reached, as there is great heterogeneity in past results (Black et al., 2011). Evidence presented by previous research is patchy and varied, as some studies find that temperature and precipitation changes influence migration directly, while others do not, or only have indirect effects (Helbling et al., 2023). This is not surprising as different studies apply different methodologies in how they estimate climate change, investigate different types of migration, or consider only one country while others consider many (Piquet, 2010; Black et al., 2011; Hoffmann et al., 2020; Helbling et al., 2023).

Moreover, the academic field of migration studies is currently dominated by international migration rather than internal migration, even though, more people move internally than internationally (King, 2012; UNDP, 2009; IOM, 2020). This also applies to past research on climate-induced migration, where the vast majority solely considers international migration while few other studies only estimate internal migration by urbanisation at the national level as a proxy (Cattaneo & Peri, 2016; Henderson et al., 2017; Hoffmann et al., 2020, 2021). Although evidence for internal migration due to temperature and precipitation changes is suggested to be more robust than for international migration, the limited measurement of internal migration makes it likely that results are biased (Hoffmann et al., 2021). Namely, using urbanization at the national level as a proxy overlooks other forms of internal migration such as rural-to-rural, urban-to-rural, and urban-to-urban.

Changing temperature- and precipitation levels can either directly influence migration decisions, for instance by posing an immediate threat to health or well-being, or indirectly through the interplay with other migration drivers such as economic, political, and social conditions (Hoffmann et al., 2020; Helbling et al., 2023). Previously found direct- and indirect effects of temperature and precipitation changes on migration will be set out hereafter, and testable hypotheses will be formulated based on this. Due to the lack of consensus, hypotheses are formulated on the basis of where the majority of evidence points to. As temperature and precipitation changes are suggested to have different effects on migration, past results will be described separately.

A multitude of studies found higher temperatures to be directly related to increased migration across and within country borders, or from rural to urban areas, as migration within country borders is usually measured with urbanization as a proxy due to a lack of data (Marchiori et al., 2012; Mueller et al., 2014; Bohra-Mishra et al., 2014; Coniglio & Pesce, 2015; Backhaus et al., 2015; Cattaneo & Peri, 2016; Mastrorillo et al., 2016; Peri & Sasahara, 2019; Hoffmann et al., 2021). In contrast, Beine & Parsons (2015), Cai et al. (2016), and Nawrotzki et al. (2015) find no direct effect of temperature and changes on migration, but only through the detrimental effect on agricultural productivity. This is in line with other studies that found agricultural productivity to be the channel through which temperature changes affect migration (Marchiori et al., 2012; Mastrorillo et al., 2016; Maurel & Tuccio, 2015; Falco et al., 2019). But then again, Marchiori et al. (2017) found no significant relationship between temperature-induced decreases in agricultural productivity and international migration. Overall, however, the majority of past research suggests that temperature increases are related to increased migration. Therefore, the first hypothesis is as follows:

H1: Regional temperature increases are related to increased migration from that region.

Similarly to temperature, evidence of the effect on migration of changing precipitation levels is ambiguous as well. This is partly caused by varying measures of precipitation. In the simplest measure, precipitation levels are averaged over certain periods of time, while in another measure, the year-to-year variation in the level of precipitation is used (Berlemann & Steinhardt, 2017). Using the simpler measure, several studies found that a decline in the level of precipitation is directly related to increased migration, international as well as internal (Henry et al., 2004b; Grey & Mueller, 2012; Afifi et al., 2014; Dallmann & Millock, 2017; Nawrotzki & DeWaard, 2018). Similarly, Munshi (2003, Gray (2009), and Beine and Parsons (2015) suggest that decreased precipitation leads to increased migration, but again only through its adverse effect on agricultural productivity. In contrast, an increase in the level of precipitation is suggested to be related to increased migration as well (Hunter et al., 2014; Nawrotzki et al., 2015; Gray & Wise, 2016). Illustrating the lack of consensus once more, Mueller et al. (2014) and Bohra-Mishra et al. (2017) reported no robust effect of precipitation-level changes on migration, while Nawrotzki et al., (2015) found no evidence for agricultural productivity to be the channel through which precipitation affects migration.

Not surprisingly, meta-analyses bring forward that changes in the level of precipitation tend to have a negligible direct impact on migration. Across and even within studies, the direction of the estimated effect of changes in the level of precipitation differs substantially (Hoffmann et al., 2021; Bertoli et al., 2022). Changes in the level of precipitation are suggested to have less strong effects on migration than temperature changes because across regions both increased and decreased precipitation can increase migration, depending on the context. In other words, not specifically increases or decreases in precipitation are suggested to be related to increased migration, but *changes* in precipitation are. As a result, aggregating data from these different contexts would lead to a weakening of the average estimated effect. This is less of an issue for changes in temperature because increases in temperature are more clearly linked with negative impacts on livelihoods (Hoffmann et al., 2021).

To circumvent this methodological problem, many studies use the other described measurement of precipitation: year-to-year variation in the level of precipitation, which is independent of the direction of the change in the level of precipitation. In doing so, the vast majority of studies find that increased precipitation variability is related to increased migration, within as well as between countries (Muelle & Osgood, 2009; Marchiori et al., 2012; Bohra-Mishra et al., 2014; Coniglio & Pesce, 2015; Maurel & Tuccio, 2016; Mastrorillo et al., 2016; Hoffmann et al., 2022). Although found significant, the effects of precipitation on migration are often smaller in magnitude than those of temperature effects. Similarly, the effects of precipitation changes are found to be weaker when temperature changes are controlled for, while effects of temperature changes are estimated stronger when precipitation changes are controlled for (Bohra-Mishra et al., 2014; Backhaus et al., 2015; Hoffmann et al., 2020; Hoffmann et al., 2021). Nevertheless, from described results, it is derived that year-to-year variation precipitation is more clearly related to increased migration than an increase in precipitation is. As a result, a change in precipitation is expected to be linked to migration through a quadratic non-linear relationship, in which both decreases and increases in precipitation lead to increased migration. Therefore, the second hypothesis will state that not simply an *increase* in precipitation is related to migration, but a *change* in precipitation is related to migration in the following manner:

H2: Regional precipitation changes are non-linearly related to increased migration from that region

In sum, the majority of quantitative studies find that temperature increases and precipitation changes positively affect migration, either directly or indirectly. Despite the small amount of research that considers internal migration combined with its limited measurement, migration due to temperature and precipitation changes seems to be often regional, while results on the effects on international migration are more ambiguous (Hoffmann et al., 2021; Bertoli et al., 2022). However, because of contextual variations between these studies, it should be emphasized that heterogeneity between these contexts could make these results context-dependent (Hoffmann et al., 2020; Yavçan, 2021).

2.4 Heterogeneity

As stated above, migration responses to temperature and precipitation change are suggested to potentially be context-specific, due to the manifold pathways through which temperature and precipitation changes together with other factors could affect migration (Hoffmann et al., 2020; Helbling et al., 2023; Bertoli et al., 2022). Because of this, it is also explored whether the relationship of a changing climate with migration differs between different regions.

So, changing temperature and precipitation are generally found to be affecting migration flows. However, some studies report stronger effects for middle-income countries than for low-income countries, bringing forward that financial constraints can immobilize people despite adverse effects of temperature and precipitation changes (Cattaneo & Peri, 2016; Peri & Sasahara, 2019; Bazzi, 2017). This could especially be the case whenever the impacts of changing temperatures and precipitation result in a reduction of the very capital required to enable a move (Government Office for Science, 2011; Beine & Parson, 2015). Similar evidence has been found for migration within country borders, where the poorest and most vulnerable regions are characterized by climate-related immobility after temperature and precipitation changes (Henry et al., 2004b; Adger et al., 2014; Nawrotzki & DeWaard, 2018; Biella et al., 2022). This is again also in line with the migration hump theory and the aspiration/ability model, which predicts that migration from low-income countries increases with socioeconomic development, as it enables more individuals to match their migration aspirations with the ability to migrate (Taylor, 1996; Carling & Schewel, 2018; Hoffmann et al., 2020; Yavçan, 2021). Therefore, the relationship of changing temperatures and precipitation with migration may differ between regions, depending on the relative wealth of those regions. To investigate whether this heterogeneity between regions exists, the following hypotheses are formulated:

H3: The effect of regional temperature increase on migration is stronger for relatively wealthy regions

H4: The effect of regional precipitation change on migration is stronger for relatively wealthy regions

Heterogeneity between regions might also exist in terms of how exposed or vulnerable the population is to changing temperatures and precipitation. Namely, temperature and precipitation changes particularly induce migration from countries where agriculture represents the main source of income for the population (Falco et al., 2019; Mastrorillo et al., 2016). Agricultural work, especially in low- and middle-income countries, is affected by a changing climate more than work that takes place indoors due to exposure to and dependency of people and land on weather conditions (Falco et al., 2019). This is supported by results that have found that temperature and precipitation changes affect migration particularly in agriculture-dependent countries (Maurel & Tuccio, 2016; Cai et al., 2016). Again, similar results have been found for migration within country borders, as changing temperatures and precipitation induce migration especially from rural regions (Nawrotzki et al., 2015; Bohra-Mishra et al., 2014; Biella et al., 2022). These results suggest that agriculture is a channel through which temperature increases and precipitation changes affect migration. To be able to suggest that this mechanism exists, possible heterogeneity between regions is investigated by means of the following hypotheses:

H5: The effect of regional temperature increase on migration is stronger for relatively agriculture-dependent regions

H6: The effect of regional precipitation change on migration is stronger for relatively agriculture-dependent regions

There might also exist heterogeneity in the relationship of temperature and precipitation with migration between different initial climates. A rise in average temperature might be particularly detrimental in already hot climates (Beine & Parsons, 2015). Results found by Bohra-Mishra et al. (2014) support this, as they found that an increase in temperature is related to migration more strongly at initial higher levels of temperature. As a result, temperature increases may have stronger effects on migration in relatively hot climates than in relatively cold climates. Similarly, as mentioned, the effect of precipitation on migration may also depend on the initial level of precipitation. Namely, in conditions that are relatively dry, a further decline in precipitation is likely to increase migration, whereas, in relatively wet conditions, migration is likely to increase with a further increase in precipitation (Bohra-Mishra et al., 2014). Whether the initial temperature and level of precipitation - and thereby heterogeneity between regions - affect the relationship of temperature increases and precipitation changes with migration, is explored by means of the following hypotheses:

H7: The effect of temperature increase on migration is stronger for relatively hot regions

H8: In relatively dry regions, only regional precipitation decreases are related to increased migration (negative effect)

H9: In relatively wet regions, only regional precipitation increases are related to increased migration (positive effect)

3. Methodology

3.1 Data and variables

To test the hypotheses through empirical analyses, a novel panel dataset is assembled, which covers 53 African countries and 529 sub-national regions from 2000 to 2020. The sub-national regions are defined following the regional distributions used by Smits (2016) and do not necessarily follow administrative units of countries. The African countries and sub-national regions included in the sample are graphically mapped in appendix 8.2.

The focus is on Africa for several reasons. First, due to limited resources, data on population distribution, temperature and precipitation could only be computed for one continent. The choice for Africa was made because of its characteristics that make the continent particularly relevant for this research. The inhabitants of Africa are, on average, the poorest people on the planet, with more than 40 percent of the population living in poverty (Beegle et al., 2016). Combined with heavy reliance on agricultural production for many countries, the prevalence of poverty makes Africa expected to be highly vulnerable to climate change as even small changes in weather conditions can significantly impact human well-being (Marchiori et al., 2012; Global Data Lab, 2023d). Furthermore, given the scenarios set by the IPCC (2022) that predict temperature and precipitation patterns to change even further, Africa is among the regions that face the largest number of ecological threats. Therefore, a relative lack of resilience will lead to worsening food and water insecurity, competition over resources, and civil unrest, which in turn might lead to increased displacement and migration in many African countries (IEP, 2022). This predicted vulnerability is also translated into the share of people that want to migrate from a specific region. Nowhere it is greater than West Africa, Sub-Saharan Africa, and North Africa (OECD, 2015; Carling & Talleraas, 2016). As a result, to be able to increase resilience among the most vulnerable, it is particularly valuable to identify the role that changing climate conditions play in determining migration in Africa.

The complete dataset is compiled by merging different sources of data, the Global Data Lab, the World Bank, and NASA. A full overview of the dependent, independent, and control variables, their measurement, and their source can be found in appendix 7.2, but will also described in detail hereafter.

3.1.1 Dependent variable

To begin with, because sub-national data on migration does not exist, data on population count per sub-national region is derived from gridded spatial population data of NASA’s The Gridded Population of the World, Version 4 (GPWv4, 2018), in which a proportional allocation gridding algorithm has been used to estimate population per pixel (CEISIN, 2018). The spatial population data is consistent with national censuses and population registers, but also adjusted to match the United Nations World Population Prospects country totals. For use in this research, the gridded population values are aggregated at the sub-national level for all available years: 2000, 2005, 2010, 2015, and 2020. Now, to estimate migration, the 5-year percentage change in the population living in a sub-national region is calculated for each of the intervals of years, with the year 2000 as a baseline. This is calculated as a percentage change to correct for the initial size of the population of the region in question. Next to that, all values are multiplied by minus 1, so that the variable reflects emigration instead of immigration. Although migration is measured at the regional level, note that this migration can be within country borders as well as across country borders, as the destination of migrants is unknown.

However, other factors can affect population changes over time as well, and those need to be accounted for. Namely, the interplay between fertility and mortality rates affects population counts over time as well. Therefore, to be able to attribute changes in the population counts to migration, fertility and mortality rates will be included as control variables. As sub-national data on fertility rates are only available for a very limited range of countries and years of the sample and sub-national mortality rates are not available at all, country-level data on national fertility and mortality rates are used for both countries and sub-national regions within that country. By doing so, this research adopts the assumption made by Alessandrini et al. (2020) of similarity in fertility and mortality rates between the country- and regional level in the process of estimating migration. Any resulting under- or overestimation of fertility and mortality for specific region-years would create a bias in the measurement of migration, as observed migration would partly be driven by positive or negative population growth.

The correlation between fertility rates at the country- and regional level for the limited sample on which sub-national data is available would give an indication of the reasonability to assume similarity between the two. As shown in appendix 7.3, observations on fertility rates in common sub-national regions and years are correlated with a value of 0.96, which suggests an almost perfect linear relationship between the two. Although this correlation covers only a part of the whole sample, it shows that for this research it is not unreasonable to assume similarity between country- and regional level fertility rates. A similar correlation test is not possible for mortality rates, as data on sub-national mortality rates are not available at all. Therefore using country-level mortality rates for sub-national regions within that country might be a too strong assumption.

3.1.2 Independent variables

For the independent variables on climate necessary for hypotheses 1 and 2, data on average temperature and precipitation is collected from NASA’s Modern Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). MERRA-2 data is gridded reanalysis satellite data, which in this case estimates for each grid cell the monthly average air temperature at two meters above the surface and daily precipitation in millimetres. Using MERRA-2 climate data offers several advantages over alternatives often used in similar research for estimating temperature and precipitation. Namely, temperatures and precipitation changes are often based on weather station measurements, which has some major disadvantages (see Marchiori et al., 2012; Nawrotzki et al., 2015; Cattaneo & Peri, 2016; Falco et al., 2019). Weather stations come in and out of existence, and are far rarer in lower-income and sparsely populated regions. As a result, data is often unbalanced and not continuous (Auffhammer et al., 2013). Gridded weather station data, which are based on interpolations between weather stations, are used as an alternative to be able to create a balanced and continuous panel. Nevertheless, these data still suffer from the unequal distribution of weather stations as weather variables can vary significantly in magnitude across space, especially precipitation (Hoffmann et al., 2021; Dell et al., 2014). Next to that, the entry and exit of weather stations create artificial variations in measurements, which can lead to measurement errors, and thus biased results (Auffhammer et al., 2013). Another challenge concerns areas where there are more grid cells than underlying stations, which is particularly a problem for areas with sparse coverage. In that case, data can only be computed at large enough units, for instance on country-level.

As a result, satellite data offers significantly higher resolution and accuracy across the board in certain areas of sparse coverage of weather stations, especially if data are only required after 2000 (Dell et al., 2014). However, despite that satellite data is more accurate across the board, it is less accurate than observations of a weather station relatively close to that weather station. By combining information from satellites and ground weather stations, gridded reanalysis data corrects for this. The relatively low-income and rural profile of Africa, the focus of this research, therefore, makes MERRA-2 more suitable and reliable than weather stations for estimating temperature and precipitation. As a result, measurements of temperature and precipitation are standardized and continuous, which results in a balanced panel across time and space.

Again, for use in this research, this data is aggregated yearly at the sub-national level, where temperature is measured as a spatially weighted annual average in Celsius and precipitation is measured by the spatially weighted annual total millimetres in that sub-national region. This research aims to explore the effect of gradual climate change on migration, so year-to-year fluctuations in temperature and precipitation can distort the investigation of the longer-term effects of a changing climate on migration. Moreover, changing temperatures and precipitation might affect migration only over longer timespans when accumulated to greater changes or because of the initial persistence in staying of populations. To address these issues and to harmonize this section of the dataset with the one on population, the average temperature and precipitation over the antecedent 5 years are calculated for the years 2000, 2005, 2010, 2015, and 2020. Thereafter, to reflect gradual climate change, the percentage change in the 5-year average of temperature and precipitation is taken for each of the intervals of years, with the year 2000 as a baseline. This is calculated as a percentage change as opposed to an absolute change to standardize the values of the variables, which is particularly necessary for precipitation. Otherwise, the effect of a millimetre increase in precipitation annually would be uninterpretable small. To clarify how the percentage change in the 5-year average of temperature and precipitation is taken for each of the intervals of years, an example is shown below. For instance, to obtain the observation for the percentage change in temperature for 2005, the following computation is made:

$$\Delta TEMP(2005) = \frac{AVG(2001 \ldots 2005) - AVG(1996 \ldots 2000)}{AVG(1996 \ldots 2000)} * 100$$

For exploring possible heterogeneity between regions, several variables and modifications to the total sample are introduced. First, to test hypotheses 3 and 4 - whether temperature increases and precipitation changes are stronger related to increased migration from relatively wealthy regions - the sub-national log gross income per capita in 1000 US Dollars (2011 PPP) of that region is used. Adopting the approach of Cattaneo and Peri (2016) and Falco et al. (2019), the total sample is split into two based on the average income over the ten years preceding the year 2000. The regions with a log GNI per capita below the median value will be categorized as relatively unwealthy regions, whereas regions with a value above the median value will be categorized as relatively wealthy regions. Note that the categorization on the basis of wealth is relative to the African continent, and income in these regions is generally speaking still low compared to most parts of the rest of the world.

Secondly, a measure of agricultural dependence is necessary to test hypotheses 5 and 6 - whether the effects of temperature increases and precipitation changes on migration are stronger in agriculture-dependent regions. To this end, the percentage of the population employed in agriculture in that region will be used. As sub-national data on employment in agriculture is only available for a very limited range of regions and years, national data on the percentage of the population employed in agriculture is used for regions within that country. Subsequently, following the methodological approach of Cai et al. (2016), the total sample will be split based on the average population employed in agriculture over the decade preceding the year 2000. Regions with a value above the median will be considered relative agriculture-dependent regions, whereas regions with a value below the median will be considered relative agriculture-independent regions.

Similar to fertility rates, the correlation between the share of the population employed in agriculture at the country- and regional level for the limited sample of which sub-national data is available gives an indication of whether it is valid to assume similarity between the two. As shown in appendix 7.3, the country- and regional level data is correlated with a value of 0.77, which is fairly high. Although this is not as high as the correlation of fertility rates and also only covers a part of the total sample, the vast majority of regions will be labelled as either relatively agriculture-dependent or agriculture-independent correctly. Therefore, the assumption of similarity between the national and sub-national data is adopted.

Thirdly, to test hypothesis 7, the sample is again split into two, this time on the basis of average temperature between 1990 and 2000. Regions with a value for the level of temperature below the median value will be categorized as relatively cool regions, while regions with a value above the median value will be categorized as relatively hot regions. For hypotheses 8 and 9, the same process is repeated, but now with the average level of precipitation between 1990 and 2000. Regions with a value for the level of precipitation below the median value will be labelled as relatively dry regions, whereas regions with a value above the median value will be labelled as relatively wet regions.

3.1.3 Control variables

Several control variables will be included in the analysis to enhance the internal validity of the study by limiting the influence of confounding and other extraneous variables. As mentioned, fertility and mortality rates are included to isolate migratory effects in changing population counts. Next to that, temperature and precipitation changes are always both included in the model to make the effects of the variables more representative as every region experiences both simultaneously as well.

Usually, when investigating a certain migratory driver, migration studies also include control variables on other drivers in their analysis. However, as this research explores the extent to which a changing climate is a migratory driver, including non-climatic controls should be done with caution, as these could induce a bias in the estimation due to an over-controlling problem (Hsiang, 2016, Falco et al., 2019). The reason for this is that most would-be relevant controls on the most important other migratory drivers, such as income, health, conflict, and political regime, may be endogenous and affected by climate variables themselves, which would make them "bad controls" (Black et al., 2011; Angrist & Pischke, 2009). Including them anyway could have the effect of partially eliminating the explanatory power of the climate variables, even if a changing climate is the underlying fundamental cause of a change in the dependent variable (Dell et al., 2014). On the other hand, excluding controls entirely from the analysis might lead to incomplete conclusions about the relationship of temperature and precipitation with migration because it could depend on the simultaneous influence of the control variables. Therefore, this research follows the recommendations of Hoffmann et al. (2021) and Falco et al. (2019) for future studies to first employ a parsimonious model that quantifies climate change impacts on migration without over-controlling for other migratory drivers. Thereafter, controls will be included to better isolate the effects of the climate variables on migration from other factors that might influence the relationship. This way, the total effect of climate change on migration is assessed more accurately.

To this end, the control variables that are used represent the most important other migratory drivers according to theory and past studies (Black et al., 2011; Falco et al., 2019; Czaika & Reinprecht, 2022). The first control variable is actually already introduced, namely, to estimate economic conditions, the sub-national log gross income per capita in 1000 US Dollars (2011 PPP) of that region is used. Income disparities between regions and countries have always been an important motivator for migration, and therefore it is important to control for income when exploring the effects of a changing climate on migration. The second control variable concerns health, which is measured by the sub-national life expectancy at birth in that region. Undesirable health standards have been found to leave people with the desire to migrate, which makes controlling for it important. The third control variable measures the extent to which conflict and violence are absent in that region. As everyone may can imagine and also supported by empirical studies, conflicts and wars are indeed primary drivers of migration (Brzoska & Fröhlich, 2016; Adhikari, 2019). Because sub-national data is not available in a usable manner for conflicts and violence, standardized national data of the World Bank on the variable ‘political stability and the absence of violence/terrorism’ is used for regions within that country. Due to how the variable is defined, a lower score indicates relatively more violence and conflict within that region. The fourth and final control variable included in this research concerns the political conditions in that region. This is estimated by the extent to which the population experiences freedom and lives under a democratically chosen and accountable government. As, again, sub-national data is not available on this, standardized national data of the World Bank on the variable ‘voice and accountability’ is used for regions within that country. This should be less of an issue for this variable compared to the other variables for which country-level data is used as all regions within a country usually live under the same national government and therefore also experience somewhat the same extent of freedom.

All in all, the novel dataset represents significant progress over datasets used in similar studies because of its sub-national character and its improved measurement of temperature and precipitation. However, the dataset is not without its limitations. Particularly, the dataset can be improved where country-level data is used for regions within that country to approximate fertility rates, mortality rates, agricultural dependence, conflict, and political regimes in the absence of good-quality sub-national data.

3.2 Empirical model

To explore the relationship of regional temperature and precipitation change with migration, a panel regression analysis will be performed. Before determining the specifications of the model for this panel regression analysis, the panel is first checked against the methodological assumptions to strengthen the internal validity of the empirical model.

3.2.1 Checking assumptions

As shown and explained in 7.4, the methodological assumptions of no autocorrelation-, homoscedasticity-, independence-, and normality in error terms are violated. To correct these violations, cluster-robust standard errors are included in the panel regression model. That clustered standard errors are needed can also be reasoned intuitively. Namely, including clustered standard errors is appropriate when observations on the dependent, independent, and control variables are not independently distributed across regions. This is clearly the case, as adjacent regions often have a somewhat similar climate, while regions within the same country are more likely to have relatively similar values for income and life expectancy. Moreover, as country-level data on fertility rates, mortality rates, agricultural dependence, the absence of conflict, and voice and accountability is used for regions within that country, clustered standard errors help to mitigate problems arising from the assumption of similarity in national- and sub-national data on these variables.

The assumption of no endogeneity usually cannot be respected with full confidence in the sphere of social science, which limits the extent to which a causal relationship between the independent and dependent variables can be established. This is, in general, a stubborn and hard-to-solve problem because explanatory variables are often not truly exogenous to the response variable. Luckily respecting the assumption of no endogeneity is less of an issue for this particular research. Namely, as changes in temperature and precipitation are driven by random geophysical developments, there exists no reverse causality between temperature and precipitation changes and migration. As a result, this research is able to establish a causal relationship between a changing climate and migration, if such exists.

The final assumption to be respected is the assumption of no multicollinearity. This is violated whenever predictor variables included in the same model are highly correlated with each other. When this would be ignored, determining the individual effect of these variables is made difficult as results will be less reliable. To check whether the predictor variables are not highly correlated with each other, a correlation matrix between the dependent, independent, and control variables is calculated, which is displayed in appendix 7.3. Variables with a correlation value above 0.8 or minus 0.8 are generally considered highly correlated. Because of this, the correlation between mortality and life expectancy stands out with a value of -0.876. This is a fairly logical correlation because a decline in mortality rates is linked, by definition, to an increase in life expectancy, as fewer people dying means more people getting older (Roser et al., 2013). To prevent biased results, one of these two variables needs to be excluded from the model. Because mortality rates are essential in estimating migration, the variable on life expectancy is dismissed. As none of the other variables that are simultaneously included in the model shows a high correlation with each other, the assumption of no multicollinearity is respected by only excluding life expectancy from the model.

Every empirical analysis suffers from the omitted variable bias. This bias occurs if one or more relevant variables are left out of the model, which leads to the possibility that the model attributes the effect of the missing variables to those that are included. A key advantage of using a panel regression model in the analysis is that can account for unobserved entity-level differences through the inclusion of region-fixed effects. Region-fixed effects concern region-specific and time-invariant components that are unknown and therefore cannot be controlled for otherwise. That including region-fixed effects is preferable for this analysis as opposed to just a simple pooling of the observations is also statistically shown and explained in appendix 7.5. An additional statistical instrument necessary to minimize the omitted variable bias, is to include time-fixed effects in the model. Using time-fixed effects excludes the possibility that a found relationship between a changing climate and migration is actually explained by events that happened simultaneously, such as changing global economic conditions, technological changes, or changing migratory laws in countries of origin and potential destinations. By including time-fixed effects, events that have happened simultaneously are controlled, and found effects of temperature and precipitation changes on migration are not actually explained by these factors.

3.2.2 Model specification

So, to generate results for testing the hypotheses, first, a panel regression model without control variables on other migratory drivers is run, including the dependent variable on migration controlled for fertility- and mortality rates, and the independent variables on the change in temperature and precipitation. Thereafter, the same model is run, but now also includes the control variables on economic conditions, health, conflict, and political regime. From this, the preferred model is chosen, on the basis of the influence that the control variables have on the effects of temperature and migration. The resulting functional form of the model is as follows, in which the quadratic variable on precipitation and the control variables are placed in half square brackets for now:

$$MIG_{it} = \beta_0 + \beta_1 TEMP_{it} + \beta_2 PRCP_{it} + [\beta_3 (PRCP_{it})^2] + \beta_4 FERT_{it} + \beta_5 MORT_{it} + [\beta_6 INC_{it} + \beta_7 VOICE_{it} + \beta_8 CONF_{it}] + \sum_{i=2}^{529} (\beta_{(i+7)} D_i) + year + u_{it}$$

In which:

- MIG_{it} is the dependent variable on migration;
- β_0 is the intercept;
- $\beta_1 TEMP_{it}$ is the parameter associated with the independent variable on temperature change;
- $\beta_2 PRCP_{it}$ is the parameter associated with the independent variable on precipitation change, and $\beta_3 (PRCP_{it})^2$ is associated with quadratic version of this variable;
- $\beta_4 FERT_{it}$ is the parameter associated with the control variable on fertility rates;
- $\beta_5 MORT_{it}$ is the parameter associated with the control variable on mortality rates;
- $\beta_6 INC_{it}$ is the parameter associated with the control variable on income;
- $\beta_7 VOICE_{it}$ is the parameter associated with the control variable on voice and accountability;

- $\beta_8 CONF_{it}$ is the parameter associated with the control variable on political stability and the absence of violence/terrorism;
- $\sum_{i=2}^{529}(\beta_{(i+7)}D_i)$ is the parameter associated with the dummy variable to allow region fixed effects in the model;
- $year$ is the parameter associated with time-fixed effects;
- and u_{it} is the region- and time-specific error term

To investigate that indeed not an *increase* but a *change* in precipitation is related to increased migration, not only the variable on the change in precipitation can be used. Namely, hypothesis 2 suggests a quadratic effect of changing precipitation on migration, where both positive and negative changes are related to increased migration. To approximate this non-linear effect, the change in precipitation is squared and added into an additional model alongside the regular change in precipitation. If indeed not an increase but a change in precipitation is related to increased migration, the model without the quadratic precipitation variable would show that an increase in precipitation is not related to increased migration, while the model with the quadratic precipitation variable would in fact suggest that both positive and negative increases in precipitation lead to increased migration. This is not replicated for the variable on the change in temperature, as hypothesis 1 simply suggests that an increase in temperature leads to increased migration. Therefore, it is expected that an increase in temperature is related to increased migration in both the model with and without the squared precipitation change.

Thereafter, the preferred empirical model will allow the investigation of heterogeneity between regions in how temperature and precipitation changes affect migration through the use of sub-samples. To test hypotheses 3 and 4, the preferred model with the quadratic variable on precipitation is run for both the sample including the relatively wealthy regions as well as for the sample including the relatively unwealthy regions. It is expected that the sample including the relatively wealthy regions will display a stronger effect of temperature and precipitation changes on migration, as populations within these regions would have a higher capability of migrating. To test whether differences in effects between the two samples are statistically significant, the approach of Rooks et al. (2016) is adopted, in which a third model is run for the total sample using a dummy variable and interaction effects. The dummy variable indicates whether an observation belongs to the relatively unwealthy or relatively wealthy sample of regions, and is multiplied with all the independent variables. The significance of this interaction term indicates whether the differences in effects between the two samples are significant.

To test hypotheses 5 and 6, a similar process is repeated but now the preferred model with the quadratic variable on precipitation is run for both the sample including the relatively agricultural independent regions as well as the relatively agricultural dependent regions. Following the hypotheses, it is expected that temperature increases and precipitation changes are stronger related to increased migration for the sample including the relatively agricultural dependent regions than for the sample including the relatively agriculture-independent regions. Again, the differences in effects between the two samples are tested for significance by a third model for the total sample with a dummy variable on agricultural dependence and interaction effects between this dummy variable and the independent variables.

For testing hypothesis 7, the preferred model with the quadratic variable on precipitation is now run for both the sample including relatively cool regions and the sample including relatively hot regions. If the hypothesis is correct, the effect of an increase in precipitation on migration would be stronger in relatively hot regions. Differences in effects between the two samples are again tested for significance through a third model including a dummy variable on whether an observation belongs to the relatively cool or relatively hot sample and interaction terms between this dummy variable and the independent variables.

Finally, to test hypotheses 8 and 9, the preferred model with the quadratic variable on precipitation is run once more for both the sample including relatively dry and the sample including the relatively wet regions. In contrast to temperature, the effect on migration is not expected to be stronger in relatively wet regions compared to relatively dry regions but is expected to change direction. In other words, where for the total sample it is expected that positive as well as negative increases in precipitation are related to increased migration, for the sample with relatively dry regions it is expected that only a decrease in precipitation is related to increased migration, whereas for the sample with relatively wet regions a further increase in precipitation is expected to increase migration. Again, differences in effects between the two samples are tested for significance by running a third model including a dummy variable on whether an observation belongs to the relatively dry or relatively wet sample and interaction terms between this dummy variable and the independent variables.

4. Results

4.1 Descriptive statistics

Before the results are computed, descriptive statistics are consulted to get a sense of the data on hand and the distribution of the variables. From table 1 on the descriptive statistics can be seen that the sample is strongly balanced with almost complete data for every variable on the 529 African sub-national regions for four time periods. The dependent variable on migration, which is measured by negative population change, has a mean value of -13.5. This means that the regions included in the sample actually have, on average, experienced a population growth of 13.5 percent. Because also the standard deviation is 13.4, the majority of regions in the sample have experienced positive population growth. This shows once again the importance of controlling for fertility and mortality rates to isolate migratory effects in changing population counts.

The descriptive statistics on the change in temperature and precipitation show that regions have experienced increases as well as decreases in both temperature and precipitation with a mean change relatively close to zero. Next to that, the statistics on the temperature and precipitation level indicate that cool and hot as well as dry and wet regions are represented in the sample. While the mean level of precipitation is similar to the global mean level of precipitation, the mean temperature level is quite high compared to the rest of the world (IPCC, 2021). This may result in context-specific results and could therefore affect the reproducibility of the upcoming results with a sample of regions with relatively lower levels of temperature.

Variable	**Obs.**	**Mean**	**Min**	**Max**	**Median**	**Std. Dev.**
Migration (negative population change)	2152	-13.5	-189	43.3	-12.5	13.4
Temperature change	2152	0.367	-5.72	8.2	0.333	1.91
Temperature level	2152	24.3	10.9	30.2	25	3.36
Precipitation change	2152	8.35	-71.2	204	3.27	36.1
Precipitation level	2152	1010	1.26	4205	942	654
Mortality rates	2152	306	87.6	730	294	104
Fertility rates	2152	4.79	1.36	7.62	4.9	1.24
Agricultural dependency	2152	50.2	4.85	90.3	49	20.4
Income	2128	7.94	6.3	11	7.79	0.86
Life expectancy	2128	59.3	40.7	77.1	59.1	7.05
Political stability and absence of violence/terrorism	2136	-0.748	-3.14	1.06	-0.591	0.9
Voice and accountability	2136	-0.668	-2.17	0.955	-0.699	0.675

Table 2: Descriptive statistics

The statistics on the mean-, minimum-, and maximum values combined with the standard deviations of the control variables and agricultural dependency show that the regions included in the sample have widely different contexts with respect to these variables. This supports the inclusion of variables to control for these different contexts, and also underwrites the relevance of an investigation into heterogeneity in the relationship of changing temperature and precipitation with migration with respect to income and agricultural dependency.

4.2 Results

The results section will be structured as follows: First, the main results will determine whether to include control variables into the model and compute results for testing hypotheses 1 and 2; Secondly, the section on heterogeneity will compute results by using the described split samples on income, agricultural dependency, temperature level, and precipitation level to test hypotheses 3 through 9; Note that an independent variable is considered significantly related to the independent variable whenever the p-value of the coefficient is below 0.05, while a p-value below 0.10 indicates a relationship that is somewhat significant. As the p-value measures how likely it is that any observed effect is due to chance, a p-value below 0.05 indicates that there is a less than 5 percent chance that the observed effect is random.

4.2.1 Main results

As also stated above, this section of the results will first determine whether to include control variables in the model. Thereafter, results of the preferred panel regression model are displayed to test hypotheses 1 and 2. In table 2, the regression results of the models that include only the linear variable on precipitation are displayed, where the first model does not include control variables on other potential drivers of migration and the second model does include these control variables. Note that all models hereafter will include region-fixed effects, time-fixed effects, and cluster-robust standard errors as motivated in the methodology.

To assess the validity of measuring migration by negative population change, it is important to first interpret the control variables on fertility and mortality. While both the fertility rate and the mortality rate are significantly related to migration in these models, the fertility rate is suggested to be negatively related to migration whereas the mortality rate is positively related to migration. As the dependent variable on migration is measured as negative population change, it makes sense that the fertility rate is negatively related to the dependent variable because, by definition, an increase in fertility increases the population count. The same goes for mortality rates, only the other way around: An increase in mortality decreases the population count by definition, and therefore mortality rate is positively related to the dependent variable. As a result, controlling for fertility- and mortality rates combined with them being significantly related to the dependent variable in an intuitive way, increases the validity of the measurement of migration.

Dependent var = Migration (negative population change)	(1) Linear precipitation, without control variables	(2) Linear precipitation, with control variables
Temperature change	0.129	0.227*
	(0.092)	(0.092)
Precipitation change	-0.002	0.002
	0.129	(0.003)
Fertility rate	-3.150***	-2.025**
	(0.665)	(0.644)
Mortality rate	0.015***	0.015***
	(0.003)	(0.003)
Income		-4.774**
		(1.803)
Political stability and the absence of violence/terrorism		2.779***
		(0.638)
Voice and accountability		-3.511***
		(0.776)
Observations	2152	2114
R^2	0.133	0.165
Time-fixed effects	Yes	Yes
Regions-fixed effects	Yes	Yes
Cluster-robust SE	Yes	Yes

Note: standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.10

Table 2: Regression results of models with only linear variable on precipitation, and without (1) and with (2) control variables

Now, by comparing the two models to determine whether to include control variables in the model, it is striking that temperature change is only significantly related to migration when controlled for the other migratory drivers. As in this case, the inclusion of control variables when they are significantly related to the dependent variable can affect the effect of the main variables. The variables on other migratory drivers help to isolate the relationship between temperature change and migration by controlling for other factors that influence the relationship. The fact that temperature change is only significant when control variables are included suggests that the relationship between the main variables and the dependent variable is more nuanced and depends on the simultaneous influence of the control variables. Failing to account for these control variables may lead to biased and incomplete conclusions about the effects of temperature and precipitation on migration. Therefore, the model with control variables is accepted as the preferred model for use hereafter.

Next to being significantly related to migration, temperature change is also positively related to migration. Namely, an increase of one percent in temperature is suggested to lead to a 0.227 percent increase in migration. Although this is in line with hypothesis 1, the hypothesis will only be accepted after temperature change is also significantly and positively related to migration in the model including the squared variable on precipitation change. The results are also in line with the expectation formulated by hypothesis 2, as an *increase* in precipitation is suggested to be unrelated to an increase in migration, because the variable precipitation change is not significantly related to the dependent variable on migration. To explore whether a *change* in precipitation does lead to an increase in migration, a third model is run where the squared precipitation change is included as an additional variable, of which results are shown in table 3.

Dependent var = Migration (negative population change)	(2, repeated) Linear precipitation, with control variables	(3) Non-linear precipitation, with control variables
Temperature change	0.227*	0.314**
	(0.092)	(0.099)
Precipitation change	0.002	0.011*
	(0.003)	(0.006)
Precipitation change^2		-0.0001*
		(0.00002)
Fertility rate	-2.025**	-1.906**
	(0.644)	(0.644)
Mortality rate	0.015***	0.015***
	(0.003)	(0.003)
Income	-4.774**	-4.644**
	(1.803)	(1.789)
Political stability and the absence of violence/terrorism	2.779***	2.833***
	(0.638)	(0.641)
Voice and accountability	-3.511***	-3.684***
	(0.776)	(0.789)
Observations	2114	2114
R^2	0.165	0.166
Time-fixed effects	Yes	Yes
Regions-fixed effects	Yes	Yes
Cluster-robust SE	Yes	Yes

Note: standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.10

Table 3: Regression results of models with linear variable on precipitation (2; repeated) and with quadratic variable on precipitation (3)

By comparing the results of this model with the model with only the linear variable on precipitation change, several things stand out. First, the effects as well as the significance of the control variables remain virtually unchanged by the addition of the squared version of the variable precipitation change.

Secondly, and most importantly, precipitation change is suggested to be related to migration through a non-linear quadratic relationship, as both precipitation change as well as the squared precipitation change are significantly related to the dependent variable. Specifically, the quadratic fit of the model displayed in figure 1 indicates that precipitation change is related to migration through an inverted U-shape.

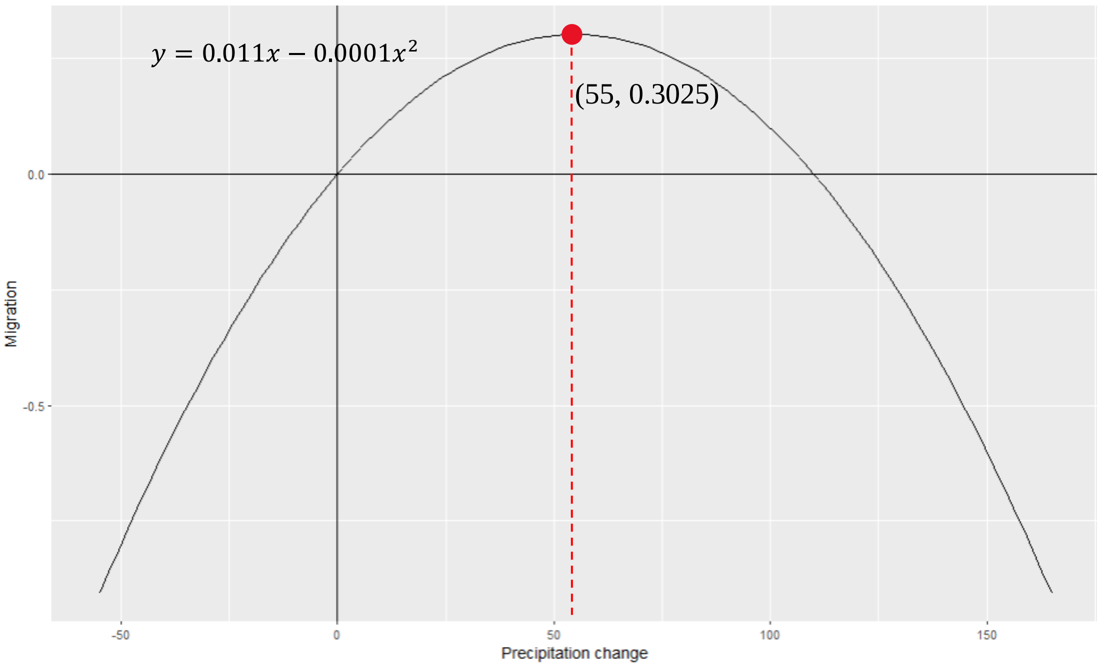


Figure 1: Quadratic fit of the relationship between precipitation change and migration

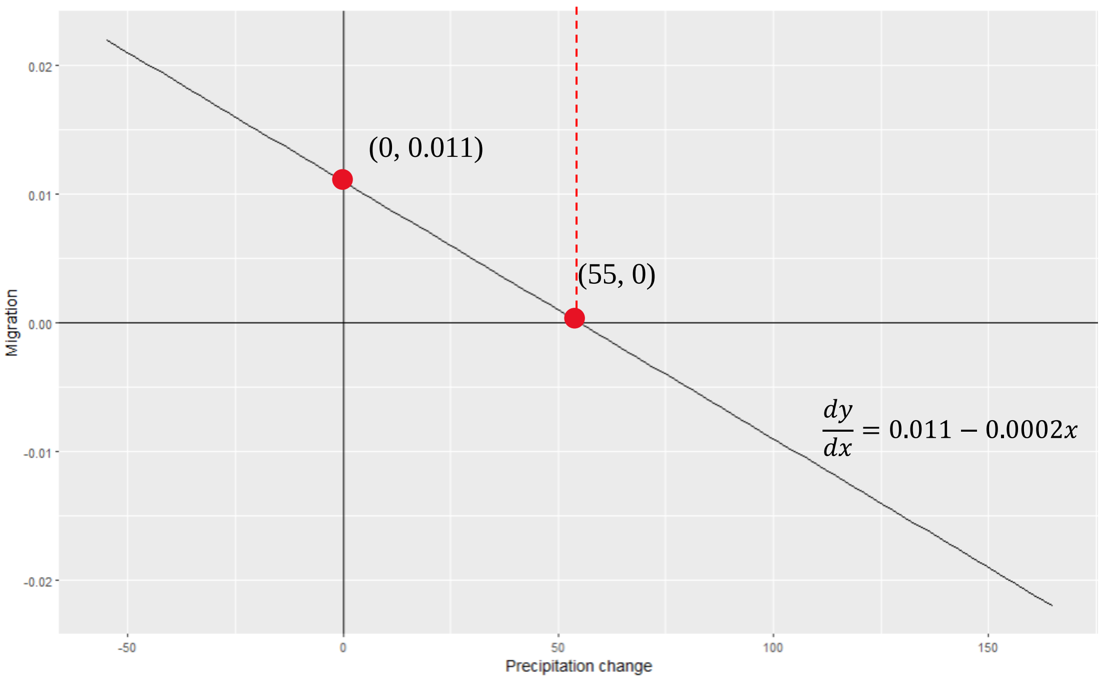


Figure 2: Effect distribution of precipitation change on migration

Because of the non-linear nature of the relationship, the effect of precipitation change on migration differs for each value of precipitation change. Shown visually in figure 2 is the distribution of the isolated effect that a change in precipitation has on migration for different values for the variable. This distribution of effects is obtained from taking the derivative of the part of the functional form that is associated with the effect of precipitation change on migration, of which the process is shown by equations 1 and 2. The distribution of effects for every value for precipitation change is shown by equation 3, and is found by filling the observed coefficients from table 3 into equation 2.

$$MIG_{it} = \beta_2 PRCP_{\ it} + \beta_3 (PRCP_{\ it})^2 \tag{1}$$

$$\frac{dMIG_{it}}{dPRCP_{it}} = \beta_2 + 2 * \beta_3 PRCP_{it} \tag{2}$$

$$\frac{dMIG_{it}}{dPRCP_{it}} = 0.011 - 0.0002 * PRCP_{it} \tag{3}$$

The resulting equation provides the function through which a change in precipitation is associated with a change in migration. As highlighted in figure 2, this results in a basic effect of 0.011 when precipitation change is zero percent. Additionally, the function of equation 3 indicates that with every percentage point increase in the precipitation change, the basic effect is decreased by 0.0002. Therefore, decreases in precipitation sort greater effects on migration than the basic effect, while increases in precipitation decrease the basic effect on migration. As also highlighted in figures 1 and 2, an increase in precipitation of 55 percent is a turning point, where the effect of precipitation change on migration even becomes negative.

The finding that the effect decreases if the change in precipitation increases, can possibly be explained by the suggestion that people may still aspire to migrate but increasingly lose the ability to do so due to the detrimental consequences of greater changes in precipitation. Nevertheless, a deeper look into the data finds that only 137 of the 2114 observations concern changes greater than 55 percent, and so a change in the level of precipitation has led to increased migration in virtually all regions included in the sample. Inversely, with every percentage point decrease in the precipitation change, the basis effect increases by 0.0002, meaning that the effect of precipitation change on migration becomes stronger when the change in precipitation becomes increasingly negative. So, while both a precipitation decrease as well as a precipitation increase up to 55 percent are related to increased migration, precipitation decreases sort stronger effects on migration than precipitation increases. To conclude: because a change in precipitation is overwhelmingly related to an increase in migration, hypothesis 2 - regional precipitation changes are related to increased migration from that region - is accepted.

Moreover, the size of the effect as well as the significance of temperature change has increased after the inclusion of the squared precipitation change, while the direction of the effect has remained the same. Now, an increase in temperature of one percent leads to an increase in migration of 0.314 percentage point. Therefore, the discovered non-linear and significant relationship of precipitation change with migration has further isolated and strengthened the effects of temperature change on migration. This supports the findings of the meta-analyses of Hoffmann et al. (2021), which suggest that the effect of temperature changes is found to be stronger when precipitation changes are controlled for. As temperature change is significantly and positively related to migration in the model with and without the quadratic precipitation change, hypothesis 1 - regional temperature increase is related to increased migration from that region - is accepted.

In addition, because the quadratic relationship between precipitation change and migration is significant and improves the explanatory power of temperature change as well, the squared variable on precipitation is included in the preferred model for the investigation of heterogeneity in the relationship between regions hereafter.

4.2.2 Heterogeneity

In this section of the results, the preferred model will allow the investigation of heterogeneity between regions in how temperature and precipitation changes are related to migration through the use of sub-samples. To test whether differences in effects between the two samples are statistically significant, the approach of Rooks et al. (2016) is adopted, in which a third model is run for the total sample using a dummy variable and interaction effects as explained in the methodology. The significance of this interaction term indicates whether the differences in effects between the two samples are significant.

To start with, the total sample is divided on the basis of the median value of income to test hypotheses 3 and 4. Following the hypotheses, it is expected that the sample including the relatively wealthy regions will display a stronger effect of temperature and precipitation changes on migration, as populations within these regions would have a higher capability of migrating. When comparing the results of the two samples in table 4, it is first important to note that the effects of all control variables significantly differ between the two models. This indicates that it is crucial to control for these factors, as otherwise the differences in the relationship of temperature and precipitation change with migration would possibly be caused by confounding factors that also differ between the two samples.

Because the effect of temperature change significantly differs between the two samples, it can be stated that an increase in temperature has slightly stronger effects on migration in relatively unwealthy regions. Next to that, precipitation changes seem to be only related to migration in relatively unwealthy regions. Specifically, the effects of precipitation change significantly differ between the two samples, where the effect is similar but stronger for the sample with relatively unwealthy regions compared to the results of the total sample, whereas the effect disappears entirely for the sample with relatively wealthy regions. Remarkable is that these results suggest heterogeneity between regions to exist in a way that is exactly opposite to what was predicted by hypotheses 3 and 4. Namely, the hypotheses stated that the effects of temperature increase and precipitation on migration would be stronger in relatively wealthy regions, as the population within these regions would have a higher ability to migrate. Therefore, hypotheses 3 and 4 are rejected.

That the results indicate that actually the effect of temperature increase on migration is stronger for the sample with relatively unwealthy regions, while the effect of precipitation change is

Dependent var = Migration (negative population change)	(1) Below median income sample of regions	(2) Above median income sample of regions	(3) P-value: Is the difference in coefficients significant?
Temperature change	0.369**	0.349*	*
	(0.114)	(0.166)	
Precipitation change	0.040***	0.003	***
	(0.009)	(0.007)	
Precipitation change^2	-0.0003***	-0.00000	***
	(0.00004)	(0.00002)	
Fertility rate	-1.434(*)	-4.934**	***
	(0.735)	(1.572)	
Mortality rate	0.023***	0.011*	**
	(0.003)	(0.004)	
Income	-0.064	-9.354**	*
	(1.852)	(3.061)	
Political stability and the absence of violence/terrorism	1.865**	4.972***	*
	(0.628)	(1.345)	
Voice and accountability	-5.371***	-0.246	***
	(0.904)	(1.367)	
Observations	1054	1060	2114
R^2	0.234	0.227	-
Time-fixed effects	Yes	Yes	Yes
Regions-fixed effects	Yes	Yes	Yes
Cluster-robust SE	Yes	Yes	Yes

Note: standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.10

Table 4: Regression results of the model run for the sub-sample with relatively unwealthy regions (1) and relatively wealthy regions (2), and differences in effects are tested for significance (3)

even only found for this sample, is possibly explained by two reasons. First, these findings suggest that the decision whether to migrate depends on how resilient the population is to a changing climate. Intuitively, the population of relatively unwealthy regions are less able to adapt to a changing climate and deal with its adverse consequences due to a lack of financial resources. As a result, an increase in temperature or change in precipitation in these regions increases the aspiration to migrate due to a lack of resilience. This explanation is supported by the fact that the size of the effects on migration is also greater for the sample with relatively unwealthy regions than for the total sample, which is shown already in table 3. Additionally, this is supported as well by the results for the total sample shown in table 3, specifically by the significance and direction of the coefficient of the control variable on income. The effect of this variable indicates that an increase of 1000 US dollars in income decreases migration by 4.6 percent. These results indicate that an increase in financial resources increases resilience after a changing climate. This suggests resilience to be a moderator of the relationship.

Secondly, a possible reason why past research did find that temperature and precipitation changes are stronger related to increased migration from relatively wealthy regions while this research does not, is that past studies often considered migration across country borders, as opposed to migration within country borders. These studies argue that populations in relatively unwealthy countries are often unable to migrate after temperature and precipitation changes due to a lack of financial resources (Cattaneo & Peri, 2016; Bazzi, 2017; Peri & Sasahara, 2019; Hoffmann et al., 2020). However, migrating within country borders requires much less financial resources than migrating across country borders (Massey et al., 1993). Therefore, the ability to migrate could be a less decisive factor than was predicted by the hypotheses compared to aspiring to do so in the decision of whether to migrate within borders. This could be a distinct difference between the relationship of a changing climate with migration across country borders and migration within country borders, respectively. As a result, the fact that the results suggest that temperature increases and precipitation changes are stronger related to increased migration from relatively unwealthy regions instead of the other way around, might be caused by the sub-national nature of this research.

To test hypotheses 5 and 6, a model is run for each of the two sub-samples created on the basis of the median value of agricultural dependence. It is expected that temperature increases and precipitation changes are stronger related to increased migration in relatively agriculture-dependent regions compared to relatively agriculture-independent regions. The results displayed in table 5 suggest that this is indeed the case, as both the variable on temperature change as well as the variables on precipitation change are only significantly and positively related to migration for the sample with agriculture-dependent regions. However, interpreting this difference should be done with caution as the third model indicates that the difference in the effects of temperature change between the two samples is not significant.

As a result, hypothesis 5 - the effect of regional temperature increase on migration is stronger for agriculture-dependent regions - is accepted with the caveat that the difference between the two samples is not found to be statistically significant. In contrast, the differences between the two samples in the effects of the variables on precipitation change are, in fact, significant. Therefore, hypothesis 6 - the effect of regional precipitation change on migration is stronger for agriculture-dependent regions – is accepted. Additionally, as temperature increase and precipitation change are only significantly related to increased migration for relatively agriculture-dependent regions, agriculture seems to be a mechanism through which temperature and precipitation changes affect migration.

To test hypothesis 7 - the effect of temperature increase on migration is stronger for relatively hot regions -, again a model is run for each of the two sub-samples created on the basis of the median

Dependent var = Migration (negative population change)	(1) Below median agricultural dependency sample of regions	(2) Above median agricultural dependency sample of regions	(3) P-value: Is the difference in coefficients significant?
Temperature change	0.289	0.315***	-
	(0.178)	(0.090)	
Precipitation change	-0.008	0.045***	***
	(0.008)	(0.008)	
Precipitation change^2	0.00002	-0.0003***	***
	(0.00002)	(0.00005)	
Fertility rate	-4.459*	-2.753***	***
	(1.868)	(0.452)	
Mortality rate	-0.001	0.018***	***
	(0.007)	(0.003)	
Income	-6.617*	-1.495*	-
	(3.280)	(0.678)	
Political stability and the absence of violence/terrorism	3.645**	0.745**	*
	(1.159)	(0.272)	
Voice and accountability	-3.731**	-3.596***	-
	(1.260)	(0.644)	
Observations	1068	1046	2114
R^2	0.240	0.227	-
Time-fixed effects	Yes	Yes	Yes
Regions-fixed effects	Yes	Yes	Yes
Cluster-robust SE	Yes	Yes	Yes

Note: standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.10

Table 5: Regression results of the model run for the sub-sample with relatively agriculture-independent regions (1) and relatively agriculture-dependent regions (2), and differences in effects are tested for significance (3)

level of temperature, and differences in effects are tested for significance. In contrast to what was expected, the effect of temperature increase on migration does not differ significantly between relatively cool and relatively hot regions. That suggests that the initial level of temperature of a region does not affect the relationship between temperature change and migration, and therefore an increase in temperature seems to be similarly detrimental to livelihoods in relatively cooler and relatively hotter regions. As a result, hypothesis 7 is rejected. However, it is important to note that the mean level of temperature for the sample with relatively cool regions is with 21.2 degrees Celsius still well above the global mean temperature of around 14 degrees Celsius (Menne et al., 2018). Therefore, it cannot

Dependent var = Migration (negative population change)	(1) Below median level of temperature sample of regions	(2) Above median level of temperature sample of regions	(3) P-value: Is the difference in coefficients significant?
Temperature change	0.387**	0.398***	-
	(0.146)	(0.140)	
Precipitation change	0.013(*)	0.013(*)	-
	(0.009)	(0.008)	
Precipitation change^2	-0.0001(*)	-0.0001(*)	-
	(0.00003)	(0.0001)	
Fertility rate	-2.085**	-1.299(*)	-
	(0.748)	(1.595)	
Mortality rate	0.030***	-0.030(*)	(*)
	(0.005)	(0.003)	
Income	-4.006	-4.813(*)	-
	(2.449)	(2.555)	
Political stability and the absence of violence/terrorism	4.265***	2.396**	-
	(1.134)	(0.758)	
Voice and accountability	-0.486	-6.010***	***
	(1.006)	(1.052)	
Observations	1075	1039	2114
R^2	0.216	0.196	-
Time-fixed effects	Yes	Yes	Yes
Regions-fixed effects	Yes	Yes	Yes
Cluster-robust SE	Yes	Yes	Yes

Note: standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.10

Table 6: Regression results of the model run for the sub-sample with relatively cool regions (1) and relatively hot regions (2), and differences in effects are tested for significance (3)

be ruled out that the hypothesis would hold when comparing truly cool regions with truly hot regions, instead of relatively cool and hot regions.

Finally, the total sample is divided on the basis of the median level of precipitation to test hypothesis 8. For the sample with relatively dry regions, it is expected that only a decrease in precipitation is related to increased migration, whereas for the sample with relatively wet regions, a further increase in precipitation is expected to increase migration. To begin with the sample of relatively dry regions, the expectation is not reflected by the results, displayed in table 7. Similar to the results for the total sample, a decrease as well as an increase in precipitation is related to increased migration. While the derivative of the quadratic variable on precipitation change indicates that

Dependent var = Migration (negative population change)	(1) Below median level of precipitation sample of regions	(2) Above median level of precipitation sample of regions	(3) P-value: Is the difference in coefficients significant?
Temperature change	0.507**	0.359**	(*)
	(0.159)	(0.136)	
Precipitation change	0.015**	0.024*	*
	(0.006)	(0.011)	
Precipitation change^2	-0.00003(*)	-0.0001	(*)
	(0.00002)	(0.0001)	
Fertility rate	-5.321***	3.855**	**
	(1.086)	(1.456)	
Mortality rate	0.010***	-0.009(*)	**
	(0.002)	(0.011)	
Income	-6.653***	-2.322	(*)
	(1.985)	(2.528)	
Political stability and the absence of violence/terrorism	0.547	7.437***	***
	(0.635)	(1.124)	
Voice and accountability	0.295	-10.855***	***
	(0.770)	(1.377)	
Observations	1052	1062	2114
R^2	0.171	0.306	-
Time-fixed effects	Yes	Yes	Yes
Regions-fixed effects	Yes	Yes	Yes
Cluster-robust SE	Yes	Yes	Yes

Note: standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.10

Table 7: Regression results of the model run for the sub-sample with relatively dry regions (1) and relatively wet regions (2), and differences in effects are tested for significance (3)

decreases do sort stronger effects on migration than precipitation increases, it is only above an increase in precipitation of 250 percent that a precipitation change leads to decreased migration. Therefore, hypothesis 8 - in relatively dry regions, regional precipitation decreases are related to increased migration - is rejected, despite that precipitation decreases lead to slightly more migration than precipitation increases do in relatively dry regions.

When comparing the results for the sample with relatively wet regions to the results for the sample with relatively dry regions, several significant differences occur. First, an increase in temperature seems to be stronger related to increased migration in relatively dry regions, although the difference is only significant at the 0.10 level. Secondly, and most importantly, the non-linear

Dependent var = Migration (negative population change)	(2, repeated) Above median level of precipitation sample of regions	(4) Above median level of precipitation sample without squared variable
Temperature change	0.359**	0.271*
	(0.136)	(0.124)
Precipitation change	0.024*	0.013*
	(0.011)	(0.006)
Precipitation change^2	-0.0001	
	(0.0001)	
Fertility rate	3.855**	3.652*
	(1.456)	(1.457)
Mortality rate	-0.009(*)	-0.008(*)
	(0.011)	(0.011)
Income	-2.322	-2.352
	(2.528)	(2.526)
Political stability and the absence of violence/terrorism	7.437***	7.491**
	(1.124)	(1.115)
Voice and accountability	-10.855***	-10.774**
	(1.377)	(1.384)
Observations	1062	1062
R^2	0.306	0.305
Time-fixed effects	Yes	Yes
Regions-fixed effects	Yes	Yes
Cluster-robust SE	Yes	Yes

Note: standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, (*) p<0.10

Table 8: Regression results for the sub-sample with relatively wet regions with (2, repeated) and without (4) the squared variable on precipitation change

relationship between precipitation change and migration does not hold anymore for the sample with relatively wet regions, because the squared precipitation is no longer significantly related to migration. The fact that the non-quadratic precipitation change is still significantly related to migration suggests that for relatively wet regions a linear relationship between precipitation change and migration exists. However, due to the presence of the quadratic variable in the model, this effect cannot be interpreted as such. Therefore, an additional model is run for the sample with relatively wet regions, but now excluding the quadratic version of precipitation change. The results shown in table 8 indicate that precipitation change indeed is related to migration through a linear relationship for the sample with relatively wet regions. Therefore, only an increase, and not an decrease, in precipitation is related to increased migration in relatively wet regions. Note that this is exactly in line with the expectation on relatively wet regions formulated in hypothesis 9. As a result, hypothesis 9 - in relatively wet regions, regional precipitation increases are related to increased migration - is accepted.

These findings in combination with the sizes of the effects of precipitation change on migration suggests that populations in relatively dry regions are more vulnerable to precipitation decreases than populations in relatively wet regions, while somewhat equally vulnerable to precipitation increases.

5. Conclusion and discussion

5.1 Conclusion

This thesis investigates the extent to which climate change is a migratory driver in low- and middle-income countries by estimating the effect of sub-national regional temperature and precipitation changes on migration from that region. To this end, the preceding research has focused on migration due to gradual changes in temperature and precipitation in Africa.

Compared to existing research, the contribution of this thesis is twofold. First, this is the first multi-country research that explores the extent to which climate change is a migratory driver at the sub-national level. This is made possible by the novel sub-national dataset on temperature, precipitation, and population distribution in Africa used in this research. Secondly, while gradual climate change is increasingly recognised as an important driver of migration, no clear consensus on the impact of temperature and precipitation changes on migration has been reached due to great heterogeneity in past results (Black et al., 2011; Helbling et al., 2023). As a result, the results of this research contribute to the academic field by providing new and distinctive evidence on the relationship between climate change and migration.

The empirical analysis of data on 529 African regions observed from 2000 through 2020 finds that temperature increases and precipitation changes are significantly and predominantly positively related to increased migration. Namely, a one percent increase in regional temperature over 5 years is related to an increase of 0.31 percentage point in migration from that region. Next to that, precipitation decreases as well as precipitation increases up to 55 percent lead to increased migration, where decreases in precipitation sort increasingly stronger effects on migration than increases. As a result, a change in precipitation over 5 years is linked to migration through a non-linear quadratic relationship. These findings are in line with conclusions drawn by a multitude of authoritative studies that considered migration on the country-level, particularly with those of Marchiori et al. (2012), Backhaus et al. (2015), Cattaneo & Peri (2016), and Hoffmann et al., (2022). Moreover, part of the lack of consensus in past results is caused by some studies that found no significant relationship between precipitation change and migration but estimated the effect of precipitation change on migration as a linear relationship (Mueller et al., 2014; Bohra-Mishra et al., 2017; Hoffmann, 2021). By estimating the effect of precipitation change on migration as a non-linear relationship, one should find that both decreases as well as increases lead to increased migration across different contexts.

Interestingly, the results on heterogeneity indicate that the relationship of a changing climate with migration differs between, respectively, relative unwealthy and wealthy regions, relative agriculture-independent and agriculture-dependent regions, and relatively dry and wet regions. For starters, the effects of temperature increases on migration are slightly stronger in relatively unwealthy regions compared to relatively wealthy regions, whereas precipitation changes are related to increased migration only in relatively unwealthy regions. This is exactly the opposite of what the expectation embodied by the hypotheses stated and what past research found at the country level, which was that temperature and precipitation changes particularly induce migration from relatively wealthy regions or countries due to the increased capability to migrate. This finding suggests that the ability to migrate could be a less decisive factor to migrate or not at the sub-national level compared to the country level due to the lower costs of migration within borders. As a result, the capability to migrate being less of an issue for internal movements combined with the ability of the novel dataset to detect next to international movements also internal movements could explain the finding that an increase in temperature or a change in precipitation increases the aspiration to migrate especially from relatively unwealthy regions due to the relative lack of resilience of the population. Resilience is, therefore, suggested to be a moderator of the relationship.

Next to that, the results suggest that temperature increases and precipitation changes only lead to increased migration in relatively agriculture-dependent regions. Agriculture, therefore, seems to be a channel through which temperature and precipitation changes affect migration. This is in line with past country-level studies that also argued agriculture to be the mechanism in the relationship between a changing climate and migration, particularly Falco et al. (2019), Mastrorillo et al. (2016), and Cai et al. (2016). Furthermore, no difference between relatively cool and relatively hot regions is found in the relationship of a changing climate with migration. Finally, the results indicate that for relatively wet regions only precipitation increases are related to increased migration. In contrast, for relatively dry regions both precipitation decreases and increases up to 250 percent are related to increased migration, despite the fact precipitation decreases sort stronger effects on migration than precipitation increases. This finding suggests that populations in relatively dry regions are more vulnerable to precipitation decreases than relatively wet regions, while equally vulnerable to precipitation increases.

By relating found results to the research question, it is concluded that a changing climate indeed has been a migratory driver in the African regions over the past two decades, in particular for relatively unwealthy- and agriculture-dependent regions. Namely, temperature increases and precipitation changes are overwhelmingly related to increased migration, while for relatively wet regions only increasing precipitation is related to increased migration.

5.2 Discussion

This section will first reflect on the limitations of the research by examining the internal and external validity of the empirical analysis. Thereafter, suggestions for the way forward for future research on this topic are given, and policy implications of the results are formulated.

The internal validity of the research concerns the extent to which the observed results also represent the truth for the regions included in the sample, and are not due to methodological errors. Several factors could have affected this internal validity. For starters, by estimating migration through negative population growth of that region while controlling for fertility- and mortality rates, estimates could differ from true migration numbers. Namely, only registered people are detected with this statistic. As only half of Africa's population is registered at birth, true emigration from regions is likely to be underestimated with chosen measurement (United Nations, 2022). Besides that, although measuring migration at the regional level is a significant improvement over measuring migration at the country level, it is not able to detect within-region movements. Despite that the regional level has been chosen because temperature and precipitation levels usually do not vary much within these regions, within-region movements can still be driven by a changing climate. For instance, people employed in agriculture could move to areas where they are less exposed and vulnerable to weather and climate, such as cities or places with more predictable access to water. Therefore, measuring migration at the individual level would improve the estimation of true migration numbers. The difficulty of estimating migration accurately is widely acknowledged, and is a key area where methodological innovation is necessary to improve the evidence base (Ardittis & Laczko, 2020).

Moreover, country-level data on fertility- and mortality rates are used for regions within that country due to the absence of good quality sub-national data. By doing so, this research adopts the assumption made by Alessandrini et al. (2020) of similarity in fertility and mortality rates between the country- and regional level in the process of estimating migration. Any resulting under- or overestimation of fertility and mortality for specific region-years would create a bias in the measurement of migration, as observed migration would partly be driven by positive or negative population growth. While the assumption of similarity between the regional level and the country level seems to be reasonable for fertility rates, this might be too strong of an assumption for mortality rates, conflicts, and political regimes. The results of this research, therefore, should be tested for robustness when sub-national data on fertility- and mortality rates, as well as sub-national data on agricultural dependency, conflicts, and political regimes, become available.

As explained, the findings on heterogeneity between relatively cool and hot regions could be dependent on the basis on which the total sample is divided into two sub-samples. The total sample is divided on the basis of the mean level of temperature, however, the mean value for the sample with relatively cool regions is still well above the global mean temperature. Therefore, it cannot be ruled out that the results would change when comparing truly cool and hot regions instead of relatively cool and hot regions. In its turn, this affects the external validity as well, as one cannot conclude from these results that for every other continent relatively cool and hot regions also do not differ in the relationship between a changing climate and migration.

Namely, external validity concerns the extent to which the findings of this research can be generalized to other contexts. Africa is just like other continents a unique place and is characterized by the prevalence of poverty, a fast-growing population, political instability, and widespread conflict. By controlling for income, fertility- and mortality rates, political regime, and conflicts on the one hand and region- and time-fixed effects on the other, it is attempted to make the results less dependent on the African context. However, many unobserved factors such as history and culture make Africa unique and could, as a result, have affected the outcomes. Therefore, it cannot be excluded that the results are still context-dependent, and the same analysis would sort different results for other continents. Next to that, one should not assume that the observed effects are exactly the same across regions in Africa. Africa is far from a homogenous place, and even within countries, the regional effects of climate change on migration can still differ, even though the omitted variable bias is attempted to be minimalized.

In addition, to estimate the effects of changing climate on migration even more accurately, further research is necessary. Specifically, this research should be reproduced when sub-national data on mentioned variables become available and for other continents to test the robustness of found results. Next to that, it should be interesting to investigate additional potential channels at the regional level through which a changing climate affects migration by means of a more in-depth mediation analysis. However, this is also conditional on the availability of sub-national data that would enable such a study. Finally, as this research has focused on migration due to gradual changes in climate, in particular temperature and precipitation changes, future research should adopt a similar sub-national approach to explore the effects of other types of climate-variability on migration, such as extreme weather, droughts and disasters. Together, a more complete picture of the total effects of climate change on migration can be painted.

In conclusion, as migration and climate change are two of the largest global challenges humanity is currently facing, the findings of this research could have significant policy implications. Namely, policymakers could, in combination with climate models, use the results on the extent to which a changing climate is a migratory driver to predict future population distributions and which regions will experience immigration or emigration and act accordingly to mitigate adverse consequences for affected populations. Due to the projected size of the African population and the size of found effects, climate-induced migration in Africa can roughly go up to a hundred million people in the coming decades. To prevent or mitigate migration, the results suggest that policymakers should focus on increasing resilience to a changing climate, especially in agriculture-dependent regions.

6. Bibliography

Abel, G. J., Brottrager, M., Cuaresma, J. C., & Muttarak, R. (2019). Climate, conflict and forced migration. Global environmental change, 54, 239-249.

Adhikari, P. (2013). Conflict-induced displacement, understanding the causes of flight. American Journal of Political Science, 57(1), 82-89.

Adams, H., & Adger, W. N. (2013). The contribution of ecosystem services to place utility as a determinant of migration decision-making. Environmental Research Letters, 8(1), 015006.

Adger, W. N., Huq, S., Brown, K., Conway, D., & Hulme, M. (2003). Adaptation to climate change in the developing world. Progress in development studies, 3(3), 179-195.

Afifi, T. (2011). Economic or environmental migration? The push factors in Niger. International Migration, 49, e95-e124.

Afifi, T., Liwenga, E., & Kwezi, L. (2014). Rainfall-induced crop failure, food insecurity and out-migration in Same-Kilimanjaro, Tanzania. Climate and Development, 6(1), 53-60.

Alessandrini, A., Ghio, D., & Migali, S. (2020). Estimating net migration at high spatial resolution. Publications Office of the European Union.

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Appadurai, A. (2004). The capacity to aspire: Culture and the terms of recognition. Culture and public action, 59, 62-63.

Ardittis, S., & Laczko, F. (2020). Introduction-Migration policy in the age of immobility. Migration Policy Practice, 10(2), 2-7.

Auffhammer, M., Hsiang, S. M., Schlenker, W., & Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. Review of Environmental Economics and Policy.

Backhaus, A., Martinez-Zarzoso, I., & Muris, C. (2015). Do climate variations explain bilateral migration? A gravity model analysis. IZA Journal of Migration, 4(1), 1-15.

Bakewell, O. (2010). Some reflections on structure and agency in migration theory. Journal of ethnic and migration studies, 36(10), 1689-1708.

Bazzi, S. (2017). Wealth heterogeneity and the income elasticity of migration. American Economic Journal: Applied Economics, 9(2), 219-255.

Beegle, K,; Christiaensen, L., Dabalen, A., Gaddis, I., (2016). Poverty in a Rising Africa: Africa Poverty Report. Washington, DC: World Bank.

Beine, M., & Parsons, C. (2015). Climatic factors as determinants of international migration. The Scandinavian Journal of Economics, 117(2), 723-767.

Berlemann, M., & Steinhardt, M. F. (2017). Climate change, natural disasters, and migration—a survey of the empirical evidence. CESifo Economic Studies, 63(4), 353-385.

Bertoli, S., Docquier, F., Rapoport, H., & Ruyssen, I. (2022). Weather shocks and migration intentions in Western Africa: Insights from a multilevel analysis. Journal of Economic Geography, 22(2), 289-323.

Biella, R., Hoffmann, R., & Upadhyay, H. (2022). Climate, Agriculture, and Migration: Exploring the Vulnerability and Outmigration Nexus in the Indian Himalayan Region. Mountain Research and Development, 42(2), R9-R21.

Bircan, T., Purkayastha, D., Ahmad-yar, A. W., Lotter, K., Iakono, C. D., Göler, D., ... & Ünver, Ö. (2020). Gaps in Migration Research: Review of migration theories and the quality and compatibility of migration data on the national and international level. HumMingBird Paper, (July).

Black, R., Adger, W. N., Arnell, N. W., Dercon, S., Geddes, A., & Thomas, D. (2011). The effect of environmental change on human migration. Global environmental change, 21, S3-S11.

Bohra-Mishra, P., Oppenheimer, M., Cai, R., Feng, S., & Licker, R. (2017). Climate variability and migration in the Philippines. Population and environment, 38, 286-308.

Bohra-Mishra, P., Oppenheimer, M., & Hsiang, S. M. (2014). Nonlinear permanent migration response to climatic variations but minimal response to disasters. Proceedings of the National Academy of Sciences, 111(27), 9780-9785.

Boyd, M. (1989). Family and personal networks in international migration: recent developments and new agendas. International migration review, 23(3), 638-670.

Brzoska, M., & Fröhlich, C. (2016). Climate change, migration and violent conflict: vulnerabilities, pathways and adaptation strategies. Migration and Development, 5(2), 190-210.

Cai, R., Feng, S., Oppenheimer, M., & Pytlikova, M. (2016). Climate variability and international migration: The importance of the agricultural linkage. Journal of Environmental Economics and Management, 79, 135-151.

Carling, J., & Schewel, K. (2018). Revisiting aspiration and ability in international migration. Journal of Ethnic and Migration Studies, 44(6), 945-963.

Carling, J., & Talleraas, C. (2016). Root causes and drivers of migration. Oslo: Peace Research Institute Oslo (PRIO), 2016, 1-44.

Castles, S., & Van Hear, N. (2011). Root causes. Global migration governance, 287-306.

Cattaneo, C., & Peri, G. (2016). The migration response to increasing temperatures. Journal of development economics, 122, 127-146.

CIESIN - Columbia University (2018). Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC).

Coniglio, N. D., & Pesce, G. (2015). Climate variability and international migration: an empirical analysis. Environment and Development Economics, 20(4), 434-468.

Czaika, M., & Reinprecht, C. (2022). Migration Drivers: Why Do People Migrate?. Introduction to Migration Studies, 49.

Czaika, M., & Reinprecht, C. (2022a). Why do people stay put in environmentally stressful regions? Cognitive bias and heuristics in migration decision-making. Regional Environmental Change, 22(3), 84.

Dallmann, I., & Millock, K. (2017). Climate variability and inter-state migration in India. CESifo Economic Studies, 63(4), 560-594.

Dell, M., Jones, B. F., & Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. Journal of Economic literature, 52(3), 740-798.

Esipova, N., Ray, J., & Pugliese, A. (2011). Gallup World Poll: the many faces of global migration. United Nations.

Esipova, N., Pugliese, A., & Ray, J. (2018). More than 750 million worldwide would migrate if they could. WORLD, 13, 14.

Falco, C., Galeotti, M., & Olper, A. (2019). Climate change and migration: is agriculture the main channel?. Global Environmental Change, 59, 101995.

Fischer, P.A., Martin, R., Straubhaar, T. (1997) Should I stay or should I go. In: Hammar T, Brochmann G, Tamas K (eds) International migration, immobility and development. Berg Press, Oxford, pp 49–90

Global Data Lab (2023). “Share of the population working in agriculture (men/women, %)” Area Database, v.4.2.1. Institute for Management Research, Radboud University, Nijmegen, The Netherlands.

Global Data Lab (2023a). “Fertility rates (births per woman)” Area Database, v.4.2.1. Institute for Management Research, Radboud University, Nijmegen, The Netherlands.

Global Data Lab (2023b). “Log gross national income per capita in 1000 US Dollars (2011 PPP)” Subnational HDI, v.7.0. Institute for Management Research, Radboud University, Nijmegen, The Netherlands.

Global Data Lab (2023c). “Life expectancy” Area Database, v.4.2.1. Institute for Management Research, Radboud University, Nijmegen, The Netherlands.

Global Data Lab (2023d). “GDL Vulnerability Index”. Institute for Management Research, Radboud University, Nijmegen, The Netherlands.

Global Modeling and Assimilation Office (GMAO) (2015). MERRA-2 instM_2d_asm_Nx: 2d,Monthly mean, Single-Level, Assimilation, Single-Level Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC)

Global Modeling and Assimilation Office (GMAO) (2015a). MERRA-2 tavgM_2d_flx_Nx: 2d, Monthly mean, Time-Averaged, Single-Level, Assimilation, Surface Flux Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC)

Government Office for Science (2011) Foresight: Migration and Global Environmental Change Final Project Report. Government Office for Science, London.

Gray, C. L. (2009), "Environment, Land, and Rural Out-migration in the Southern Ecuadorian Andes", World Development 37, 457–468.

Gray, C., & Mueller, V. (2012). Drought and population mobility in rural Ethiopia. World development, 40(1), 134-145.

Gray, C., & Wise, E. (2016). Country-specific effects of climate variability on human migration. Climatic change, 135, 555-568.

Gubhaju, B., & De Jong, G. F. (2009). Individual versus household migration decision rules: Gender and marital status differences in intentions to migrate in South Africa. International migration, 47(1), 31-61.

Hagen-Zanker, J. (2008). Why do people migrate? A review of the theoretical literature. A Review of the Theoretical Literature (January 2008). Maastricht Graduate School of Governance Working Paper No.

Harris, J. R., & Todaro, M. P. (1968). A two sector model of migration with urban unemployment in developing economies.

Helbling, M., Rybski, D., Schewe, J., Siedentop, S., Glockmann, M., Heider, B., Stroms, P. (2023). Measuring the effect of climate change on migration flows: Limitations of existing data and analytical frameworks. PLOS Climate, 2(1), e0000078.

Henderson, J. V., Storeygard, A., & Deichmann, U. (2017). Has climate change driven urbanization in Africa?. Journal of development economics, 124, 60-82.

Henry, S., Piché, V., Ouédraogo, D., & Lambin, E. F. (2004a). Descriptive analysis of the individual migratory pathways according to environmental typologies. Population and Environment, 25(5), 397-422.

Henry, S., Schoumaker, B., & Beauchemin, C. (2004b). The impact of rainfall on the first out-migration: A multi-level event-history analysis in Burkina Faso. Population and environment, 25(5), 423-460.

Hoffmann, R., Dimitrova, A., Muttarak, R., Crespo Cuaresma, J., & Peisker, J. (2020). A meta-analysis of country-level studies on environmental change and migration. Nature Climate Change, 10(10), 904-912.

Hoffmann, R., Šedová, B., & Vinke, K. (2021). Improving the evidence base: A methodological review of the quantitative climate migration literature. Global Environmental Change, 71, 102367.

Hunter, L. M., Luna, J. K., & Norton, R. M. (2015). Environmental dimensions of migration. Annual Review of Sociology, 41, 377-397.

Hunter, L. M., Nawrotzki, R., Leyk, S., Maclaurin, G. J., Twine, W., Collinson, M., & Erasmus, B. (2014). Rural outmigration, natural capital, and livelihoods in South Africa. Population, space and place, 20(5), 402-420.

Hsiang, S. (2016). Climate econometrics. Annual Review of Resource Economics, 8, 43-75.

Institute for Economics & Peace (IEP) (2022). Ecological Threat Report: Analysing Ecological Threats, Resilience & Peace, Sydney, October 2022.

International Organization for Migration (IOM) (2020). World Migration Report 2020 - International Organization for Migration. United Nations Digital Library.

IPCC (2021). Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 3−32, doi:10.1017/9781009157896.001.

IPCC (2022). Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA,

Islam, M. R. (2018). Climate change, natural disasters and socioeconomic livelihood vulnerabilities: migration decision among the char land people in Bangladesh. Social Indicators Research, 136, 575-593.

Khavarian-Garmsir, A. R., Pourahmad, A., Hataminejad, H., & Farhoodi, R. (2019). Climate change and environmental degradation and the drivers of migration in the context of shrinking cities: A case study of Khuzestan province, Iran. Sustainable Cities and Society, 47, 101480.

King, R. (2012). Willy Brandt Series of Working Papers in International Migration and Ethnic Relations 3/12.

Lee, E. S. (1966). A theory of migration. Demography, 3, 47-57.

Lee, H., Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P., ... & Zommers, Z. (2023). AR6 Synthesis Report: Climate Change 2023. Summary for Policymakers.

Marchiori, L., Maystadt, J. F., & Schumacher, I. (2012). The impact of weather anomalies on migration in sub-Saharan Africa. Journal of Environmental Economics and Management, 63(3), 355-374.

Marchiori, L., Maystadt, J. F., & Schumacher, I. (2017). Is environmentally induced income variability a driver of human migration?. Migration and Development, 6(1), 33-59.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., & Pellegrino, A. (1999). Worlds in motion: understanding international migration at the end of the millennium: understanding international migration at the end of the millennium. Clarendon Press.

Massey, D., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A. & Taylor, J.E. (1993). 'Theories of international migration: a review and appraisal'. Population and Development Review, 19(3): 431–66.

Mastrorillo, M., Licker, R., Bohra-Mishra, P., Fagiolo, G., Estes, L. D., & Oppenheimer, M. (2016). The influence of climate variability on internal migration flows in South Africa. Global Environmental Change, 39, 155-169.

Maurel, M., & Tuccio, M. (2016). Climate instability, urbanisation and international migration. The Journal of Development Studies, 52(5), 735-752.

McAuliffe, M., & Triandafyllidou, A. (2022). 4 Migration Research and Analysis: Recent United Nations Contributions. World Migration Report, 2022(1), e00025.

McLeman, R., & Smit, B. (2006). Migration as an adaptation to climate change. Climatic change, 76(1-2), 31-53.

Menne, M. J., Williams. C. N., Gleason, B. E., Rennie, J. J, & Lawrimore, J. H. (2018). The Global Historical Climatology Network Monthly Temperature Dataset, Version 4. J. Climate, in press.

Migali, S., Natale, F., – (eds.), Population exposure and migrations linked to climate change in Africa, EUR 30881 EN, Publications Office of the European Union, Luxembourg, 2021, JRC126594

Mueller, V., Gray, C., & Kosec, K. (2014). Heat stress increases long-term human migration in rural Pakistan. Nature climate change, 4(3), 182-185.

Mueller, V. A., & Osgood, D. E. (2009). Long-term consequences of short-term precipitation shocks: evidence from Brazilian migrant households. Agricultural Economics, 40(5), 573-586.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. The Quarterly Journal of Economics, 118(2), 549-599.

Nawrotzki, R. J., & DeWaard, J. (2018). Putting trapped populations into place: climate change and inter-district migration flows in Zambia. Regional environmental change, 18, 533-546.

Nawrotzki, R. J., Hunter, L. M., Runfola, D. M., & Riosmena, F. (2015). Climate change as a migration driver from rural and urban Mexico. Environmental Research Letters, 10(11), 114023.

OECD (2015), Connecting with Emigrants: A Global Profile of Diasporas 2015, OECD Publishing, Paris.

Perch-Nielsen, S.L., Bättig, M.B., & Imboden, D. (2008). Exploring the link between climate change and migration. Climatic change, 91, 375-393.

Peri, G., & Sasahara, A. (2019). The impact of global warming on rural-Urban migrations: Evidence from global big data (No. w25728). National Bureau of Economic Research.

Piguet, E. (2010). Linking climate change, environmental degradation, and migration: a methodological overview. Wiley Interdisciplinary Reviews: Climate Change, 1(4), 517-524.

Reuveny, R. (2007). Climate change-induced migration and violent conflict. Political geography, 26(6), 656-673.

Ritchey, P. N. (1976). Explanations of migration. Annual review of sociology, 2(1), 363-404.

Rooks, G., Klyver, K., & Sserwanga, A. (2016). The context of social capital: A comparison of rural and urban entrepreneurs in Uganda. Entrepreneurship theory and Practice, 40(1), 111-130.

Max Roser, Esteban Ortiz-Ospina and Hannah Ritchie (2013) - "Life Expectancy". Published online at OurWorldInData.org.

Schmidt, A. F., & Finan, C. (2018). Linear regression and the normality assumption. Journal of clinical epidemiology, 98, 146-151.

Sjaastad, L. A. (1962). The costs and returns of human migration. Journal of political Economy, 70(5, Part 2), 80-93.

Smits, J. (2016). GDL Area Database. Nijmegen, The Netherlands, 16-101.

Stark, O., & Taylor, J. E. (1991). Migration incentives, migration types: The role of relative deprivation. The economic journal, 101(408), 1163-1178.

Taylor, J. E. (1996). Development strategy, employment and migration: insights from models (Vol. 1). OECD Publishing.

UNDP (2009). Human Development Report 2009: Overcoming Barriers: Human Mobility and Development. Palgrave Macmillan, Houndmills, Basingstoke, Hampshire, UK and New York, NY, USA, 217 pp

United Nations (2022) Children’s Fund, A Statistical Update on Birth Registration in Africa, UNICEF, New York.

United Nations (2023). Global Issues. United Nations. Retrieved from: https://www.un.org/en/global-issues

World Bank (2023). “Employment in agriculture (% of total employment)”. World Development Indicators, The World Bank Group, Washington, D.C., United States.

World Bank (2023a). “Fertility rate (births per woman)”. World Development Indicators, The World Bank Group, Washington, D.C., United States.

World Bank (2023b). “Mortality rate (per 1000 male/female adults)”. World Development Indicators, The World Bank Group, Washington, D.C., United States.

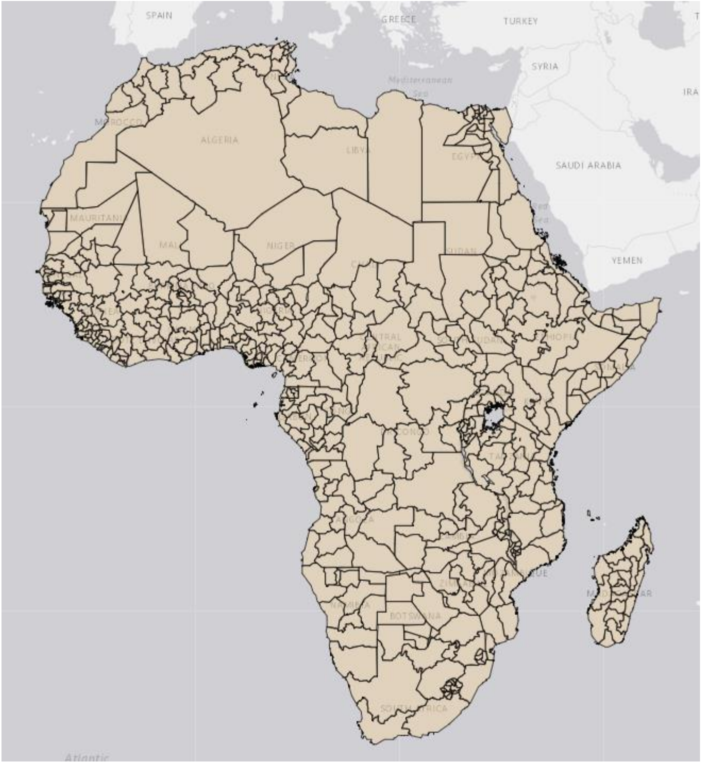
World Bank (2023c). “Political stability and absence of violence/terrorism: Estimate”. World Development Indicators, The World Bank Group, Washington, D.C., United States.

World Bank (2023d). “Voice and accountability: Estimate”. World Development Indicators, The World Bank Group, Washington, D.C., United States.

Yavçan, B. (2021). The influence of long-term environmental trends on migration dynamics. Working Paper, MAGYC project.

7. Appendices

7.1 Countries and sub-national regions included in the sample



7.2 Overview variables

Type of variable	Indicator	Measurement	Source
Dependent variable	Migration (negative population change)	*Negative change in population count living in the region (%)*	NASA SEDAC GPWv4.11, 2018 (CEISIN – Columbia University, 2018)
Independent variable	Temperature change	*Change in 5 year average temperature per region (%)*	NASA Earth Data; MERRA-2 (GMAO, 2015)/Own calculation
Independent variable	Precipitation change	*Change in 5 year average temperature per region (%)*	NASA Earth Data; MERRA-2 (GMAO, 2015a)/Own calculation
Independent variable	Agricultural dependency	*National share of the population working in agriculture (%)*	World Bank; World Development Indicators (World Bank, 2023)
Independent variable	Agricultural dependency	*Sub-national share of the population working in agriculture (%)*	Global Data Lab; Area Database (v.4.2.1) (Global Data Lab, 2023; Smits, 2016)
Independent variable	Wealth	*Log gross national income per capita in 1000 US Dollars (2011 PPP)*	Global Data Lab; Subnational HDI (v7.0) (Global Data Lab, 2023b; Smits, 2016)
Control variable	Fertility	*National fertility rate (births per woman)*	World Bank; World Development Indicators (World Bank, 2023a)
Control variable	Fertility	*Sub-national fertility rate (births per woman)*	Global Data Lab; Area Database (v.4.2.1) (Global Data Lab, 2023a; Smits, 2016)
Control variable	Mortality	*Mortality rate (per 1000 adults)*	World Bank; World Development Indicators (World Bank, 2023b)
Control variable	Economic conditions	*Log gross national income per capita in 1000 US Dollars (2011 PPP)*	Global Data Lab; Subnational HDI (v7.0) (Global Data Lab, 2023b; Smits, 2016)
Control variable	Health	*Life expectancy at birth*	Global Data Lab; Area Database (v.4.2.1) (Global Data Lab, 2023d; Smits, 2016)
Control variable	Conflict	*Political stability and absence of violence/terrorism*	World Bank; World Development Indicators (World Bank, 2023c)
Control variable	Political conditions	*Voice and accountability*	World Bank; World Development Indicators (World Bank, 2023d)

7.3 Correlation matrix with dependent, independent, and control variable (mentioned correlations are highlighted)

Correlation	*Migration (negative population change)*	*Temperature change*	*Precipitation change*	*Mortality rates*	*Fertility rates - national*	*Fertility rates - sub-national*	*Agricultural dependency - sub-national*	*Agricultural dependency - national*	*Income*	*Life Expectancy*	*Political stability and absence of violence/terrorism*	*Voice and accountability*
Migration (negative population change	*1.00*											
Temperature change	*0.137*	*1.00*										
Precipitation change	*-0.059*	*-0.655*	*1.00*									
Mortality rates	*-0.042*	*0.120*	*-0.154*	*1.00*								
Fertility rates - national	*-0.805*	*-0.083*	*0.005*	*0.278*	*1.00*							
Fertility rates - sub-national	*-0.767*	*-0.109*	*0.091*	*-0.162*	*0.962*	*1.00*						
Agricultural dependency - sub-national	*-0.535*	*-0.043*	*0.064*	*-0.159*	*0.589*	*0.575*	*1.00*					
Agricultural dependency - national	*-0.613*	*-0.066*	*0.034*	*0.308*	*0.735*	*0.573*	*0.772*	*1.00*				
Income	*0.577*	*0.095*	*-0.054*	*-0.270*	*-0.712*	*-0.619*	*-0.797*	*-0.830*	*1.00*			
Life Expectancy	*0.354*	*-0.051*	*0.109*	*-0.876*	*-0.635*	*-0.268*	*-0.157*	*-0.540*	*0.512*	*1.00*		
Political stability and absence of violence/terrorism	*0.260*	*0.087*	*-0.109*	*-0.073*	*-0.466*	*-0.296*	*-0.237*	*-0.434*	*0.421*	*0.246*	*1.00*	
Voice and accountability	*0.193*	*0.010*	*-0.034*	*-0.037*	*-0.306*	*-0.156*	*-0.282*	*-0.328*	*0.200*	*0.171*	*0.613*	*1.00*

7.4 Assumption tests

To test if the methodological assumptions are respected or violated, several statistical tests are performed. First, to check whether the assumption of homoscedasticity holds, a Breusch-Pagan test is performed. The results in table A1 indicate that the assumption is violated and heteroscedasticity within the sample is present as the p-value is below the 0.05 threshold. The presence of heteroscedasticity means that standard deviations of the dependent variable are non-constant over different values of independent variables. Failing to correct for this leads to biased estimates of standard errors, inefficient point estimates, and an overestimation of the goodness of fit of the model. To this end, the cluster-robust standard errors included in the model are heteroscedasticity-consistent standard errors.

Breusch-Pagan test on homoscedasticity	**Outcome**
BP-coefficient	31.459
P-value	5.114e-5
Assumption respected or violated?	Violated

Table A3: Breusch-Pagan test on homoscedasticity of error terms

Secondly, to check whether the assumption of no autocorrelation holds true, a Durbin-Watson test is performed. The test, of which the results are displayed in table A2, finds that there is serial correlation between idiosyncratic error terms. This means that error terms can somewhat be predicted by the previous one. In other words, observations of a variable are similar as a function of the time lag between each other. Failing to account for this results in estimates that are no longer the Best Linear Unbiased Estimators (BLUE). While this does not lead to biased coefficients of the effects, the standard errors tend to be underestimated and p-values overestimated. Again, to this end, the cluster-robust standard errors included in the model are also serial correlation-robust standard errors.

Durbin-Watson test on autocorrelation	**Outcome**
DW-coefficient	0.68707
P-value	2.2e-16
Assumption respected or violated?	Violated

Table A2: Durbin-Watson test on autocorrelation of error terms

Thirdly, the normality of the error terms can be estimated through a Q-Q plot. The assumption of normality is tested by comparing a theoretical normal distribution of error terms with the actual distribution of the error terms. Figure A1 displays the Q-Q plot, in which can immediately be seen that the actual distribution of errors does not follow the line of the theoretical normal distribution of errors.

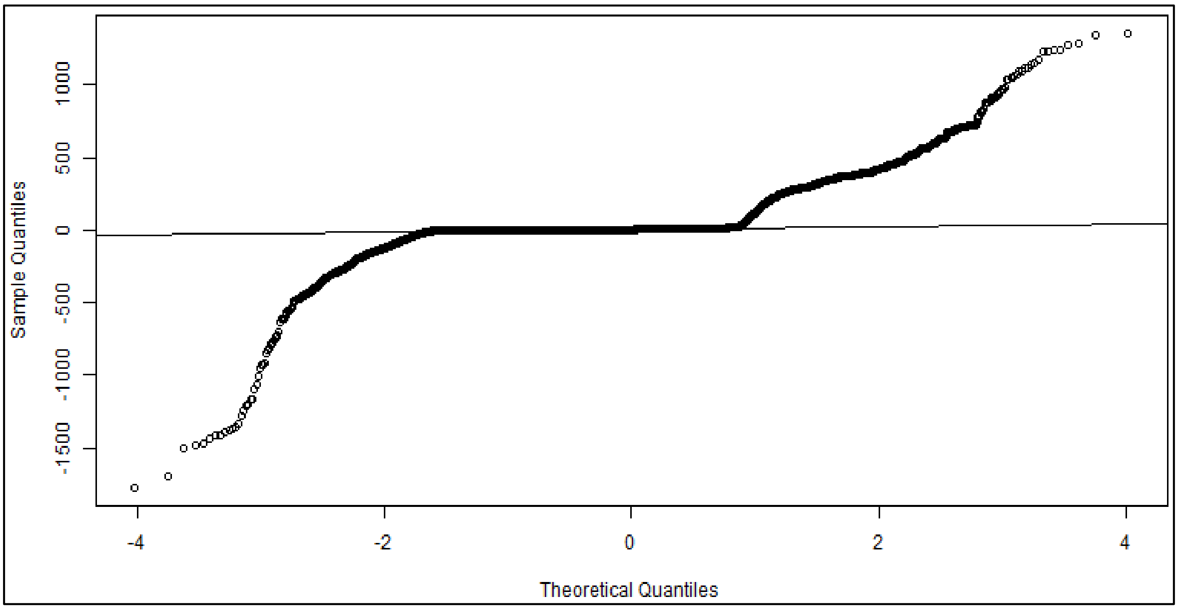


Figure A4: Normal Q-Q plot on normality of error terms

As a result, the assumption of normality of error terms is violated. However, as estimates are still rather robust under departures from normality for large samples such as to one used in this research, no altercations to the model have to be made (Schmidt & Finan, 2018).

Finally, to check whether the error terms are independently distributed, a Chow test is performed. The findings, shown in table A3, indicate that the error terms are, in fact, not distributed independently. This means that error terms are clustered at the regional level. Failing to account for this would lead to biased conclusions. Therefore, the cluster-robust standard errors included in the model are also consistent in the presence of clusters.

Chow test on independence	**Outcome**
F-statistic	23.837
P-value	2.2e-16
Assumption respected or violated?	Violated

Table A3: Chow test on independence of error terms

7.5 Test to determine whether to use pooled or region-fixed effects model

To find out whether it is preferable to simply pool the observations or include region-fixed effects, first the Hausman test is performed to determine whether a pooled model or either a random- or fixed-effects model fits the data the best. While a pooled model assumes that each region has the same intercept, random-effects allow intercepts to differ between regions. Usually, this is a strong assumption to make. Indeed, table A1 shows that a random-effects model is more efficient for the data on hand, as the p-value is below 0.05. This suggests that there is individual-specific heterogeneity that leads to heteroscedasticity, which can be captured by the random effects model.

Hausman test – Pooled vs Random-effects model	**Outcome**
Chi-square	50.43
P-value	1.189e-08
Which model is more efficient?	Random-effects model

Table A4: Hausman test – Pooled model vs Random-effects model

Now that is established that a random-effects model is preferred, another test has to be performed to determine whether to use a random- or a fixed-effects model. The results shown in table A2 indicate that the fixed-effects model is more appropriate to use than the random-effects model, as the p-value is again below the 0.05 threshold. This is because, apparently, the random-effects assumption of no correlation between the independent variables and region-specific unobserved heterogeneity is violated. A fixed-effects model eliminates this unobserved region-specific heterogeneity by including a dummy representing all regions in the model. As a result, unobserved region-specific heterogeneity no longer leads to biased results.

Hausman test – Random-effects vs fixed effects	**Outcome**
Chi-square	19.913
P-value	2.2e-16
Which model is more appropriate?	Fixed-effects model

Table A5: Hausman test –Random-effects model vs fixed-effects model

If the random-effects assumption holds, researchers often prefer the random-effects model as it allows independent variables in the model that do not change over time. However, the analysis of this research does not include variables that are stable over time, therefore, the fixed-effects model is happily adopted.