Climate and conflicts

April 3, 2024

Abstract

The climate crisis has led to increased occurrences of extreme heat and drought worldwide. In regions vulnerable to these changes, like the Horn of Africa, governmental instability, internal displacements and violent conflicts worsen the situation. However, the relation between weather, conflict, and forced displacement remains inadequately understood due to the complexities of identifying individual causal pathways. In our research, we study the role of environmental and socioeconomic determinants of conflict in Somalia between 1997 and 2022. In particular, we seek how the occurrence of conflicts in Somali regions is influenced by local climate anomalies and incoming displacements to those regions. Additionally, we investigate how climate anomalies and conflicts impact on internal displacements. Given the scarcity of displacement data, our objective is to evaluate if climate anomalies in neighboring regions can substitute for displacements. We hypothesize that conflicts are influenced not only by local climate anomalies but also by those in nearby areas. In fact, anomalies in one region can lead to increased displacement, fostering tensions in arrival regions and potentially trigger conflicts.

1 Introduction

The global rise in average temperature due to the climate crisis has led to more frequent high-temperature extremes and intensified drought conditions. The Horn of Africa is a region particularly vulnerable to these changes, exacerbated by governmental instability, civil war, and conflicts. Somalia, in particular, grapples with a political crisis, further destabilizing the region. The intricate connection between climate and conflicts is influenced by numerous factors. Causal linkages include losses in agricultural production, stock market fluctuations, migration, and internal displacements, all intricately interconnected, making it challenging to establish direct cause-effect relationships. In addition, the socioeconomic situation of the region plays a significant role, since it can mitigate the impact of weather anomalies and reduce the consequences.

A substantial body of research indicates a correlation between climate anomalies and violent conflicts on a global scale [2, 3]. However, it is essential to examine whether this relationship exists for climatic variables and local conflicts at the regional and sub-regional level. Our research studies the role of environmental and socioeconomic determinants of conflict in the Horn of Africa, particularly in Somalian regions, between the years 1997-2022. Several reasons drive the selection of our geographic area. First, that area has recently seen a concerning concurrent rise in the severity of civil conflicts and extreme weather events. Second, the region under consideration heavily depends on agriculture, with rainfed agriculture being a dominant source of livelihood for the majority of the population. As a result, this area is particularly susceptible to weather-related shocks, such as droughts and irregular rainfall patterns. These environmental challenges can profoundly affect food security, livelihoods, and overall socioeconomic stability in the region. Furthermore, this area includes economically disadvantaged and politically unstable countries. These nations often lack the resources and infrastructure to effectively mitigate and adapt to climate change impacts, which can exacerbate existing socioeconomic challenges.

Earlier research conducted in the East African region has indicated that both unusually high temperatures and periods of extreme rainfall variation can elevate the likelihood of violent conflicts [6, 7]. Likewise, drought episodes have been associated with an increased risk of conflicts [4].

We investigate a possible transmission mechanism: we hypothesize that a potential linkage between civil conflict and drought is through the consequences of internal displacements, as evidenced in the schematic graph we outline in 1. More precisely, we propose that conflicts become more prevalent

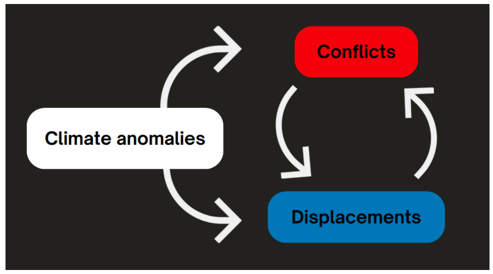

Figure 1: Multiple interaction pathways

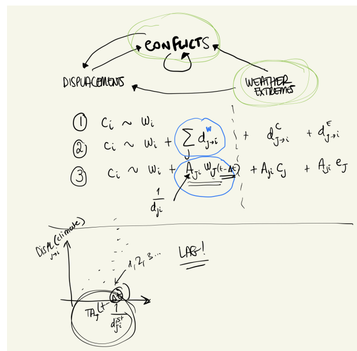

Figure 2: Elisa's graph just for us

due to drought-related displacements. This is motivated by the fact that incoming displacements can potentially heighten demands on limited resources, including essentials such as food and drinking water.

In this work, we will test the previous hypotheses, estimate the effects associated with the proposed links, forecast the potential impact of projected weather anomalies on conflicts, and ultimately draw conclusions relevant to policy-making. Our research aims to contribute to the existing literature in distinct ways. First, we aim to assess the role of environmental and socioeconomic determinants of conflict in the Horn of Africa. We investigate whether the relationship between weather anomalies and conflict holds at the local level in the case of Somalia. We analyze variations in conflict incidence at the sub-national level, an aspect often overlooked in most cross-country studies in the conflict literature. Second, we propose a potential key mechanism through which weather anomalies translate into conflicts in Somalia. To this end, we investigate whether climatic conditions, by influencing drought severity and the likelihood of armed conflict, play a significant role as explanatory factors for internal displacements. In doing so, our research extends the conflict literature by examining how the influx of large numbers of migrants to a new location can serve as incentives for individuals to participate in conflicts. Third, given the frequent difficulty in obtaining comprehensive displacement data, our goal is to explore whether climate and conflict data might serve as complementary data. We aim to examine the connection between climate-induced displacements and weather anomalies in the region where these displacements originate, coupled with the distance between the origin and destination regions. Similarly, we check whether a correlation exists between conflict-induced displacements, and the actual number of conflicts in the region of origin of the displacements. In other words, we seek to identify new variables that can effectively proxy displacement data, which is often difficult to retrieve.

Our research questions are:

- Do weather anomalies have an effect in creating new conflicts or exacerbating existing ones?
- Which weather anomalies have the greater influence in increasing the number of conflicts?
- Do weather anomalies and conflicts have an effect in increasing internal displacements?
- Do internal displacements, along with climate variables, have an effect in the occurrence of conflicts?

- What is the temporal lag between the weather anomalies and the occurrence of conflicts and between displacements and conflicts?
- Can weather anomalies of other regions serve as a proxy for displacement data?

2 Data

Our conflict data is taken from the Armed Conflict Location and Event Dataset (ACLED 2011), which records protests and violent events at a sub-national level, from year 1997. KK: **descriptive stats? more info?**

The weather-related variables are constructed from climatic data provided by the University of East Anglia Climatic Research United (UEA-CRU 2011). We employed version 4.07 of the UEA-CRU time-series datasets, which report average temperatures and total precipitation by months at data points of a high-resolution grid (of 0.5 × 0.5 degree). These datasets rely on measurements collected from weather stations distributed across the globe.

The data we used is taken from the Protection and Return Monitoring Network (PRMN) survey, collected by the United Nations High Commissioner for Refugees (UNHCR) since 2016. PRMN/ UNHCR capture this information at the regional level by conducting interviews with IDP heads of household. This is done primarily at points of arrival (i.e., destination) or by interviewing key informants at IDP settlements, transit centers, and other strategic locations and detailed information on the survey can be found in (UNHCR Somalia, 2017). Currently one main reason for each movement is recorded during PRMN interviews, and it is reported in four categories (drought, flood, conflict, and other). However, often the real driver for displacement may be combination of closely interrelated factors.

Data for GDP and population density is taken from AidData's GeoQuery.KK: **do we also have socioecon? add citation.**

3 Methods

To achieve our objectives, we examine the relationship between weather anomalies, defined in the next section, and conflicts through different regression models. We use region panel data on precipitation and temperature, which are available from the year 1901KK: **we need to keep conistency on the reported dates.** and can be used to obtain long-term averages.

In our initial analysis, the target variable is the number of violent conflict events in Somalia's administrative regions during specific month-year time periods. The explanatory variables are region- and time-specific temperature and precipitation anomalies, as well as drought length defined based on temperature anomalies.

To construct the Temperature Anomaly (TA) variable, we used monthly averages of daily maximum temperatures in order to capture the temperatures during the daytime when evaporation tends to be at its peak. In particular, we transformed the gridded temperature data into a single, centered data point for each administrative unit. We achieved this by employing Spatial Areas of Interest (AOIs) that are derived from shapefiles encompassing the boundary polygons of Somali regions. We computed the long-term averages $\mu^T_{i,m,y}$ for region i during the month-year (m, y) time period by utilizing data spanning from 1901 to 2022. These averages were used to compute the standardized difference with the maximum temperature $T_{i,m,y}$ for each region and for each month. Due to the presence of noisy data in monthly anomaly time series, which can hide patterns related to large-scale climate variability, we employed a rolling mean technique with a four-month window ($n = 4$) to mitigate this noise and facilitate the identification of climate patterns [4].

A similar approach was employed to construct the Precipitation Anomaly (PA) variable (see equations below).

$$\begin{aligned} TA^n_{i,m,y} &= \frac{1}{n}\sum_n \frac{T_{i,m,y} - \mu^T_{i,m}}{\sigma^T_{i,m}} \\ PA^n_{i,m,y} &= \frac{1}{n}\sum_n \frac{R_{i,m,y} - \mu^R_{i,m}}{\sigma^R_{i,m}} \end{aligned} \tag{1}$$

The Drought Lenght (DL) variable was introduced to account for the effects of the prolonged duration of the drought period, and is defined as the number of consecutive months with positive temperature anomalies.

To address potential issues related to omitted or unobserved variables, we control for region-fixed and time-fixed effects, which remove the impact of characteristics that remain constant over time and across regions. In our simplest version of the modelKK: **haven't talked about a model yet.**, we exclusively introduce weather-related explanatory variables while considering the number of violent conflict events as the dependent variable. To be more specific, we constructed the conflict variable by summing the count of conflict events in each region per month.

3.1 Climate and displacements effect on conflicts

The simplest model we employ has the following estimation equation:

$$Conflicts_{i,m,y} = \alpha + \beta_1 TA_{i,m,y} + \beta_2 PA_{i,m,y} + \beta_3 DL^{TA}_{i,m,y} + \psi_i + \theta_{m,y} + \epsilon_{i,m,y} \tag{2}$$

where ψ_i is the $n \times 1$ vector of region fixed effects, $\theta_{m,y}$ is the $my \times 1$ vector of time fixed effects and $\epsilon_{i,m,y}$ is an idiosyncratic error term. The total number of observations in the panel is $N \times T$, where $N = 18$ is the number of entities, i.e. regions of Somalia, and $T = 84$ is the number of months between 2016-01 and 2022-12.

Our assumption is that the omitted effects of the model can be arbitrarily correlated with the climate variables. We conducted a pFtestKK: **cite package** to determine whether the fixed effect model was a more suitable choice compared to the OLS model. The results indicated that the fixed effect model provided a better fit to the data, as it exhibited a p-value significantly lower than 0.05. Consequently, we opted for the fixed effects model following the rejection of the null hypothesis by the Hausman test KK: **cite**, which favored the random effects model over the fixed effects model.

We performed the regression using the package plm [1], the traditional R package for panel data estimation. Models are estimated using the lm() function to transformed data, i.e. the data obtained by subtracting the average over time and over region to every variable, which is usually termed (time-) demeaning. By demeaning the data, we remove across-region (time-invariant) variation so that we can observe only within-regions differences. Similarly, we remove region-invariant variation to observe within-time differences.

In this preliminary analysis, we deliberately omitted factors strongly correlated with conflicts, such as political and economic determinants. We acknowledged that this decision would result in a significantly poorer fit of the data. However, our objective was to ensure that any explained variance of the target variable (i.e. conflicts), no matter how minimal, could be attributed solely to climatic factors.

In a second model specification, we included the sum of the displacements arriving to a region as additional explanatory variable. The regression equation reads:

$$Conflicts_{i,m,y} = c + \alpha TA_{i,m,y} + \beta PA_{i,m,y} + \gamma DL^{TA}_{i,m,y} + \delta SumIDPs_{i,m,y} + \psi_i + \theta_{m,y} + \epsilon_{i,m,y} \tag{3}$$

3.2 Climate and conflict effect on displacements

Before employing the displacements variable as a regressor, we checked whether we could predict the intensity of the displacements from the climatic variables, and subsequently from the conflict variable.

We hypothesized that there could be a correlation between the number of arrivals to a region with the climatic variables of the region of origin of the displacement, weighted by the inverse of the distance squared. In fact, it is plausible to expect an increase in the number of displacements from a region with higher anomalies (in temperature or precipitation), and to have more displacements coming from nearby regions. In this section, we exclusively examined displacements that were reported to be due to drought $Disp^W_{ij}$.

The regressions were performed with different lags in the anomalies of temperature, precipitation and drought length, and with different functions KK: **what does this mean?** of the inverse distance. We included in the model a set of control variables to account for the regional effects, i.e. the population density and the GDP of the origin and destination regions.

The regression equation for this model reads:

$$
\begin{aligned}
Log(Disp^W_{ij}) = \alpha + \beta_1 * \frac{TA_{lag2,i}}{dist^2_{ij}} + \beta_2 * TA_{lag2,i} + \beta_3 * PA_{lag2,i} + \beta_4 * DL_{lag2,i} + \\
+ \beta_5 * \frac{1}{dist_{ij}} + \beta_6 * \frac{1}{dist^2_{ij}} + \beta_7 * Conflicts_i + \beta_8 * PopDensity_i + \\
+ \beta_9 * PopDensity_j + \beta_{10} * GDPmean_i + \beta_{11} * GDPmean_j
\end{aligned}
\tag{4}
$$

First, we perform the regression excluding all pairs of regions for which there are no displacements. In this case we are not trying to estimate if displacements occur but, for the displacements that do occur, we are estimating their intensity. Second, we replace the variable $Log(Disp^W_{ij})$ with $Log(Disp^W_{ij} + 1)$, in order to keep all the data points.

KK: **how are the displacements per region due to conflicts? is it again banadir the only place where people go to?**

In a similar manner, we are interested to check whether we can predict conflict-related displacements from the conflict data. In this section we exclusively examined displacements reported to be due to conflicts. The regression equation now reads:

$$
\begin{aligned}
Log(Disp^C_{ij}) = \alpha + \beta_1 * \frac{Conflicts_i}{dist^2_{ij}} + \beta_2 * TA_{lag2,i} + \beta_3 * PA_{lag2,i} + \beta_4 * DL_{lag2,i} + \\
+ \beta_5 * \frac{1}{dist_{ij}} + \beta_6 * \frac{1}{dist^2_{ij}} + \beta_7 * Conflicts_i + \beta_8 * PopDensity_i + \\
+ \beta_9 * PopDensity_j + \beta_{10} * GDPmean_i + \beta_{11} * GDPmean_j
\end{aligned}
\tag{5}
$$

3.3 Spatial lag model

The traditional fixed effects model can be extended to include a spatially lagged dependent variable among the regressors, as well as a spatially autocorrelated error term (SARAR model). The simplest representation of a linear regression model with both fixed spatial and time effects takes the following form:

$$
\begin{aligned}
Conflicts_{i,m,y} &= c + \rho W_i Conflicts_{i,m,y} + \alpha TA_{i,m,y} + \beta PA_{i,m,y} + \gamma DL^{TA}_{i,m,y} + \psi_i + \theta_{m,y} + u_{i,m,y} \\
u_{i,m,y} &= \lambda W_i u_{i,m,y} + \epsilon_{i,m,y}
\end{aligned}
\tag{6}
$$

where W_i is the $N \times N$ spatial weight matrix, ρ and λ are the unknown spatial autoregressive parameters to be estimated. We chose to implement the spatial weight matrix W_i as a normalised inverse distance matrix, calculated between the centroids of the regions. This choice is motivated by the fact that regions in closer proximity have a greater influence compared to those situated at greater distances.KK: **this should be the opening of the paragraph to justify the model choice.**

The presence of the spatial lag introduces a form of endogeneity that violates the assumption of the standard regression model that the regressors are uncorrelated to the error term. To account for this endogeneity, we employed a classical R package for the estimation of spatial panel data, namely splm [5]. This package implements Maximum Likelihood (ML) estimation for autoregressive models (numerical maximization of the log-likelihood), including both a spatially lagged dependent variable and a spatially autocorrelated error term.

In our model, we incorporate population density as a control variable. This decision is driven by the consideration that conflicts are more prone to arise in densely populated areas.

We performed a Spatial Durbin Model regression, to account for spatial lag effects of climate in other regions:

$$
C_i = \alpha + \beta X_i + \gamma W X_j + \rho W C_j + \psi_i + \theta_t + \epsilon_{i,t} \tag{7}
$$

KK: **this goes to results.**However, we did not find statistical significance for this variable, probably due to the fact that the effect of population density (that changes little over time) is accounted for in the fixed effects.

4 Results

4.1 Climate effect on conflicts

In the heatmap below, we present the results of the first model, which employed a panel data approach incorporating fixed effects and exclusively utilized climatic variables as explanatory factors. This analysis spanned seven different time lags for between climatic variables and the conflict one.

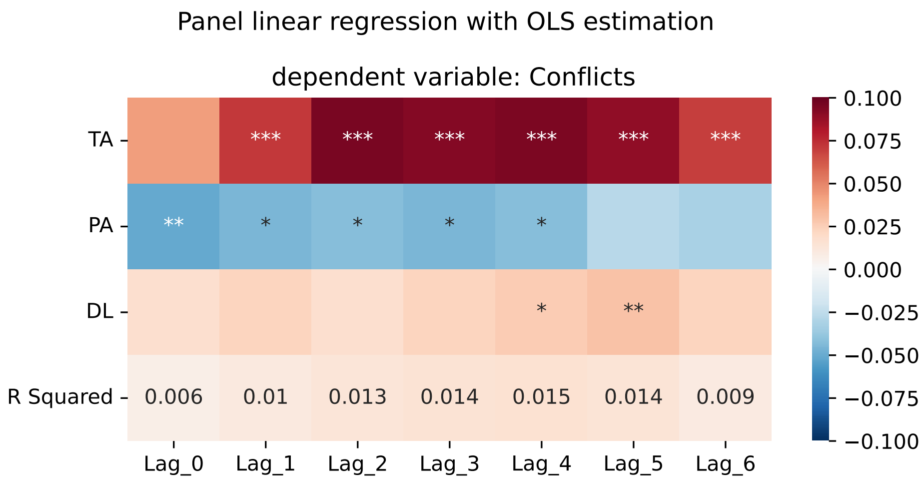

Figure 3:

We found that TA is significant for all lags with a positive coefficient, and PA for most lags with a negative coefficient, DL is significant and positively correlated to conflicts for higher lags. This suggests that an increase in TA and DL, as well as a decrease in PA, is related to an increase in the number of conflicts. These results confirms the findings of similar papers addressing the relationship between climate and conflicts in Somalia. However, given the low goodness of fit of the model, for any time-lag, we are not able to assess there is an effect of climate on conflict. It is plausible to think that the climate variables we used are not good proxies for drought and aridity of the soil, which might be the responsible factors for losses in agricultural production and the consequent tensions over resource access which we believe are channels leading to more conflicts.

In the following table, we display the outcomes of the model integrating an additional explanatory variable, namely the sum of displacements arriving to a given region, within a given month. All explanatory variables are lagged by one month with respect to the conflict variable.

Table 1:

	Dependent variable:
	conflicts
TA_lag1	0.059***
	(0.022)
PA_lag1	−0.025**
	(0.013)
DL_lag1	0.027
	(0.021)
sum_disp	0.035***
	(0.013)
Observations	1,512
R^2	0.015
Adjusted R^2	−0.058
F Statistic	5.321*** (df = 4; 1407)
Note:	*p<0.1; **p<0.05; ***p<0.01

In terms of sign and magnitude of the coefficients of the climate variables, the results of the previous model are confirmed. However, the model is again poorly fitting the data.

4.2 Analysis of displacement flows

As a preliminary analysis we explored the displacements among regions and within years, as a directed network where the nodes are the regions and the edges weights are the number of displacements. We computed the mean network degree (i.e. the average number of edges per node in the graph, which in our case corresponds to the number of individuals fleeing from their region of origin, averaged by all regions). We did so to investigate the timeseries of displacements in Somalia, to check if there was any kind of seasonal behaviour, or if there were years with an anomalously high value of displacements. We also computed the mean self-weight of the graph.

We found that there are two peaks around years 2017 and 2022, corresponding to years with severe drought. To check if there were relevant differences in the displacement time series at the subregional level with respect to the national averages, we computed the in and out degree for each subregion, finding that the behaviour of the timeseries is qualitatively confirmed at different scales (national, regional).

The relative weight of self-loops shows an increase in the ratio between internal and total displacements from 2016 to 2022, suggesting a growing preference for internal displacement within one's region of origin.

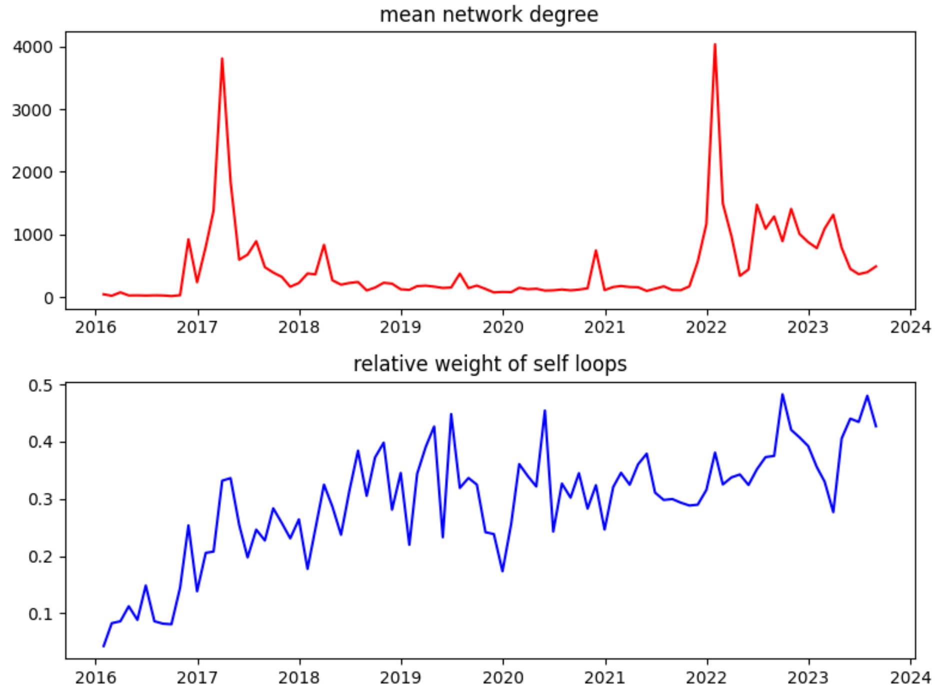

Figure 4: The mean network degree highlights two peaks of displacements in the years 2017 and 2022. The relative weight of self-loops shows an increase in the ratio between internal and total displacements.

To check if there were relevant differences in the displacement time series at the subregional level with respect to the national averages, we computed the in and out degree for each subregion, finding that the behaviour of the timeseries is qualitatively confirmed at different scales. The accumulated displacements to the district of Banadir consist of almost 25% of the totality of destinations, as evident in the following timeseries of cumulative displacements.

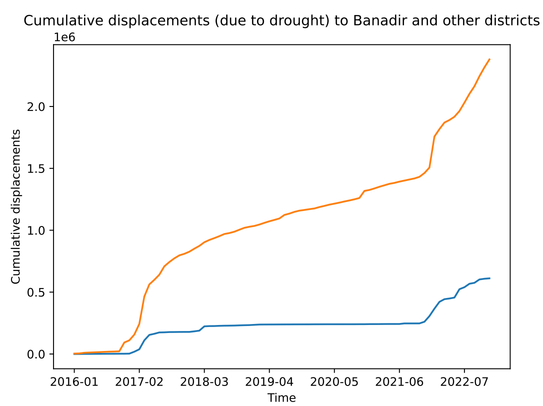

Figure 5:

We computed the difference between out degree and in degree to see which regions were a typical destination, and which were mostly the origin of the displacements, finding that region of the capital Banadir is the greatest attractor of displacements. The following plot shows a logarithmic decrease in displacements with an increase in distance between sub-regions. Additionally, in recent years there tends to be more displacements directed towards closer districts, or more internal displacements within the same district, confirming the observations derived from the variation of weight of self-loops.

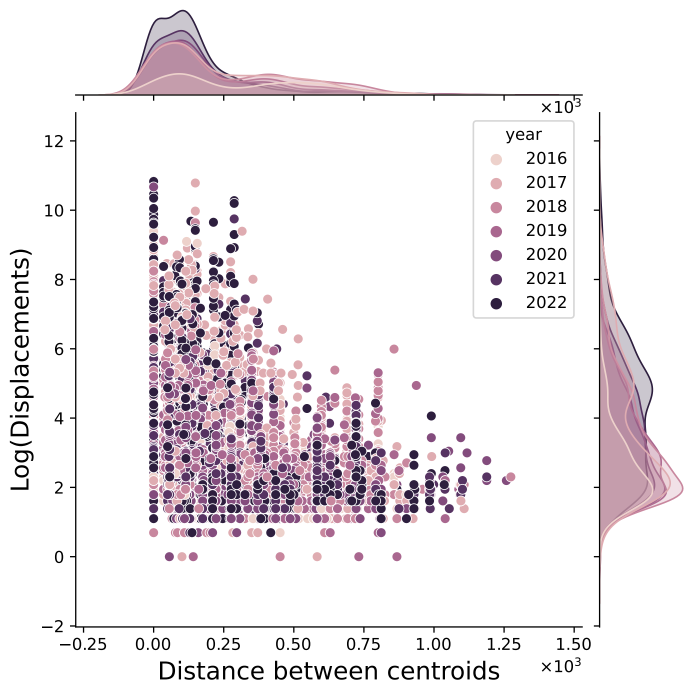

Figure 6: Displacements logarithmically decreasing with the distance between the sub-regions of origin and destination

4.3 Climate effects on displacements

We plotted the $Log(Disp_{ij}^{W})$ against $\frac{TA_{lag2}}{dist^2}$ for each region in 2022, which is the year in which there were the most displacements due to drought.

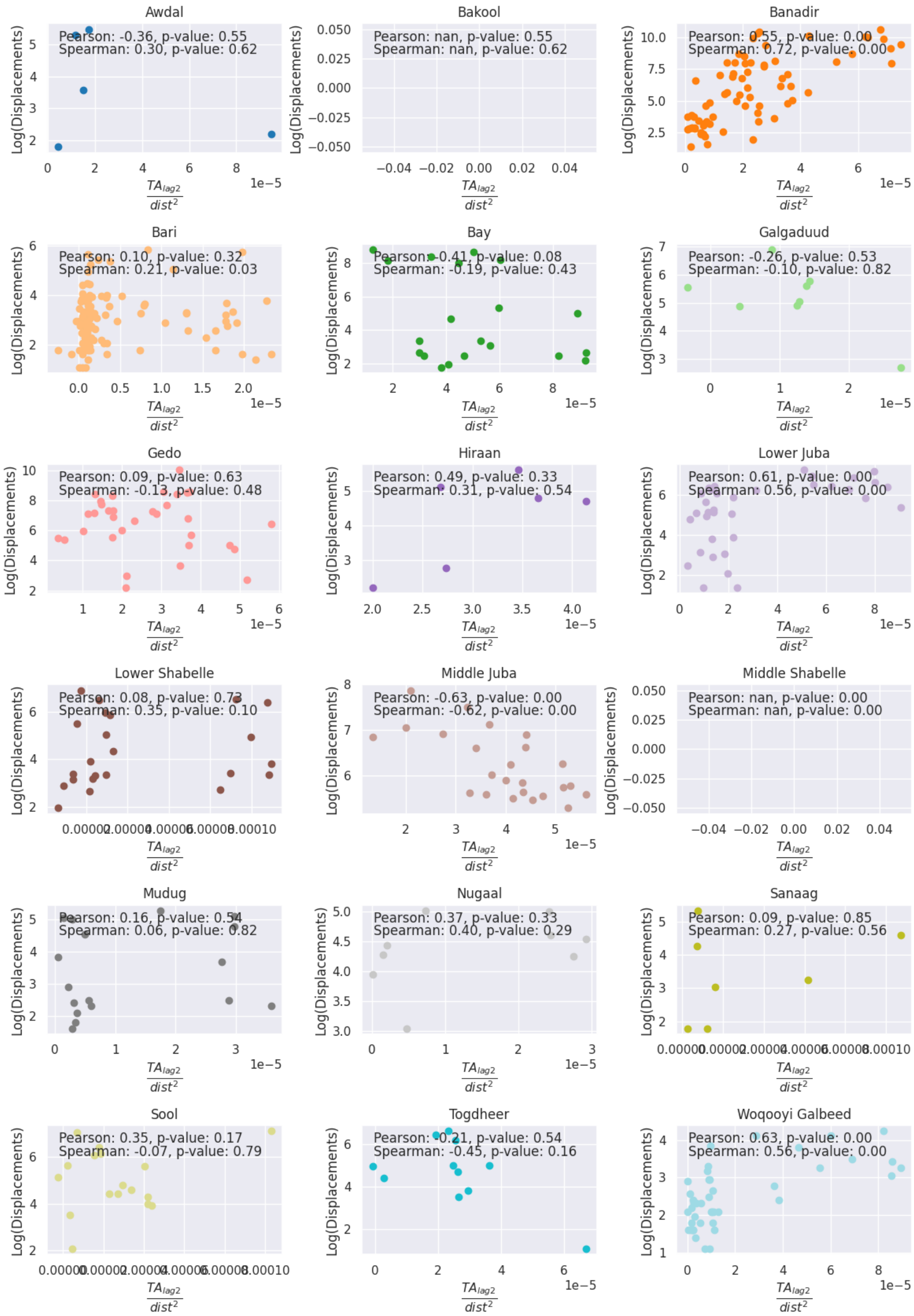

Figure 7: The region of the capital Banadir shows the expected behaviour

From the above plots, the destination region that best exhibits the expected behaviour is Banadir region, where there is also the highest number of arrivals. We computed the Pearson correlation for different lags of the TA variable in each region, finding that only few combinations of lags and regions have a significant positive correlation when considering the years from 2016 to 2022. By considering only year 2022 the positive correlation increases and becomes significant for more regions, at all lags of TA.

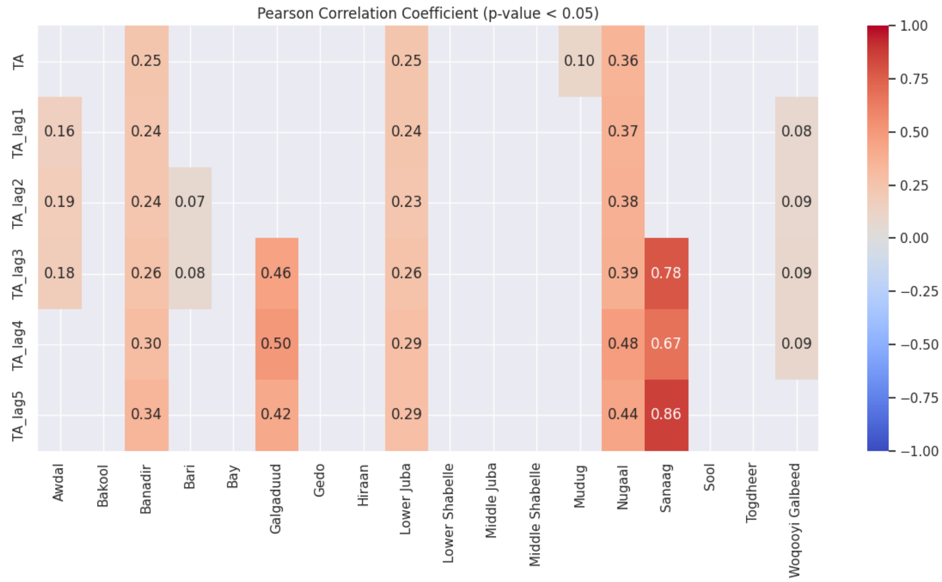

Figure 8: Only a subset of regions and lags show a significant Pearson correlation, when looking at all the years together

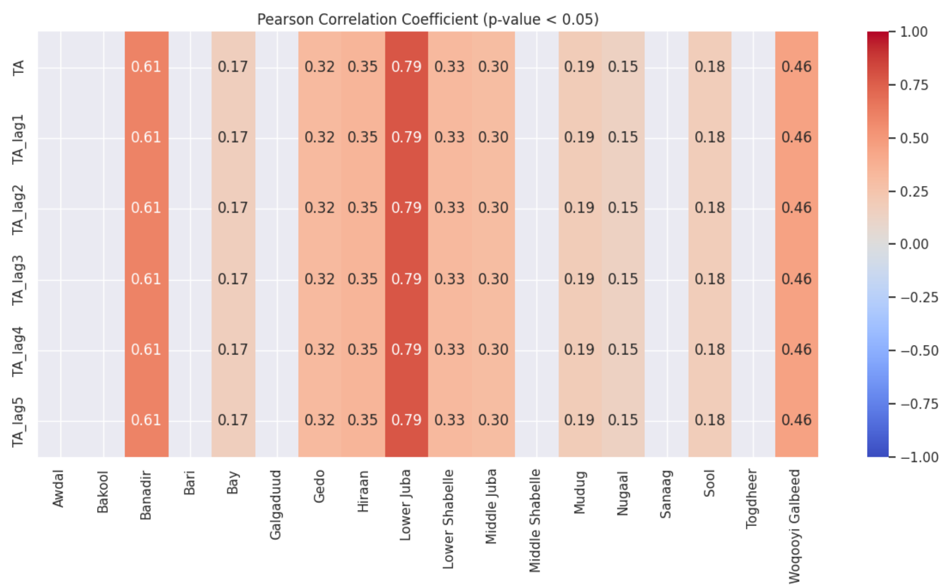

Figure 9: More regions and lags become significant when considering only 2022

We investigated the log of the displacements whose destination was Banadir region in 2022, as a function of this new variable $\frac{TA_{lag2}}{dist^2}$. When considering all years from 2016 to 2023, we get comparable results to the ones on data of 2022 only, but with lower R^2:

Table 2: KK: **somehow i'm worried we are doing something wrong. we are predicting displacements due to drought, and temperature anomalies are not explaining the phenomenon.**

	Dependent variable:		
	Log(Displacements$_{Drought}$)		
	(1)	(2)	(3)
$\frac{TA\_lag2}{distance^2}$	0.582***		0.022
	(0.034)		(0.061)
TA_lag2		0.037**	0.032
		(0.018)	(0.023)
PA_lag2		0.053**	0.053**
		(0.023)	(0.023)
DL_lag2		−0.009	−0.009
		(0.010)	(0.010)
inv_distance^2		0.233***	0.229***
		(0.010)	(0.016)
conflicts	0.028	0.046**	0.046**
	(0.021)	(0.021)	(0.021)
gdp_mean_origin	0.007	0.136	0.136
	(0.135)	(0.132)	(0.132)
gdp_mean_destination	−0.467***	−0.286**	−0.287**
	(0.117)	(0.114)	(0.114)
population_density_origin	−0.022	−0.164	−0.165
	(0.158)	(0.154)	(0.154)
population_density_destination	0.619***	0.379***	0.379***
	(0.127)	(0.124)	(0.124)
Constant	0.210***	0.227***	0.226***
	(0.006)	(0.011)	(0.012)
Observations	3,671	3,671	3,671
R^2	0.140	0.195	0.195
Adjusted R^2	0.139	0.193	0.193
Residual Std. Error	0.140 (df = 3664)	0.136 (df = 3661)	0.136 (df = 3660)
F Statistic	99.644*** (df = 6; 3664)	98.361*** (df = 9; 3661)	88.516*** (df = 10; 3660)
Note:			*p<0.1; **p<0.05; ***p<0.01

We compared different models that included only the interaction term $\frac{TA\_lag2}{distance^2}$ (first column in the table above), then only the individual terms (second column), and finally both interaction and individual terms (third column). We noticed that the interaction term was not bringing a clearer signal (better fit of the data) with respect to the individual terms, so we dropped it in the subsequent analysis.

Instead, we hypotetized that climate and conflict variables of both the origin and the destination regions could be responsible for variations in the number of displacements. We investigated the relationship between the logarithm of the number of displacements due to drought, and climatic and conflict variables in both departure and arrival of regions of displacements. Employing a linear regression approach, we incorporated control variables such as population density and GDP in both departure and arrival regions to address regional fixed effects. As done in previous models, this analysis spanned seven different time lags for both climatic and conflict variables. We show the results in the heatmap below, where on the y axis there are the independent variables, on the x axis the lags, and in the cells is reported the level of significance of each variable, whereas the colorbar maps to the coefficients of the respective variables.

The displacements considered in this analysis are attributed to drought only (self reported by the displaced people at the arrival region), encompassing all years from 2016 to 2022, totaling 3630 observations. The coefficients of the variables exhibit similar patterns across lags, suggesting that the results are stable.

Specifically, the temperature anomaly TA maintains significance across all lags: its coefficient is positive in the departure region, increasing with lag, and negative in the arrival region, decreasing with the lag. However, the interpretation of the results for Drought Length is less straightforward. It appears that people are moving from regions with shorter drought periods to those with longer ones. Nonetheless, this variable is considered less reliable, as it is computed as the cumulative of days with positive anomaly temperature.

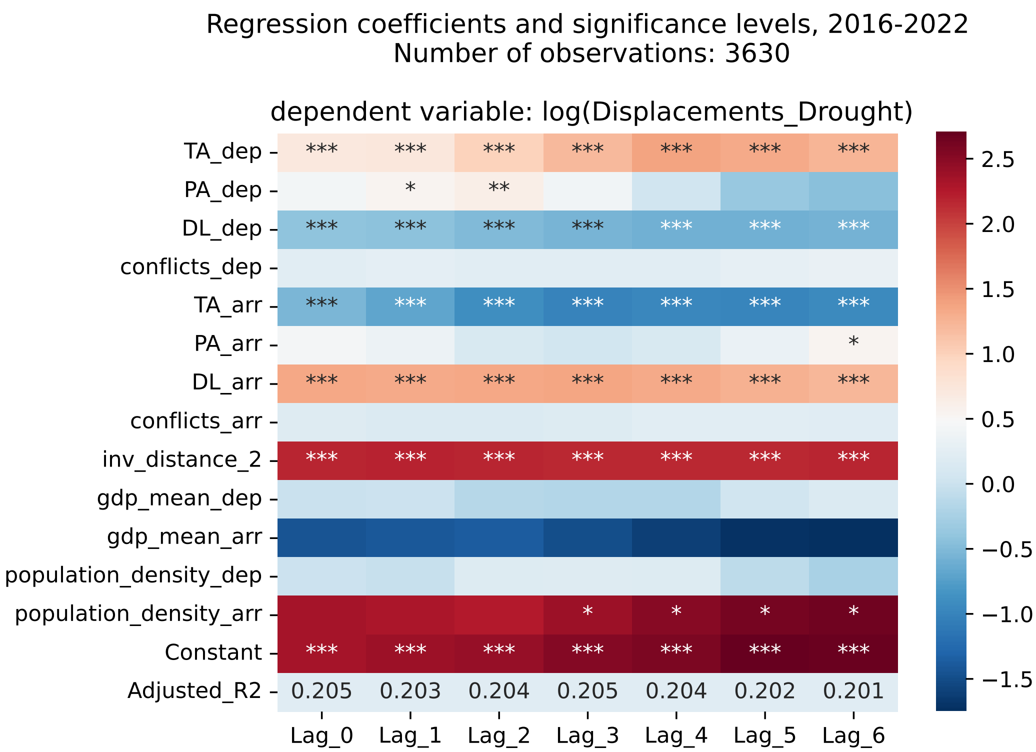

Figure 10: Heatmap of regression coefficients and significance levels **KK: so basically the time lag doesn't have much of an effect.**it seems to be more correlated with 4m window of TA.

We decided to focus on the model with a time lag of 2 months, and we substituted the inverse distance squared regressor with the simple distance. In addition, we substituted the control variable of population density, with the population count. In the left column, the Ordinary Least Squares (OLS) regression with control variables is displayed, while the right column shows the panel linear regression with fixed effects applied to both the origin region and time.

Table 3: Drought displacements between regions, OLS and plm

	Dependent variable:	
	Log(Drought_Displacements)	
	OLS	*panel linear*
	(1)	(2)
TA_lag2_dep	0.579**	−0.024
	(0.256)	(0.362)
PA_lag2_dep	0.756***	−0.238
	(0.287)	(0.315)
DL_lag2_dep	−0.140	−0.069
	(0.114)	(0.185)
conflicts_lag2_dep	0.114	0.311
	(0.228)	(0.311)
TA_lag2_arr	−0.332	0.969***
	(0.214)	(0.255)
PA_lag2_arr	−0.228	0.678**
	(0.302)	(0.312)
DL_lag2_arr	1.072***	0.630***
	(0.116)	(0.126)
conflicts_lag2_arr	0.617***	0.238
	(0.218)	(0.219)
distance	−2.672***	−2.593***
	(0.124)	(0.129)
gdp_mean_dep	−0.138	
	(0.129)	
gdp_mean_arr	1.001***	1.120***
	(0.146)	(0.141)
pop_count_dep	0.852***	
	(0.135)	
pop_count_arr	−1.052***	−0.799***
	(0.111)	(0.109)
Constant	−7.391***	
	(0.159)	
Observations	3,630	3,630
R^2	0.250	0.206
Adjusted R^2	0.247	0.181
Residual Std. Error	1.474 (df = 3616)	
F Statistic	92.582*** (df = 13; 3616)	83.134*** (df = 11; 3518)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 4: Conflict displacements between regions, OLS and plm

	Dependent variable:	
	Log(Conflict_Displacements)	
	OLS	*panel linear*
	(1)	(2)
TA_lag2_dep	0.109	0.798
	(0.537)	(0.864)
PA_lag2_dep	0.223	0.473
	(0.496)	(0.575)
DL_lag2_dep	0.496^{**}	0.456
	(0.206)	(0.376)
conflicts_lag2_dep	0.743^{**}	-0.626
	(0.376)	(0.512)
TA_lag2_arr	-0.900^{**}	-0.051
	(0.447)	(0.576)
PA_lag2_arr	0.565	0.702
	(0.589)	(0.622)
DL_lag2_arr	-0.415^{*}	-0.621^{**}
	(0.229)	(0.256)
conflicts_lag2_arr	1.559^{***}	1.368^{***}
	(0.349)	(0.416)
distance	-3.526^{***}	-3.282^{***}
	(0.228)	(0.252)
gdp_mean_dep	-0.182	
	(0.208)	
gdp_mean_arr	0.827^{***}	1.454^{***}
	(0.222)	(0.248)
pop_count_dep	1.602^{***}	
	(0.261)	
pop_count_arr	-1.345^{***}	-1.875^{***}
	(0.264)	(0.296)
Constant	-6.844^{***}	
	(0.275)	
Observations	1,199	1,199
R^2	0.390	0.303
Adjusted R^2	0.383	0.231
Residual Std. Error	1.520 (df = 1185)	
F Statistic	58.169^{***} (df = 13; 1185)	42.859^{***} (df = 11; 1087)
Note:		*p<0.1; **p<0.05; ***p<0.01

4.3.1 District level

The tables below show the regression analysis done at the district level. For what concerns the drought-displacement regression, all weather variables exhibit significance, with positive coefficients observed in the departure region and negative coefficients in the arrival region for Temperature Anomaly (TA) and Precipitation Anomaly (PA), whereas Drought Length (DL) has a positive coefficient. This suggests a trend of displacement from regions experiencing high temperature and precipitation anomalies to those with lesser ones. While the positive coefficient for TA aligns with expectations, interpreting PA is challenging, as both high and low rainfall anomalies could trigger displacements. Additionally, there is signal for conflicts in both departure and arrival districts, both with positive coefficients, likely attributed to spatial autocorrelation. However, in the fixed effects regression, PA and conflict variables are no longer significant.

In the second regression analysis focusing on Conflict-displacements, climate variables lose significance, while conflict variables in both the source and destination districts exhibit increased significance. Interestingly, in the fixed-effects regression, the conflict variable in the source region loses significance, suggesting that fixed effects may mask certain signals.

In all regressions, as expected, the variable of distance between districts is significant and negatively correlated with displacements, indicating that longer distances deter migration. The control variables of population and gdp show that people tend to go from more to less populated districts, and from poorer to richer districts.

By comparing the results obtained at regional level with the ones at district level, we notice that some variables become statistically significant (climate anomalies of the departure and arrival region in the regressions with target drought-displacements), while others have a higher level of significance (conflicts in the departure region and climate of the arrival region in the regression with target Conflict-displacements, again referring to the OLS model without fixed effects). The R^2 decreases slightly from regional to district-level regressions, indicating a decrease in explanatory power. Temporal lag is set to 2 months in both the source and destination areas, reflecting the assumption that variables influencing displacement choices exhibit the same lag. However, slight variations in lag do not significantly alter the results.

Table 5: Drought displacements between districts, OLS and plm

	Dependent variable:	
	Log(Drought_Displacements)	
	OLS	*panel linear*
	(1)	(2)
TA_lag2_dep	0.828***	0.994***
	(0.212)	(0.244)
PA_lag2_dep	1.722***	0.247
	(0.264)	(0.274)
DL_lag2_dep	0.474***	0.426***
	(0.086)	(0.129)
conflicts_lag2_dep	0.746**	−0.055
	(0.309)	(0.376)
TA_lag2_arr	−1.209***	−1.976***
	(0.192)	(0.217)
PA_lag2_arr	−0.845***	−0.358
	(0.254)	(0.259)
DL_lag2_arr	0.459***	0.447***
	(0.090)	(0.092)
conflicts_lag2_arr	0.613**	0.409
	(0.258)	(0.256)
dist_centroids	−3.670***	−3.495***
	(0.101)	(0.105)
pop_count_dep	0.421**	
	(0.185)	
pop_count_arr	−1.271***	−0.918***
	(0.093)	(0.094)
gdp_mean_dep	−0.075	
	(0.229)	
gdp_mean_arr	1.538***	1.028***
	(0.174)	(0.171)
Constant	−7.154***	
	(0.088)	
Observations	9,461	9,461
R^2	0.240	0.185
Adjusted R^2	0.238	0.171
Residual Std. Error	1.517 (df = 9447)	
F Statistic	228.899*** (df = 13; 9447)	191.531*** (df = 11; 9302)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 7: Conflict displacements between districts, OLS and plm

	Dependent variable:	
	Log(Conflict_Displacements)	
	OLS	*panel linear*
	(1)	(2)
TA_lag2_dep	0.209	0.421
	(0.308)	(0.445)
PA_lag2_dep	0.225	−0.068
	(0.369)	(0.397)
DL_lag2_dep	0.180	0.028
	(0.133)	(0.254)
conflicts_lag2_dep	2.693***	−0.597
	(0.432)	(0.536)
TA_lag2_arr	−1.256***	−0.580*
	(0.309)	(0.350)
PA_lag2_arr	0.102	−0.175
	(0.432)	(0.416)
DL_lag2_arr	−0.495***	−0.671***
	(0.161)	(0.173)
conflicts_lag2_arr	1.780***	0.732**
	(0.305)	(0.328)
dist_centroids	−3.486***	−3.457***
	(0.171)	(0.175)
pop_count_dep	1.393***	
	(0.484)	
pop_count_arr	−1.195***	−1.371***
	(0.265)	(0.292)
gdp_mean_dep	−1.786***	
	(0.453)	
gdp_mean_arr	0.958***	1.398***
	(0.290)	(0.311)
Constant	−6.428***	
	(0.187)	
Observations	2,964	2,964
R^2	0.255	0.176
Adjusted R^2	0.252	0.128
Residual Std. Error	1.583 (df = 2950)	
F Statistic	77.763*** (df = 13; 2950)	54.384*** (df = 11; 2800)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6: Comparison of OLS model and models with different fixed effects specifications: individual, time and twoways

	Dependent variable:			
	Displacements_log			
	OLS		*panel linear*	
	(1)	(2)	(3)	
TA_lag2_dep	0.209	0.102	0.086	
	(0.308)	(0.359)	(0.347)	
PA_lag2_dep	0.225	0.217	−0.138	
	(0.369)	(0.355)	(0.411)	
DL_lag2_dep	0.180	0.658***	0.161	
	(0.133)	(0.229)	(0.135)	
conflicts_lag2_dep	2.693***	0.182	2.504***	
	(0.432)	(0.522)	(0.439)	
TA_lag2_arr	−1.256***	−0.840***	−1.717***	
	(0.309)	(0.322)	(0.348)	
PA_lag2_arr	0.102	−0.047	0.054	
	(0.432)	(0.410)	(0.442)	
DL_lag2_arr	−0.495***	−0.461***	−0.470***	
	(0.161)	(0.167)	(0.166)	
conflicts_lag2_arr	1.780***	1.692***	1.355***	
	(0.305)	(0.291)	(0.340)	
dist_centroids	−3.486***	−3.300***	−3.540***	
	(0.171)	(0.172)	(0.174)	
pop_count_dep	1.393***		1.315***	
	(0.484)		(0.487)	
pop_count_arr	−1.195***	−1.205***	−1.373***	
	(0.265)	(0.290)	(0.267)	
gdp_mean_dep	−1.786***		−1.667***	
	(0.453)		(0.454)	
gdp_mean_arr	0.958***	0.785***	1.311***	
	(0.290)	(0.302)	(0.303)	
Constant	−6.428***			
	(0.187)			
Observations	2,964	2,964	2,964	
R^2	0.255	0.178	0.254	
Adjusted R^2	0.252	0.155	0.229	
Residual Std. Error	1.583 (df = 2950)			
F Statistic	77.763*** (df = 13; 2950)	56.697*** (df = 11; 2883)	75.082*** (df = 13; 2867)	54.384***
Note:				*p<0.1; **p<

4.4 Climate and displacement effect on conflict at district level

Table 8: At the district level, without fixed effects (left column) climate variables are not significant, and most of the variance is explained by the population count and GDP variables. This is because most conflicts happen in densely populated districts, like large cities, where also GDP is higher. When accounting for this through fixed effects (for individual and time, in right column), the climate variables TA and DL become significant, but the R^2 drops to almost 0.

	Dependent variable:	
	conflicts	
	OLS	*panel linear*
	(1)	(2)
TA_lag1	0.005	0.019***
	(0.005)	(0.005)
PA_lag1	0.006	−0.007
	(0.007)	(0.005)
DL_lag1	0.002	0.013***
	(0.002)	(0.003)
sum_disp_lag1	0.128***	0.047***
	(0.021)	(0.015)
gdp_mean	0.452***	
	(0.008)	
pop_count	0.106***	
	(0.006)	
Constant	0.008**	
	(0.003)	
Observations	6,216	6,216
R^2	0.659	0.008
Adjusted R^2	0.659	−0.019
Residual Std. Error	0.047 (df = 6209)	
F Statistic	2,003.847*** (df = 6; 6209)	11.444*** (df = 4; 6055)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 9: The explained variance in the gravity model regression without fixed effects is again mainly given by population count and GDP variables. Climate and conflict variables, although significant, do not explain almost any variance.

	Dependent variable:	
	conflicts_log	
	OLS	*panel linear*
	(1)	(2)
TA_lag1_log	−0.016	0.170***
	(0.069)	(0.066)
PA_lag1_log	0.004	−0.117*
	(0.081)	(0.068)
DL_lag1_log	0.105***	0.039***
	(0.012)	(0.014)
sum_disp_lag1_log	0.096***	0.031**
	(0.015)	(0.012)
pop_count_log	0.492***	
	(0.025)	
gdp_mean_log	0.291***	
	(0.022)	
Constant	−0.985***	
	(0.162)	
Observations	6,216	6,216
R^2	0.241	0.006
Adjusted R^2	0.241	−0.020
Residual Std. Error	1.629 (df = 6209)	
F Statistic	329.022*** (df = 6; 6209)	9.758*** (df = 4; 6055)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 10: By logging the conflict as well as the control variables, the climate and incoming displacements become significant, but again they don't explain almost any variance.

	Dependent variable:	
	conflicts_log	
	OLS	*panel linear*
	(1)	(2)
TA_lag1	0.738***	0.990***
	(0.162)	(0.196)
PA_lag1	0.530**	−0.181
	(0.236)	(0.197)
DL_lag1	0.461***	0.567***
	(0.085)	(0.118)
sum_disp_lag1	4.297***	0.802
	(0.718)	(0.555)
gdp_mean_log	0.278***	
	(0.022)	
pop_count_log	0.528***	
	(0.025)	
Constant	−2.549***	
	(0.141)	
Observations	6,216	6,216
R^2	0.236	0.009
Adjusted R^2	0.235	−0.017
Residual Std. Error	1.635 (df = 6209)	
F Statistic	318.837*** (df = 6; 6209)	14.433*** (df = 4; 6055)
Note:		*p<0.1; **p<0.05; ***p<0.01

4.5 Conflict effects on displacements

In terms of regressions focused on displacements resulting from conflicts, we observe a similar signal for temperature anomalies with respect to the regressions with drought displacements as target. Additionally, there is signal for an effect of conflicts in the departure region, with a coefficient approximately equal to 1.4. Surprisingly, in a similar signal is there for conflicts in the arrival region, with a coefficient of around 0.9, contrary to the expected negative correlation.

Furthermore, the control variable for GDP in the departure region shows significance, with a coefficient of approximately -6, as well as the population density. This suggests that individuals tend to migrate from countries with lower GDP and high population density.

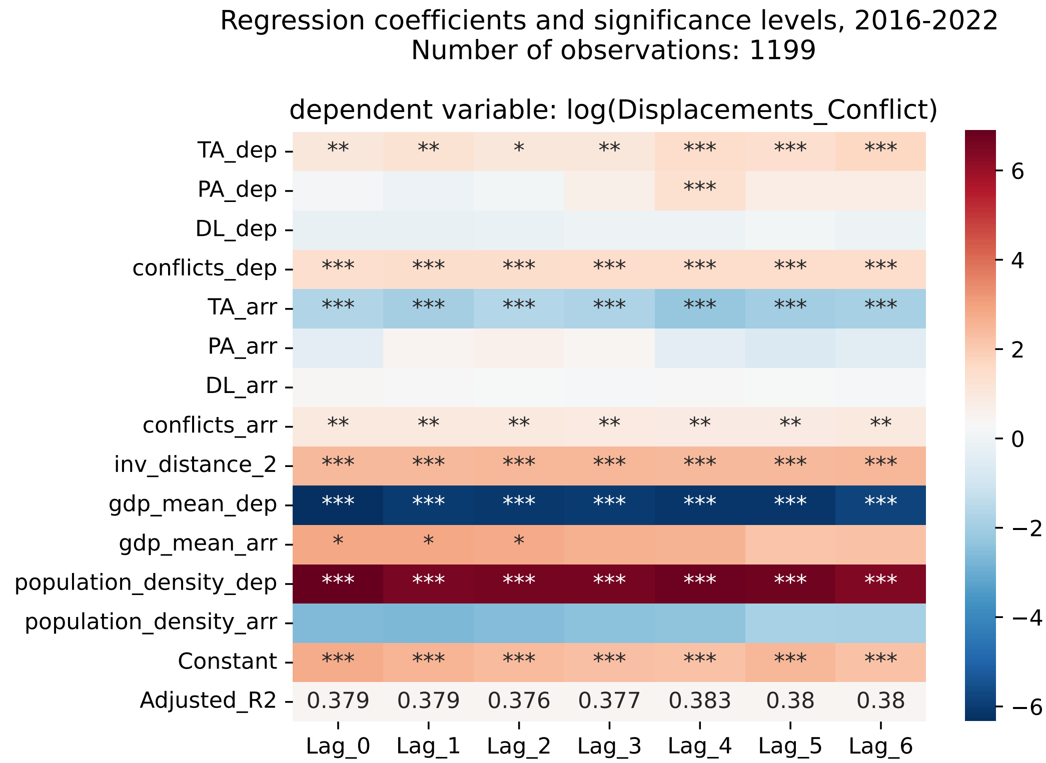

Figure 11: Heatmap of regression coefficients and significance levels. Target variable: Logarithm of the displacements due to conflicts.

See Supplementary.tex for other results, with different lags and structures of the main variable.

4.6 Climate and displacements effect on conflicts

The first model specification shows different levels of significance for the climate variables. Since for the spatial autoregressive models there is no direct equivalent to the OLS R-squared (these models are fitted by maximum likelihood) we compare them through the AIC test. With a 4-months moving average window we get:

Table 11: Regression results for different models. The climate variables are calculated with a four months moving average (n=4) and a time lag of 2 months wrt conflicts. All variables are normalised. The target variable is Log(conflicts)

	n=4				
	treat	sar	sarar	sar+disp	sarar+disp
TA_{lag2}	1.07734***	1.111402***	1.110527***	1.119684***	1.119057***
	(0.19550)	(0.188743)	(0.184342)	(0.188757)	(0.184088)
PA_{lag2}	−0.21857	−0.215161	−0.202428	−0.227648	−0.214859
	(0.17686)	(0.170744)	(0.168005)	(0.170914)	(0.168020)
DL_{lag2}	0.17093.	0.175333.	0.181892.	0.161562	0.167854.
	(0.10353)	(0.099950)	(0.099730)	(0.100392)	(0.100175)
$SumDisp$				0.385925	0.398129
				(0.272400)	(0.274032)
ρ		−0.093547	−0.026383	−0.095070	−0.024156
		(0.076016)	(0.099921)	(0.076058)	(0.099732)
λ			−0.100777		−0.106852
			(0.104956)		(0.105338)
AIC		-368.467	-367.51	-368.476	-367.626
N	1512	1512	1512	1512	1512

Signif. codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '

Table 12: Regression results for different models. The climate variables are calculated with a four months moving average (n=4) and a time lag of 2 months wrt conflicts. All variables are normalised. The target variable is Conflicts

	n=4				
	treat	sar	sarar	sar+disp	sarar+disp
TA_{lag2}	0.095080***	0.104497***	0.104173***	0.106750***	0.106514***
	(0.026418)	(0.025344)	(0.024201)	(0.025285)	(0.024064)
PA_{lag2}	−0.042341.	−0.044002.	−0.042428.	−0.047593*	−0.046215*
	(0.023899)	(0.022927)	(0.022209)	(0.022895)	(0.022130)
DL_{lag2}	0.016535	0.018547	0.020499	0.014602	0.016429
	(0.013990)	(0.013421)	(0.013362)	(0.013448)	(0.013388)
$SumDisp$				0.110251**	0.113983**
				(0.036490)	(0.036921)
ρ		−0.271***	−0.136749	−0.286874***	−0.129693
		(0.081)	(0.112933)	(0.081483)	(0.112037)
λ			−0.187		−0.212394.
			(0.116)		(0.117412)
F-test					
AIC	10367.31	-6432.34	-6433.61	-6439.47	-6441.15
N	1512	1512	1512	1512	1512

Signif. codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '

4.7 Spatial models

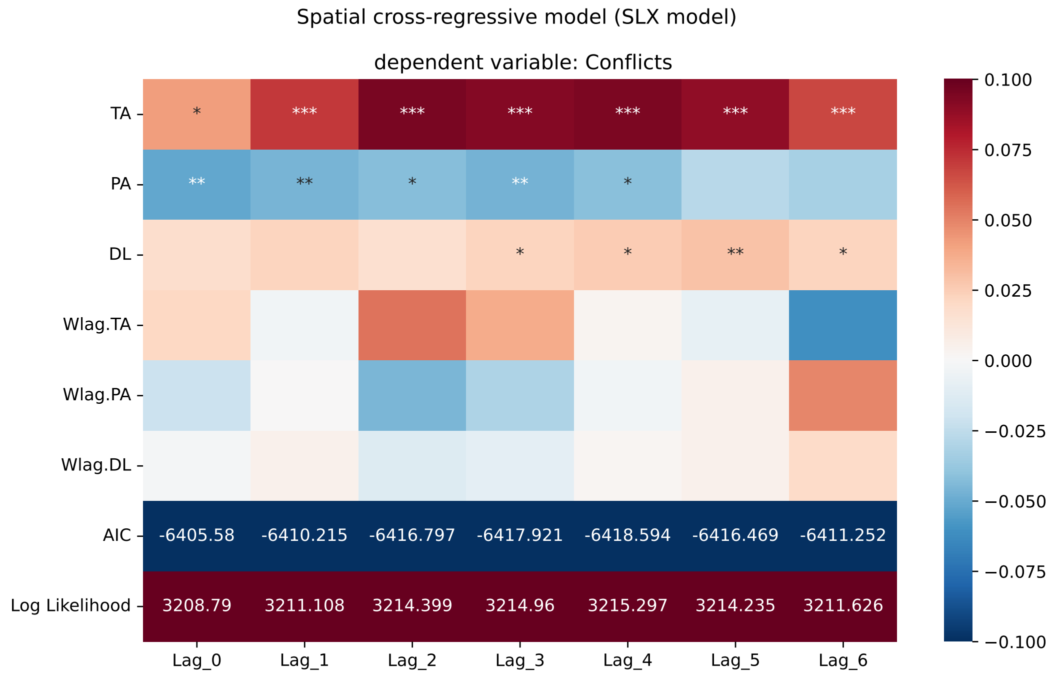

Figure 12: Standardised variables

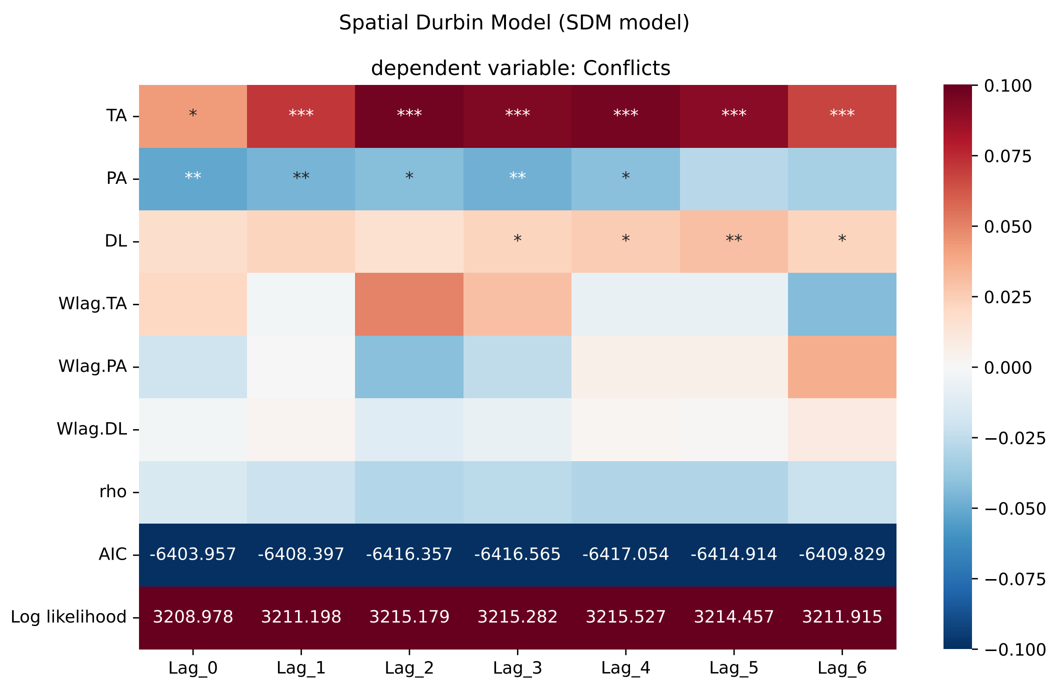

Figure 13:

5 differences with the paper "Large weather and conflict effects on internal displacement in Somalia with little evidence of feedback onto conflict"

They use data from 2016 to 2018. For the mean temperature anomalies they use the Berkeley Earth dataset: each daily time series is transformed into a series of temperature anomalies by subtracting the corresponding monthly average (with respect to 1951–1980) from each daily observation at the same station. For the precipitation observations data is taken from the Climate Hazards Groups Infrared Precipitation with Stations dataset (CHIRPS). They allow for non linearities, and find strong non linear effects.KK: **makes sense.** First, to quantify the net effect of weather shocks on IDPs, they model the total displacement out of each region as a function of weather conditions at the origin. Second, they identify the sole-extreme weather effect by estimating the same model using the subset of drought-reported displacement. This identifies the channel "weather event-displacement" by only focusing on IDPs reporting drought to be the reason for displacement. Third, they identify the effect of conflict on displacement by modelling displacement (total and conflict-reported) as a function of conflict at the origin. Fourth, they isolate the effect of displacement on conflict by modelling conflict at the destination as a function of in-migration of IDPs with and without controlling for weather. They also discuss potential endogeneity in their conflict models due to spillovers (neighboring conflict may appear to be caused by IDPs) and reverse causality (IDPs potentially choosing low conflict areas).

They find:

significant non-linear effects of temperatures and precipitation on displacement.

little indication of a significant effect of arriving displacement on the occurrence of conflict events at the destination

no statistically significant effects of conflict on displacement (due to drough+conflict+other). Once they isolate the subset of IDPs who self-report to be displaced due to conflict, they find large and significant effects

6 Conclusion

References

[1] Y. Croissant and G. Millo. Panel data econometrics in r: The plm package. *Journal of Statistical Software*, 27(2):1–43, 2008.

[2] S. Hsiang, K. Meng, and M. Cane. Civil conflicts are associated with the global climate. *Nature*, 476:438–41, 08 2011.

[3] S. M. Hsiang, M. Burke, and E. Miguel. Quantifying the influence of climate on human conflict. *Science*, 341(6151):1235367, 2013.

[4] J.-F. Maystadt and O. Ecker. Extreme weather and civil war: Does drought fuel conflict in somalia through livestock price shocks? *American Journal of Agricultural Economics*, 96:1157–1182, 03 2014.

[5] G. Millo and G. Piras. splm: Spatial panel data models in r. *Journal of Statistical Software*, 047(i01), 2012.

[6] J. O'Loughlin, F. D. W. Witmer, A. M. Linke, A. Laing, A. Gettelman, and J. Dudhia. Climate variability and conflict risk in east africa, 1990–2009. *Proceedings of the National Academy of Sciences*, 109(45):18344–18349, 2012.

[7] C. Raleigh and D. Kniveton. Come rain or shine: An analysis of conflict and climate variability in east africa. *Journal of Peace Research*, 49:51–64, 01 2012.