A Climate of War or Peace? The Effect of Droughts on Conflict Dynamics
Online Appendix

Paola Vesco, Uppsala University
paola.vesco@pcr.uu.se

April 6, 2021

1 Data

This section delves deeper into some of the data used for the analysis, and especially the construction of the agricultural indicators.

1.1 Agricultural Indicators

The models include a set of variables to account for crop harvests and failures, and to measure the effect of climate variability on crop production. Specifically, I include binary variables signaling if a mild, moderate or severe drought occurred during the main crops' growing season. To this end, I first identify the main crops by using data on crops' harvested area and production from Mapspam (International Food Policy Research Institute 2019) for the years 2000, 2005 and 2010 and fill in the missing data through multiple imputation[1] to ensure consistency with the other ViEWS datasources.

I identify the three main crops as those which cover the highest share of harvested area in each grid-cell. Second, I match each crop with its growing season calendar drawn from MIRCA (Portmann, Siebert, and Döll 2010)[2] and I code a variable to indicate whether each month falls into at least one main crops' growing season. I then integrate the monthly information about the main crops' growing season with the drought variables based on SPEI.

I interact information on drought, proxied by SPEI, with data on the main crops' growing season, to construct a measure of agricultural drought. This is obtained by considering the value of SPEI only for the months during which the growing season of the main crops is ongoing. Using this indicator, I also construct a set of measures to account for each grid-cell's adaptive capacity. Locations that have experienced a high climate variability in the past are arguably more able to adapt to shocks than locations in which climate conditions has been relatively stable. Accordingly, I include indicators to measure the average variation of agricultural drought in the previous ten years compared to the current value, as well as the count number of past years (up to 5) for which the average value of the SPEI during the growing season was lower than the median for the reference period (1990-2010). The operationalization of these and the other indicators, as well as the data sources, are listed in Table A.I.

[1] Multiple imputation (through the Amelia II package in R) applies bootstrapping and the expectation-maximization (EM) algorithm to impute each missing value (m) times to create m complete datasets (for more information about how ViEWS handles missing data, please refer to Hegre et al. 2019).

[2] Compared to other growing season calendars (e.g. SAGE), MIRCA has a more extensive collection of crops (26), it is available globally and has a very extensive spatial coverage. Moreover, MIRCA classifies crops according to the cultivation technology, enabling me to include indications of whether the main crops are irrigated or rainfed (or both).

Table A.I.

Features included in the forecasting model.

Short name	**Description**
acled_dummy_ns	Dummy encoding of acled_count_ns being greater or equal to 1. Source: ACLED (Raleigh et al. 2010).
acled_dummy_os	Dummy encoding of acled_count_os being greater or equal to 1. Source: ACLED (Raleigh et al. 2010).
acled_dummy_pr	Dummy encoding of acled_count_pr being greater or equal to 1. Source: ACLED (Raleigh et al. 2010).
acled_dummy_sb	Dummy encoding of acled_count_sb being greater or equal to 1. Source: ACLED (Raleigh et al. 2010).
count_moder_drought_prev10	Count of the months experiencing a moderate drought in the previous 10 years. Author's computation.
tlag_12_crop_sum	Total harvests of the main crops cultivated for each priogrid-cell. Source: Mapspam (International Food Policy Research Institute 2019). Lagged by 12 months.
cropprop	Proportion of the year for which the growing season is ongoing in that grid-cell, for the main crops. Source: Mapspam (International Food Policy Research Institute 2019).
decay_12_time_since_acled_dummy_pr	Exponential decay of time_since_acled_dummy_pr with half-life 12. Source: UCDP (Sundberg and Melander 2013).
tlag1_dr_mod_gs	Modest, low-intensity drought occurring in a grid-month during which the growing season is ongoing. 'Modest' corresponds to values of the SPEI (Standardized Precipitation Evapotranspiration Index, Vicente-Serrano et al., 2010) between -0.5 and -1. Lagged by 1 month.
tlag1_dr_mod_gs6	Modest, low-intensity drought occurring in a grid-month during which the growing season is ongoing. 'Modest' corresponds to values of the SPEI (Standardized Precipitation Evapotranspiration Index, Vicente-Serrano et al., 2010) between -0.5 and -1. The SPEI used for computation is provided at a six-month scale. Lagged by 1 month.
tlag1_dr_moder_gs	Moderate drought occurring in a grid-month during which the growing season is ongoing. 'Moderate' corresponds to values of the SPEI (Standardized Precipitation Evapotranspiration Index, Vicente-Serrano et al., 2010) between -1 and -1.5. Lagged by 1 month.
tlag1_dr_moder_gs6	Moderate drought occurring in a grid-month during which the growing season is ongoing. 'Moderate' corresponds to values of the SPEI (Standardized Precipitation Evapotranspiration Index, Vicente-Serrano et al., 2010) between -1 and -1.5. The SPEI used for computation is provided at a six-month scale. Lagged by 1 month.
tlag1_dr_sev_gs	Severe drought occurring in a grid-month during which the growing season is ongoing. 'Severe' corresponds to values of the SPEI (Standardized Precipitation Evapotranspiration Index, Vicente-Serrano et al., 2010) lower than -2. Lagged by 1 month.
tlag1_dr_sev_gs6	Severe drought occurring in a grid-month during which the growing season is ongoing. 'Severe' corresponds to values of the SPEI (Standardized Precipitation Evapotranspiration Index, Vicente-Serrano et al., 2010) lower than -2. The SPEI used for computation is provided at a six-month scale. Lagged by 1 month.
fvp_prop_irrelevant	Proportion of population which is assigned the status 'irrelevant'. Source: EPR (Cederman, Wimmer, and Min 2010).
fvp_prop_powerless	Proportion of population which is assigned the status 'powerless'. Source: EPR (Cederman, Wimmer, and Min 2010).

Short name	Description
fvp_ssp2_edu_sec_15_24_prop	Proportion of the population between 15 and 24 that has completed at least lower secondary schooling implies those that have completed lower or upper secondary school. Those that have attained tertiary education are included in this number. Source: IIASA (Samir and Lutz 2008).
ged_best_ns	Best estimate of fatalities for non-state violence. A non-state conflict is defined by the Uppsala Conflict Data Program (UCDP) as "the use of armed force between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year."
ged_best_os	Best estimate of fatalities for one-sided violence. One-sided violence is the use of armed force by the government of a state or by a formally organized group against civilians which results in at least 25 deaths. Extrajudicial killings in custody are excluded. Source: UCDP (Sundberg and Melander 2013).
ged_best_sb	Best estimate of fatalities for state-based violence. A state-based armed conflict is a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in one calendar year. "Armed conflict" is also referred to as "state-based conflict", as opposed to "non-state conflict", in which none of the warring parties is a government. Source: UCDP (Sundberg and Melander 2013).
growseasdummy	Binary variable indicating if the growing season is ongoing for each month and grid-cell. The growing season refers to the major crops cultivated in that cell. Source: MIRCA (Portmann, Siebert, and Döll 2010).
tlag_12_harvarea_maincrops	Harvested area of the main crops cultivated in that grid-cell. Source: Mapspam (International Food Policy Research Institute 2019). Lagged by 12 months.
tlag_12_irr_maincrops	Binary variable indicating if the cultivated crops are irrigated. Main crops are defined by their harvested area. Source: Mapspam (International Food Policy Research Institute 2019). Lagged by 12 months.
ln_fvp_timeindep	Natural log of the total number of years since the country became an internationally recognized sovereign state. Source: computed from the entrance dates set by Gleditsch and Ward 1999.
ln_fvp_timesincepreindepwar	Natural log of the total number of years since the country experienced a pre-independence war. Source: computed from the entrance dates set by Gleditsch and Ward 1999.
pgd_agri_gc	Measures the coverage of agricultural areas in each cell, extracted from the Globcover 2009 dataset v.2.3. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_barren_ih	Gives the percentage area of the cell covered by barren area, based on ISAM-HYDE landuse data. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_bdist3	Gives the spherical distance (in kilometer) from the cell centroid to the territorial outline of the country the cell belongs to. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_capdist	Gives the distance to the state's capital. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_cmr_mean	Measures the average prevalence of child malnutrition, based on raster data from the SEDAC Global Poverty Mapping project. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_excluded	Counts the number of excluded groups (discriminated or powerless) as defined in the GeoEPR/EPR data on the status and location of politically relevant ethnic groups settled in the grid cell for the given year, derived from the GeoEPR/EPR 2014 update 2 dataset. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_forest_ih	Gives the percentage area of the cell covered by forest area, based on ISAM-HYDE landuse data. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_gcp_mer	Indicates the gross cell product, measured in USD, based on the G-Econ dataset v4.0, last modified May 2011. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_imr_mean	Measures the average infant mortality rate, based on raster data from the SEDAC Global Poverty Mapping project. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_landarea	Gives the total area covered by land in the grid cell in square kilometers as defined by the CShapes dataset. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_mountains_mean	Proportion of mountainous terrain within the cell based on elevation, slope and local elevation range. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).

Short name	Description
pgd_nlights_calib_mean	Measures average nighttime light emission from the DMSP-OLS Nighttime Lights Time Series Version 4 (Average Visible, Stable Lights, Cloud Free Coverages), calibrated to account for intersatellite differences and inter-annual sensor decay using calibration values from Elvidge et.al. (2013). Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_pasture_ih	Gives the percentage area of the cell covered by pasture area, based on ISAM-HYDE landuse data. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_pop_gpw_sum	Measures population size, taken from the Gridded Population of the World version 3. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_savanna_ih	Gives the percentage area of the cell covered by grasslands, based on ISAM-HYDE landuse data. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_ttime_mean	Is an estimate of the average travel time to the nearest major city, derived from a global high-resolution raster map of accessibility developed for the EU. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_urban_gc	Measures the coverage of urban areas in each cell, based on the Globcover 2009 dataset v.2.3. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
pgd_water_ih	Gives the percentage area of the cell covered by water area, based on ISAM-HYDE landuse data. Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
rainf_maincrops	Binary variable indicating if the cultivated crops are only rainfed. Main crops are defined by their harvested area, i.e. those with the highest amount of harvest area in that grid-cell-month. Source: Mapspam (International Food Policy Research Institute 2019).
reign_age	An approximation of the leader's age calculated by subtracting the leader's birth year from year. It takes the same value for each month in a calendar year. Source: REIGN (Bell 2016).
reign_couprisk	Risk of Coup d'Etats estimated by REIGN (Bell 2016).
reign_elected	Elected is 0 for months when the incumbent leader has never been elected to the highest office and 1 if the leader has been elected to that office. Source: REIGN (Bell 2016).
reign_election_now	Is equal to 1 if a qualifying election occurs during the month observed. Source: REIGN (Bell 2016).
reign_gov_foreign_occupied	Governments occur where foreign politicians or militaries hold de facto power over a government. Source: REIGN (Bell 2016).
reign_gov_military	Indicates that a military committee runs the country. One officer typically serves as head, but this head serves the interests of the committee and his power is checked by other members of the military. Source: REIGN (Bell 2016).
reign_gov_monarchy	Indicates that the power is highly concentrated in the hands of a monarch who is much more than just a figurehead. Source: REIGN (Bell 2016).
reign_gov_personal_dictatorship	Indicates that the power is highly concentrated in the hands of a non-monarch dictator who is relatively unconstrained by a military or political party. Source: REIGN (Bell 2016).
reign_gov_provisional_civilian	Dummy variable that equals 1 if the government of the country is classified as a provisional civilian. Source: REIGN (Bell 2016)
reign_gov_warlordism	Occurs only in countries that are torn apart by conflict to the extent that they do not have a functional government. Source: REIGN (Bell 2016).
reign_irregular	Is equal to 1 in the six months following irregular elections. Elections are irregular if they are not occurring as part of an established pattern or norm for executive selection. All referendums are irregular by nature, but they are not included here. Source: REIGN (Bell 2016).
reign_lastelection	Is the number of months since the last qualifying election (election for highest office or referendum that would expand executive power), or, in the absence of previous elections, the number of months since the political system last changed (see Regime Characteristics section). Source: REIGN (Bell 2016).

Short name	Description
reign_lead_recent	Is equal to 1 in the six months following a qualifying non-referendum election. Source: REIGN (Bell 2016).
reign_male	Dummy equal ti 1 if the leader is male and 0 if the leader is female. Source: REIGN (Bell 2016).
spdist_pgd_diamsec	Spatial distance from the grid cell to the nearest secondary diamonds resource (static data). Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
spdist_pgd_petroleum	Spatial distance from the grid cell to the nearest petroleum resource (static data). Source: PRIOGRID (Tollefsen, Strand, and Buhaug 2012).
spei1_gs_prev10	Average agricultural drought for the previous 10 years, proxied by the value of SPEI during the growing season months. Source: author's computation from SPEI Drought Monitor (Vicente-Serrano, Beguería, and López-Moreno 2010).
spei1_gs_prev10_anom	Difference between the current value of agricultural SPEI (during the growing season) and the average value of SPEI during the growing season months in the previous 10 years. Source: author's computation from SPEI Drought Monitor (Vicente-Serrano, Beguería, and López-Moreno 2010).
tlag1_spei1_gsm	spei1_gsm. SPEI value for the months in which the growing season is ongoing. Source: author's computation from SPEI Drought Monitor (Vicente-Serrano, Beguería, and López-Moreno 2010). Lagged by 1 month.
spei1_gsm_cv_anom	Difference between the temporal coefficient of variation of spei1_gsm along the current year, and the average variation of spei1_gsm over the period 1990-2010. Source: author's computation from SPEI Drought Monitor (Vicente-Serrano, Beguería, and López-Moreno 2010).
spei1_gsm_detrend	Source: author's computation from SPEI Drought Monitor (Vicente-Serrano, Beguería, and López-Moreno 2010).
spei1gsy_lowermedian_count.	Count of the previous years (up to 5) during which spei1_gsm is lower than the median value for the reference period (1990-2010). Source: author's computation from SPEI Drought Monitor (Vicente-Serrano, Beguería, and López-Moreno 2010).
spei_48_detrend	SPEI (Standardized Precipitation Evapotranspiration Index) at 48-month scale, detrended. Source: author's computation from SPEI Drought Monitor (Vicente-Serrano, Beguería, and López-Moreno 2010).
splag_1_1_acled_dummy_pr	Spatial lag of acled_dummy_pr from neighborhood order 1 to 1. Source: ViEWS' computations from ACLED (Raleigh et al. 2010)
splag_1_1_ged_best_ns	Spatial lag of ged_best_ns from neighborhood order 1 to 1. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
splag_1_1_ged_best_os	Spatial lag of ged_best_os from neighborhood order 1 to 1. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
splag_1_1_ged_best_sb	Spatial lag of ged_best_sb from neighborhood order 1 to 1. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
time_since_ged_dummy_ns	Time since ged_dummy_ns=1. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_ged_dummy_os	Time since ged_dummy_os=1. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_ged_dummy_sb	Time since ged_dummy_sb=1. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_greq_25_ged_best_sb	Time since ged_best_sb is at least equal to 25. Source: UCDP (Sundberg and Melander 2013).
time_since_greq_25_ged_best_os	Time since ged_best_os is at least equal to 25. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_greq_25_ged_best_ns	Time since ged_best_ns is at least equal to 25. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_greq_25_splag_1_1_ged_best_ns	Time since ged_best_ns is at least equal to 25 in a neighboring cell. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).

Short name	Description
time_since_greq_25_splag_1_1_ged_best_os	Time since ged_best_os=1 is at least equal to 25 in a neighboring cell. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_greq_25_splag_1_1_ged_best_sb	Time since ged_best_sb=1 is at least equal to 25 in a neighboring cell. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_greq_500_ged_best_ns	Time since ged_best_ns is at least equal to 500. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
time_since_greq_500_ged_best_os	Time since ged_best_os is at least equal to 500. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_greq_500_ged_best_sb	Time since ged_best_sb=1 is at least equal to 500. Source: ViEWS' computations fromUCDP (Sundberg and Melander 2013).
time_since_greq_500_splag_1_1_ged_best_ns	Time since ged_best_ns=1 is at least equal to 500 in a neighboring cell. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
time_since_greq_500_splag_1_1_ged_best_os	Time since ged_best_os=1 is at least equal to 500 in a neighboring cell. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
time_since_greq_500_splag_1_1_ged_best_sb	Time since ged_best_sb=1 is at least equal to 500 in a neighboring cell. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
time_since_splag_1_1_ged_dummy_ns	Time since ged_dummy_ns is at equal to 1 in a neighboring cell. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
time_since_splag_1_1_ged_dummy_os	Time since ged_dummy_os is at equal to 1 in a neighboring cell. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
time_since_splag_1_1_ged_dummy_sb	Time since ged_dummy_sb is at equal to 1 in a neighboring cell. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
tlag_12_vdem_v2x_civlib	Civil liberties index, lagged by 12 months. Source: V-Dem (Coppedge 2020).
tlag_12_vdem_v2x_clphy	Physical violence index, lagged by 12 months. . Source: V-Dem (Coppedge 2020).
tlag_12_vdem_v2x_clpriv	Private civil liberties index, lagged by 12 months. . Source: V-Dem (Coppedge 2020).
tlag_12_vdem_e_v2x_freexp_altinf	Freedom of expression index, lagged by 12 months. Source: V-Dem (Coppedge 2020).
tlag_12_vdem_e_v2xel_regelec.	Regional government index, lagged by 12 months. Source: V-Dem (Coppedge 2020).
tlag_12_vdem_v2cldmovem	Freedom of movement index, lagged by 12 months. Source: V-Dem (Coppedge 2020).
tlag_12_vdem_v2xcl_rol	Equality before the law and individual liberty index, lagged by 12 months. Source: V-Dem (Coppedge 2020).
tlag_1_ged_best_ns	Time lag by 1 months of ged_best_ns. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
tlag_1_ged_best_os	Time lag by 1 months of ged_best_os. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
tlag_1_ged_best_sb	Time lag by 1 months of ged_best_sb. Source: ViEWS' computations from UCDP (Sundberg and Melander 2013).
wdi_ag_lnd_agri_zs	Agricultural land (% of land area). Source: WDI (The World Bank).
wdi_ag_lnd_totl_ru_k2	Rural land area (sq. km). Source: WDI (The World Bank).
wdi_ag_lnd_totl_ru_k2	Rural land area (sq. km). Source: WDI (The World Bank).
wdi_ag_srf_totl_k2	Country's total area, including areas under inland bodies of water and some coastal waterways. Source: WDI (The World Bank).
wdi_nv_agr_empl_kd	Agriculture, value added per worker (constant 2010 US$). Source: WDI (The World Bank).
wdi_nv_agr_totl_cn	Agriculture, value added (current LCU). Source: WDI (The World Bank).
wdi_nv_agr_totl_kd	Agriculture, value added (constant 2010 US$). Source: WDI (The World Bank).
wdi_tx_val_food_zs_un	Food exports (% of merchandise exports). Food comprises the following commodities: food and live animals, beverages and tobacco, animal and vegetable oils and fat, oil seeds, oil nuts, and oil kernels. Source: WDI (The World Bank).

2 Forecasting Models

All the forecasting models in the present analysis are based on random forest regressors. Random forest models (Breiman 2001; Muchlinski et al. 2016) are machine-learning techniques employing classification and regression trees (CART). CART predicts the outcome variable Y by linking it to some predictors X via a decision tree. The tree consists of a number of splits into branches, where each split is defined by searching all values in X to find the one which best predicts the value of Y. The tree continues to be split until some threshold is achieved (to avoid overfitting). The number of features that can be split on at each node is limited to a share of the total, ensuring that the model does not rely too heavily on any individual feature. In RF models, decision trees are combined with bootstrap-aggregating (bagging), and random feature selection. Bagging grows an ensemble of trees, each slightly different. The trees are created in parallel, i.e. independently among each other. These trees are, however, correlated as some variables are especially good at predicting Y. To avoid this correlation, each tree draws a random sample from the original data set when generating its splits, solving the correlation problem and creating a forest of uncorrelated 'random' trees.

2.1 Calibration

As for all ViEWS models (Hegre et al. 2019), the training, evaluating and testing procedures are carefully vetted to increase as much as possible the availability of data without causing any leakage, and to enhance predictive performance. The setup enables to maximize the precision of the models' evaluation and ensures to manage the data as efficiently as possible (Hegre et al. 2020).

The calibration is performed for each model to make sure that the mean (delta transformed) outcome in the calibration set is close to the average observed change in fatalities. If a model is well calibrated, the average predicted outcome for a set of cases is similar to the actual relative frequency for that set.

For each step, the calibration function is defined as:

$$Y_c = \frac{1}{m}\sum_{i=1}^{m} Y_{a,c} + \left(Y_{p,t} - \frac{1}{m}\sum_{i=1}^{m} Y_{p,c}\right) * \frac{\sigma Y_{a,c}}{\sigma Y_{p,c}} \tag{1}$$

where c denotes the calibration partition, t the test partition; a refers to actual observations, p to predicted values and σ denotes the standard deviation; $Y_{a,c}$ indicates the actual values (logged change in BRDs) observed in the calibration test, $Y_{p,t}$ refers to the predicted values in the test partition and so on.

2.2 Evaluation Metrics

I use a set of metrics to evaluate the performance of the models. The Mean Squared Error is the expected value of the squared error or loss, and corresponds to the square of the Mean Absolute Error.

I also include a new evaluation metric function to specifically measure the models' accuracy in predicting a *change* in the number of BRDs. The Targeted Absolute Distance with Direction Augmentation (TADDA) (Colaresi 2020) is a function developed to capture the models performance and accuracy in predicting the change in fatalities and not necessarily the absolute number. TADDA is simple, not as sensitive as MSE to outlier predictions and includes a penalty when prediction is of the opposite sign than the actual direction of change. The cost penalty grows when the error increases, but it is discounted to zero when the predicted change in fatalities is wrong in sign (positive or negative compared to the actual) but very close in absolute terms to the observed value [3]. For each given model

[3] Between two predictions which are equi-distant to the actual value (the error is of the same magnitude), TADDA assigns a higher score to the prediction that displays the right sign (goes in the right direction) compared to the one that goes in the opposed direction of change compared to the actual outcome; likewise, between two predictions of the same

M, TADDA is computed as:

$$\text{TADDA}^{(M)} = \frac{\sum_{i=1}^{N} |y_{\Delta,i} - f_{\Delta,i}| + |f_{\Delta,i}| I[y_{\Delta,i}^{(\pm)} \neq f_{\Delta,i}^{(\pm)}] I\left[|y_{\Delta,i} - f_{\Delta,i}| > \epsilon\right]}{N}$$

where $y_{\Delta,i}$ represents the actual change in number of BRDs, $f_{\Delta,i}$ is the forecasted change in observation i and ϵ is a small enough value to define the cost penalty threshold. In the models, I set ϵ to 0.048. Higher positive or negative TADDA values indicate higher distance between the predicted and the actual observation, hence worse predictive performance. The closer TADDA is to 0, the best is the prediction.

3 Ablation

The ablation study is accomplished in three steps. First, I train the drought_base model with all features included in Table II in the main text, and I calibrate and test the model as described in the methodology section. Second, the ablation process step-wisely trains, calibrates and evaluates a version of the whole model reduced by feature i. Finally, I compute the difference in MSE between the whole, comprehensive model with all features and the various reduced versions of the same model with features $f - i$. The variation in MSE registered for each reduced model compared to the extended one is a measure of the loss in the model's predictive performance caused by dropping that particular feature. I call this variation MAL.

Specifically, I define MAL as the change in the model MSE determined by removing the feature from that model:

$$MAL_i^{MSE} = -(MSE^{whole} - MSE_i^{abl}) \tag{2}$$

where MSE^{whole} is the MSE for the model with all features and MSE_m^{abl} the score for the model without feature i. When MSE increases as a result of the feature's removal, the contribution of that feature is positive – removing it from the model deteriorates the predictive performance. If MSE decreases, the model performs better without that feature.

The metric is positive – the dot is to the right of the central line – if the model without feature i is worse than the model including it, meaning that the feature contributes positively to the predictive performance.

sign, it prefers the one which is closer to the observed outcome.