UPPSALA UNIVERSITET

ViEWS PREDICTING CONFLICT

# Constituent models, *cm*.

## Online appendix B to ViEWS$_{2020}$: Revising and evaluating the ViEWS political Violence Early-Warning System.


Håvard Hegre[1,2], Curtis Bell[1,3], Michael Colaresi[1,4], Mihai Croicu[1], Frederick Hoyles[1], Remco Jansen[1], Maxine Ria Leis[1], Angelica Lindqvist-McGowan[1], David Randahl[1], Espen Geelmuyden Rød[1], and Paola Vesco[1]

[1]Department of Peace and Conflict Research, Uppsala University
[2]Peace Research Institute Oslo
[3]U.S. Naval War College
[4]University of Pittsburgh


January 21, 2021


**Abstract**

This appendix documents all the constituent models in the ensembles at the country-month (*cm*) level. We give an overview of the main estimation/machine-learning techniques. We describe and motivate each model, report the importance of the main features, and show prediction maps for selected steps for the 2020–2022 period.

# CONTENTS

## B-1 ESTIMATION OF CONSTITUENT MODELS

ViEWS relies on logistic regression, random forest (RF) and gradient boosting machines (GBM) models. The logit model is a generalized linear model (GLM) that performs well compared to many machine-learning techniques (Géron, 2017). Computational costs are low, and with large datasets like ours overfitting is not a serious concern. Both RF models (Breiman, 2001; Muchlinski et al., 2016) and GBM models (Friedman, 2001; Chen and Guestrin, 2016) are machine-learning techniques employing classification and regression trees (CART). In CART, the response variable $Y$ is predicted using a decision tree and some predictor variables $X$. The tree consists of a number of 'splits' into different branches. Each split is found by searching all values in $X$ to find the constant which maximally separates between the categories of $Y$. The tree continues to be split until some threshold is achieved (to avoid overfitting).

In RF models, CART trees are combined with bootstrap-aggregating (bagging), and random feature selection. Bagging creates an ensemble of trees, each slightly different. These trees are, however, correlated as some variables are especially good at discriminating $Y$. To avoid this, a random subset of variables (predictors or features) are selected for each 'split', solving the correlation problem and creating a forest of uncorrelated 'random' trees.

In GBM models, an ensemble of CART trees are trained in an additive, iterative fashion. At each iteration, a new CART tree (a new learner) is added to the ensemble, grown in such a way as to minimize a chosen, derivable, loss function capturing remaining classification or regression error. In practice, this is done by training each new tree on a regularized form of the residuals computed between the ensemble predictions at the previous step and the observed values. Each such new tree is added to the ensemble, and the iterative process is restarted (Friedman, 2001; Chen and Guestrin, 2016). This iterative improvement is referred to as 'boosting'; this iterative minimization of residual-based loss functions is equivalent with gradient descent. ViEWS uses the XGBoost implementation of GBM, an implementation that adds a regularized learning objective adding a complexity penalty to tree growing and smoothing final learned weights, as well as implementing the same random feature selection and sub-sampling of data (Chen and Guestrin, 2016).

RF and XGB models are memory-intensive to estimate given the computational capacity available for ViEWS. Hence, at the *pgm* level, we estimate them using a 'downsampled' dataset which includes all conflict events and a random sample of non-events (see details in the 'Overview of models' section in this appendix).

## B-2 HANDLING FORECASTING DYNAMICS

ViEWS employs two alternative strategies to compute forecasts for each of $s \in (1, .., 36)$ months into the future. We call these Dynamic simulation (*ds*) and One-step-ahead (*osa*).

### Dynamic simulation

The dynamic simulation (Dynasim) procedure builds on Hegre et al. (2013) and Hegre et al. (2016), and is discussed at length there.[1] The procedure involves simulating model parameters based on the estimated coefficients and the variance-covariance matrix from a model. In addition, we compute predicted probabilities for the outcomes for the first month $t$, draw outcomes, recalculate the history variables so that the input predictor matrix $X_{t+1}$ at $t+1$ reflects that draw. This is repeated for each month $s$ in the forecasting window, and for each simulated set of parameters.

If we are interested in forecasting two months into the forecasting period, we first train the constituent models, estimate the weights, and produce our ensemble one-month-ahead forecast. To produce forecasts

[1] The first author's original script 'PRIOsim' was rewritten in C# and C++ by Joakim Karlsen. The Python routines underlying the current projects are based on the 'Dynasim' reimplementation of this, written by Jonas Vestby and Frederick Hoyles.

for the next month, we need the input predictor matrices $X_{t+1}^{(k)}$. For many constituent models, these input predictors will themselves be functions of actual conflict (e.g., lagged conflict indicators, time since last conflict, spatial distance to nearest conflict). Since these do not exist for the next month (after the training window), we use the prediction as the probability of an unobserved predictor, for example for conflict at time $t+1$, when forecasting conflict at $t+2$. A simulated value is drawn from this probability, and recorded within a new simulated set of predictors $\widetilde{X}_{t+1}^{(k)}$.

The predictions for the three outcomes are obtained simultaneously within each time step. For each of these, we compute the predicted probability at $t+1$ as a function of information available at $t$, including the status for the other two outcomes. This procedure repeats for every month to the end of the forecasting window.

### 'One-step-ahead' modeling

In the 'one-step-ahead' modeling, we predict each step into the future ($t+1$, (...), $t+12$, (...) $t+36$) independently, as opposed to dynamic simulation which moves forward sequentially. We do this by estimating a set of models of the form $f_s(X_{t-s})$ where $s$ denotes the number of months into the future to forecast. In regression notation these take the form $y_{t+s} = X_t\beta_t,\ for\ s \in (1,36)$. The 'one-step-ahead' model does this by time-shifting the right-hand side variables with respect to the outcome before models are trained, making the model a link function between the future ($y_{t+s}$) and the present ($X_t$).

## B-3 OVERVIEW OF MODELS

The new ensembles include 16 models at the *cm* level and 12 at the *pgm* level. The dynasim models are derived using the 'dynamic simulation' approach described in (Hegre et al., 2013; Hegre et al., 2019). All other models are currently random forest models.

Here we present a more extensive description of each model than in the main text. We have defined a set of 'models' characterized by a set of features as well as an estimation procedure. All models are trained for each of the three different types of conflict, i.e. state-based **(sb)**, non-state **(ns)** and one-sided **(os)** violence, as well as for all steps $s$.

If nothing else is stated, the estimator is the random forest classifier (Breiman, 2001) from the scikit-learn package (Pedregosa et al., 2011). All parameters used are the package defaults, except for $n\_estimators$ which controls the number of trees in the forest and is set to 1000. With these settings the procedure works as follows: For each of 1000 trees draw a bootstrapped sample of the same size as the training data. On this sample fit a decision tree using the gini split quality criterion, considering the square root of the number of features for each split. Predicted probabilities from the forest are computed as the average of the predicted probability from each tree, which is the fraction of samples for the class in the current leaf of the tree.

Feature importances are a measure of the importance of each feature in the forest. Below, we present impurity-based features importances.[2] The importances of all features sum to one and are assigned according to the position of the feature in the trained trees, with features appearing higher up in the trees receiving a higher score as the expected fraction of samples they contribute to predicting is higher.

Below, we present a description of the model features used as predictors for the different dependent variables. The following sections will provide further details on the types of outcome which is predicted.

[2]For a critical review of impurity-based feature importances, see Strobl et al. (2007) and Strobl et al. (2008)

*Table B-1.* Models in the *cm* ensemble as of r.2020.02.01

| Model name |
|---|
| acled_violence |
| cflong |
| onset24_25_all |
| reign_coups |
| acled_protest |
| cdummies |
| demog |
| icgcw |
| neibhist |
| reign_drought |
| ds_25 |
| ds_dummy |
| all_glob |
| reign_glob |
| vdem_glob |
| wdi_all_glob |

## B-3.1 ACLED violence (acled_violence)

### Description

A model including variables that capture the recent history of violence according to UCDP definitions, but sourced from ACLED as an alternative dataset (Raleigh et al., 2010).

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

### Feature importances

Feature importances are summarized in the table below for each type of outcome. The importance of conflict history in predictions tends to slightly increase over time, while spatial lags show the opposite temporal trend. Feature importances also shows that the time-factor is relatively more important than spatial proximity in predicting violence. As expected, past conflict episodes are less important in predicting violence in neighboring countries than in the country where they occurred.

*Table B-2.* Feature importances for the ACLED violence model, **sb**, **os**, **ns**

| sb_acled_violence | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| splag_1_1_acled_count_os | 0.203784 | 0.205966 | 0.202270 | 0.198235 | 0.196691 | 0.190503 |
| splag_1_1_acled_count_sb | 0.181902 | 0.180649 | 0.180568 | 0.172854 | 0.168742 | 0.163956 |
| time_since_acled_dummy_ns | 0.148019 | 0.150630 | 0.151186 | 0.152234 | 0.163273 | 0.169117 |
| splag_1_1_acled_count_ns | 0.137840 | 0.134506 | 0.132439 | 0.131717 | 0.126310 | 0.119548 |
| time_since_acled_dummy_sb | 0.116101 | 0.115830 | 0.117793 | 0.118526 | 0.117838 | 0.121929 |
| time_since_splag_1_1_acled_dummy_ns | 0.064751 | 0.065648 | 0.065588 | 0.067864 | 0.066765 | 0.067278 |
| time_since_acled_dummy_os | 0.057924 | 0.056845 | 0.057878 | 0.061654 | 0.064601 | 0.069517 |
| time_since_splag_1_1_acled_dummy_sb | 0.056041 | 0.055758 | 0.056924 | 0.059437 | 0.057413 | 0.058867 |
| time_since_splag_1_1_acled_dummy_os | 0.033637 | 0.034168 | 0.035354 | 0.037477 | 0.038368 | 0.039284 |
| **os_acled_violence** | **1** | **3** | **6** | **12** | **24** | **36** |
| splag_1_1_acled_count_os | 0.199975 | 0.195862 | 0.194335 | 0.186287 | 0.179364 | 0.171649 |
| splag_1_1_acled_count_sb | 0.167628 | 0.166497 | 0.162252 | 0.156669 | 0.149580 | 0.143813 |
| time_since_acled_dummy_ns | 0.157826 | 0.160189 | 0.164685 | 0.170079 | 0.185615 | 0.191827 |
| splag_1_1_acled_count_ns | 0.146841 | 0.144082 | 0.141609 | 0.144233 | 0.136608 | 0.128028 |
| time_since_acled_dummy_sb | 0.082730 | 0.081223 | 0.084514 | 0.087868 | 0.090717 | 0.098982 |
| time_since_splag_1_1_acled_dummy_ns | 0.071934 | 0.071795 | 0.073297 | 0.072767 | 0.072486 | 0.070442 |
| time_since_acled_dummy_os | 0.064350 | 0.067466 | 0.067207 | 0.066817 | 0.072608 | 0.076240 |
| time_since_splag_1_1_acled_dummy_sb | 0.062310 | 0.066170 | 0.065235 | 0.067814 | 0.063691 | 0.065376 |
| time_since_splag_1_1_acled_dummy_os | 0.046405 | 0.046716 | 0.046866 | 0.047467 | 0.049332 | 0.053644 |
| **ns_acled_violence** | **1** | **3** | **6** | **12** | **24** | **36** |
| splag_1_1_acled_count_os | 0.213286 | 0.209496 | 0.208006 | 0.208822 | 0.201273 | 0.201699 |
| splag_1_1_acled_count_sb | 0.177784 | 0.174119 | 0.172055 | 0.169918 | 0.171590 | 0.159144 |
| splag_1_1_acled_count_ns | 0.158270 | 0.157109 | 0.157167 | 0.156729 | 0.144037 | 0.137438 |
| time_since_acled_dummy_ns | 0.106641 | 0.107982 | 0.106513 | 0.109618 | 0.111302 | 0.109856 |
| time_since_splag_1_1_acled_dummy_ns | 0.090735 | 0.094066 | 0.095822 | 0.098704 | 0.106291 | 0.116268 |
| time_since_acled_dummy_sb | 0.076801 | 0.083742 | 0.084110 | 0.085199 | 0.086375 | 0.093171 |
| time_since_acled_dummy_os | 0.075870 | 0.074970 | 0.077677 | 0.071350 | 0.073214 | 0.072148 |
| time_since_splag_1_1_acled_dummy_sb | 0.058390 | 0.057321 | 0.058138 | 0.059115 | 0.064959 | 0.065420 |
| time_since_splag_1_1_acled_dummy_os | 0.042223 | 0.041195 | 0.040512 | 0.040545 | 0.040959 | 0.044856 |

#### Change history

The model has been in use since r.2020.02.01

## B-3.2 Conflict history, long (cflong)

#### Description

This model includes an extensive number of features tracing the history of conflict of each country. The legacy of conflict has been shown to affect the probability of violence even long-term, whereby countries which experienced conflict in the past are more likely to be affected by violence in the present (Besley and Reynal-Querol, 2014). The model includes the counts of time (in months) since the last violent episode, for any type of violence (**sb**, **os**, **ns**) and for different BRD thresholds (causing at least 1, 25, 100, or 500 deaths in a month). Likewise, the model includes a set of lagged dependent variables, using a number of different temporal lags up to 12 months, one for each of the four intensity levels (i.e. BRDs thresholds) specified above and all three types of outcome.

#### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

#### Feature importances

The tables presented below show the feature importances for the each violent outcome. The relative importance of conflict history features in predicting each type of violence is outcome-specific; the history of state-based conflict is more important to predict state-based violence, the occurrence of one-sided violence in the past contributes more in predicting **os**, and non-state conflict history is more relevant to predict **ns**. As expected, we observe the importance of temporal lags in predictions to decrease according to the length of the selected lags, whereby shorter lags are more influential than longer ones.

*Table B-3.* Feature importances for the cflong model, **sb**

| sb_cflong | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_25_ged_best_sb | 0.057704 | 0.044439 | 0.044808 | 0.042008 | 0.036203 | 0.034864 |
| tlag_1_greq_25_ged_best_sb | 0.043727 | 0.029904 | 0.025386 | 0.023016 | 0.016914 | 0.011409 |
| time_since_greq_100_ged_best_sb | 0.041625 | 0.036336 | 0.031506 | 0.032886 | 0.034977 | 0.038602 |
| tlag_2_greq_25_ged_best_sb | 0.038593 | 0.032397 | 0.033176 | 0.024270 | 0.016895 | 0.010527 |
| tlag_3_greq_25_ged_best_sb | 0.027254 | 0.027318 | 0.023922 | 0.018262 | 0.018453 | 0.012629 |
| time_since_greq_5_ged_best_sb | 0.024641 | 0.024329 | 0.023456 | 0.023984 | 0.023350 | 0.022060 |
| tlag_4_greq_25_ged_best_sb | 0.023092 | 0.022070 | 0.020305 | 0.015058 | 0.012060 | 0.008342 |
| tlag_5_greq_25_ged_best_sb | 0.020576 | 0.020626 | 0.023502 | 0.009324 | 0.008982 | 0.011583 |
| tlag_1_greq_5_ged_best_sb | 0.019805 | 0.019306 | 0.015359 | 0.014693 | 0.010527 | 0.007037 |
| tlag_7_greq_25_ged_best_sb | 0.018487 | 0.011948 | 0.012587 | 0.009490 | 0.007591 | 0.004982 |
| time_since_ged_dummy_sb | 0.017696 | 0.020911 | 0.020260 | 0.022727 | 0.023725 | 0.019218 |
| tlag_2_greq_5_ged_best_sb | 0.016997 | 0.014707 | 0.010971 | 0.013983 | 0.008066 | 0.008050 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.016022 | 0.018631 | 0.020790 | 0.025201 | 0.032327 | 0.035323 |
| tlag_6_greq_25_ged_best_sb | 0.015638 | 0.017810 | 0.017722 | 0.008377 | 0.007011 | 0.010352 |
| time_since_greq_500_ged_best_ns | 0.015475 | 0.017961 | 0.019633 | 0.023250 | 0.027707 | 0.031002 |
| time_since_greq_500_ged_best_os | 0.015407 | 0.017035 | 0.018918 | 0.021055 | 0.023453 | 0.023910 |
| time_since_greq_500_ged_best_sb | 0.015265 | 0.016136 | 0.017920 | 0.020840 | 0.023639 | 0.025458 |
| time_since_greq_100_ged_best_os | 0.015132 | 0.017069 | 0.019018 | 0.020432 | 0.023563 | 0.026130 |
| time_since_greq_500_splag_1_1_ged_best_ns | 0.014525 | 0.016548 | 0.018032 | 0.021324 | 0.024095 | 0.028039 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.014514 | 0.015773 | 0.017270 | 0.019776 | 0.022292 | 0.026099 |
| time_since_greq_100_ged_best_ns | 0.014204 | 0.016423 | 0.018590 | 0.022393 | 0.025210 | 0.026202 |
| time_since_greq_25_ged_best_os | 0.014130 | 0.013789 | 0.016817 | 0.017592 | 0.020065 | 0.020867 |
| tlag_1_greq_1_ged_best_sb | 0.014126 | 0.012636 | 0.009972 | 0.016779 | 0.012155 | 0.006948 |
| time_since_greq_5_ged_best_os | 0.012567 | 0.013415 | 0.016257 | 0.017963 | 0.018163 | 0.018098 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.012488 | 0.013837 | 0.015707 | 0.018505 | 0.020869 | 0.023765 |

*Table B-4.* Feature importances for the cflong model, **os**

| os_cflong | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_25_ged_best_os | 0.047286 | 0.044925 | 0.036161 | 0.035936 | 0.029476 | 0.026226 |
| tlag_1_greq_25_ged_best_os | 0.031578 | 0.022868 | 0.020365 | 0.014936 | 0.005782 | 0.005207 |
| tlag_2_greq_25_ged_best_os | 0.024764 | 0.026707 | 0.014092 | 0.007902 | 0.005292 | 0.004126 |
| time_since_greq_5_ged_best_os | 0.024507 | 0.023221 | 0.023382 | 0.020764 | 0.026707 | 0.023286 |
| time_since_greq_100_ged_best_os | 0.023986 | 0.023791 | 0.025902 | 0.025800 | 0.027386 | 0.026742 |
| tlag_2_greq_5_ged_best_os | 0.022660 | 0.018753 | 0.012731 | 0.012018 | 0.007854 | 0.006093 |
| time_since_greq_500_ged_best_os | 0.022121 | 0.023729 | 0.025582 | 0.027833 | 0.029752 | 0.029437 |
| tlag_4_greq_25_ged_best_os | 0.020511 | 0.014590 | 0.015131 | 0.004894 | 0.004646 | 0.007080 |
| tlag_3_greq_25_ged_best_os | 0.019471 | 0.018419 | 0.013350 | 0.009427 | 0.005877 | 0.006127 |
| time_since_greq_100_ged_best_ns | 0.019235 | 0.020806 | 0.020309 | 0.022836 | 0.023139 | 0.025949 |
| tlag_1_greq_5_ged_best_os | 0.018902 | 0.016808 | 0.012466 | 0.010796 | 0.016382 | 0.006602 |
| time_since_greq_500_splag_1_1_ged_best_ns | 0.017926 | 0.019414 | 0.020710 | 0.024028 | 0.026254 | 0.027456 |
| time_since_greq_500_ged_best_ns | 0.017908 | 0.020084 | 0.022833 | 0.029291 | 0.035271 | 0.039371 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.017774 | 0.019552 | 0.021162 | 0.024537 | 0.025752 | 0.026726 |
| time_since_greq_500_ged_best_sb | 0.017072 | 0.018837 | 0.020155 | 0.022631 | 0.024038 | 0.024484 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.016741 | 0.017985 | 0.020249 | 0.022550 | 0.023450 | 0.026240 |
| tlag_3_greq_5_ged_best_os | 0.015823 | 0.011559 | 0.015300 | 0.014063 | 0.008804 | 0.010325 |
| time_since_greq_100_ged_best_sb | 0.015540 | 0.016124 | 0.018597 | 0.019977 | 0.021976 | 0.023360 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.015049 | 0.016357 | 0.018648 | 0.021304 | 0.021257 | 0.024646 |
| time_since_greq_25_ged_best_ns | 0.014878 | 0.016786 | 0.019011 | 0.018693 | 0.019303 | 0.022930 |
| tlag_4_greq_5_ged_best_os | 0.014449 | 0.009622 | 0.014955 | 0.012975 | 0.007143 | 0.009893 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.014343 | 0.016091 | 0.018790 | 0.019632 | 0.019596 | 0.022525 |
| time_since_greq_25_splag_1_1_ged_best_ns | 0.013388 | 0.014251 | 0.015984 | 0.018224 | 0.021090 | 0.025204 |
| time_since_greq_5_ged_best_ns | 0.013157 | 0.014452 | 0.016083 | 0.016646 | 0.018643 | 0.021331 |
| time_since_greq_25_ged_best_sb | 0.013057 | 0.013669 | 0.014412 | 0.015930 | 0.018817 | 0.020107 |

*Table B-5.* Feature importances for the cflong model, **ns**

| ns_cflong | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_25_ged_best_ns | 0.039773 | 0.033784 | 0.035608 | 0.034286 | 0.030546 | 0.030392 |
| time_since_greq_100_ged_best_ns | 0.039338 | 0.041770 | 0.036169 | 0.036656 | 0.044725 | 0.040050 |
| time_since_greq_5_ged_best_ns | 0.029618 | 0.025368 | 0.023716 | 0.027578 | 0.025569 | 0.027090 |
| tlag_2_greq_25_ged_best_ns | 0.022170 | 0.012234 | 0.005983 | 0.011857 | 0.016256 | 0.006423 |
| time_since_ged_dummy_ns | 0.021729 | 0.018596 | 0.018407 | 0.024600 | 0.023361 | 0.023516 |
| tlag_1_greq_25_ged_best_ns | 0.021469 | 0.009196 | 0.015012 | 0.011418 | 0.005438 | 0.008173 |
| tlag_1_greq_5_ged_best_ns | 0.021334 | 0.014302 | 0.015492 | 0.015131 | 0.007075 | 0.008724 |
| time_since_greq_500_ged_best_ns | 0.020191 | 0.021554 | 0.022868 | 0.026525 | 0.030166 | 0.035138 |
| tlag_2_greq_5_ged_best_ns | 0.019422 | 0.015516 | 0.008827 | 0.011237 | 0.012210 | 0.008838 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.018942 | 0.020297 | 0.020725 | 0.022044 | 0.024195 | 0.024263 |
| time_since_greq_500_splag_1_1_ged_best_ns | 0.018493 | 0.019809 | 0.020551 | 0.021537 | 0.023595 | 0.024201 |
| time_since_greq_500_ged_best_sb | 0.018031 | 0.019851 | 0.019886 | 0.023510 | 0.026813 | 0.025580 |
| time_since_greq_500_ged_best_os | 0.017787 | 0.019744 | 0.019596 | 0.023141 | 0.025614 | 0.029329 |
| time_since_greq_100_ged_best_os | 0.016580 | 0.017490 | 0.018259 | 0.021010 | 0.021970 | 0.024502 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.015664 | 0.015869 | 0.017184 | 0.018955 | 0.020093 | 0.020226 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.014989 | 0.015346 | 0.015695 | 0.017467 | 0.018551 | 0.020643 |
| time_since_greq_25_ged_best_os | 0.014887 | 0.015088 | 0.014428 | 0.017063 | 0.019177 | 0.021302 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.014695 | 0.016686 | 0.017482 | 0.018209 | 0.021372 | 0.021508 |
| time_since_greq_100_ged_best_sb | 0.014523 | 0.014970 | 0.015613 | 0.016829 | 0.018184 | 0.018536 |
| tlag_1_greq_1_ged_best_ns | 0.013041 | 0.009833 | 0.015833 | 0.015526 | 0.004981 | 0.009024 |
| time_since_greq_25_splag_1_1_ged_best_ns | 0.012653 | 0.014484 | 0.014446 | 0.017404 | 0.021645 | 0.023414 |
| time_since_greq_25_ged_best_sb | 0.012575 | 0.012748 | 0.012856 | 0.014172 | 0.015850 | 0.016482 |
| tlag_8_greq_25_ged_best_ns | 0.012476 | 0.007724 | 0.009002 | 0.005304 | 0.003705 | 0.003273 |
| time_since_greq_100_splag_1_1_ged_best_sb | 0.012301 | 0.013264 | 0.013706 | 0.014389 | 0.016666 | 0.017529 |
| tlag_8_greq_5_ged_best_ns | 0.011987 | 0.013113 | 0.005225 | 0.006836 | 0.005029 | 0.005600 |

#### Change history

The model has been in use since r.2020.02.01

### B-3.3 Onset (onset_24_25_all)

#### Description

The model includes all the features of all *cm* models, but trained on onset as the dependent variable. Onset is a binary variable coded as 1 for the first month in a two year period with at least 25 battle-related deaths (BRDs) (i.e. the previous 24 months experienced no violent episodes with at least 25 BRDs).

#### Estimation details

The model uses the ‘one-step-ahead’ approach (see Section B-2) and is estimated on data for Africa only.

#### Feature importances

The tables below report the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). Please note that in broad models like *Onset*, the tables report only the first 25 most important features, which not surprisingly coincide with conflict history variables.

*Table B-6.* Feature importances for the Onset_24_25_all model, **sb**

| sb_onset24_25_all | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_25_ged_best_sb | 0.025211 | 0.019649 | 0.017941 | 0.016636 | 0.014555 | 0.014569 |
| time_since_greq_5_ged_best_sb | 0.019907 | 0.018209 | 0.016615 | 0.015303 | 0.013470 | 0.013570 |
| time_since_greq_100_ged_best_sb | 0.017649 | 0.016271 | 0.016124 | 0.016340 | 0.015063 | 0.015485 |
| reign_precip | 0.017369 | 0.015358 | 0.016726 | 0.016690 | 0.017049 | 0.017835 |
| time_since_ged_dummy_sb | 0.017040 | 0.015450 | 0.015052 | 0.015057 | 0.014203 | 0.013802 |
| reign_lastelection | 0.015951 | 0.015683 | 0.015813 | 0.016595 | 0.013977 | 0.014775 |
| time_since_greq_500_splag_1_1_ged_best_ns | 0.015705 | 0.015677 | 0.016361 | 0.018063 | 0.016555 | 0.016814 |
| reign_irregular | 0.015392 | 0.014639 | 0.015601 | 0.016157 | 0.014801 | 0.015964 |
| reign_couprisk | 0.015342 | 0.016538 | 0.013848 | 0.016711 | 0.015701 | 0.016779 |
| reign_pctile_risk | 0.015261 | 0.016594 | 0.015183 | 0.017318 | 0.016420 | 0.016908 |
| time_since_greq_25_ged_best_os | 0.015064 | 0.014836 | 0.014383 | 0.014793 | 0.014903 | 0.015167 |
| reign_tenure_months | 0.015029 | 0.013772 | 0.014567 | 0.015823 | 0.015196 | 0.013554 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.015020 | 0.013964 | 0.013941 | 0.015121 | 0.013236 | 0.015238 |
| time_since_greq_500_ged_best_os | 0.014981 | 0.016621 | 0.017315 | 0.016088 | 0.016103 | 0.018065 |
| time_since_greq_500_ged_best_sb | 0.014937 | 0.016150 | 0.017594 | 0.017547 | 0.017100 | 0.017005 |
| time_since_ged_dummy_os | 0.014804 | 0.013718 | 0.013353 | 0.015271 | 0.013509 | 0.013726 |
| time_since_greq_100_ged_best_ns | 0.014712 | 0.015607 | 0.015073 | 0.016798 | 0.016572 | 0.015625 |
| time_since_greq_500_ged_best_ns | 0.014622 | 0.016410 | 0.017998 | 0.017153 | 0.017202 | 0.018097 |
| time_since_greq_5_ged_best_os | 0.014507 | 0.014210 | 0.014608 | 0.014575 | 0.015412 | 0.015019 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.014291 | 0.015184 | 0.017504 | 0.015519 | 0.014873 | 0.015652 |
| time_since_greq_100_ged_best_os | 0.014289 | 0.015292 | 0.015249 | 0.016358 | 0.015821 | 0.016036 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.013858 | 0.014438 | 0.015544 | 0.015646 | 0.014900 | 0.016372 |
| reign_loss | 0.013834 | 0.014899 | 0.014892 | 0.015589 | 0.015874 | 0.015915 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.013619 | 0.013950 | 0.013866 | 0.015730 | 0.013220 | 0.014406 |
| time_since_greq_5_ged_best_ns | 0.013526 | 0.014855 | 0.015398 | 0.015975 | 0.015138 | 0.015558 |

*Table B-7.* Feature importances for the Onset_24_25_all model, **os**

| os_onset24_25_all | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_25_ged_best_os | 0.023960 | 0.018528 | 0.016875 | 0.015535 | 0.014766 | 0.015163 |
| time_since_greq_5_ged_best_os | 0.017874 | 0.014260 | 0.012809 | 0.013950 | 0.013808 | 0.014608 |
| time_since_greq_100_ged_best_os | 0.015962 | 0.015915 | 0.014483 | 0.015441 | 0.015964 | 0.015089 |
| time_since_greq_500_splag_1_1_ged_best_ns | 0.015298 | 0.015821 | 0.014388 | 0.016922 | 0.016191 | 0.017022 |
| time_since_greq_500_ged_best_sb | 0.015061 | 0.016574 | 0.015917 | 0.016806 | 0.015877 | 0.016232 |
| reign_irregular | 0.014941 | 0.015797 | 0.015997 | 0.016318 | 0.014440 | 0.013962 |
| time_since_greq_100_ged_best_sb | 0.014909 | 0.014013 | 0.014034 | 0.015848 | 0.014803 | 0.015309 |
| reign_lastelection | 0.014768 | 0.016403 | 0.015255 | 0.014855 | 0.015271 | 0.013250 |
| reign_precip | 0.014740 | 0.017477 | 0.016314 | 0.016272 | 0.016148 | 0.015663 |
| reign_pctile_risk | 0.014729 | 0.015312 | 0.015196 | 0.015089 | 0.015837 | 0.018589 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.014314 | 0.014954 | 0.016066 | 0.017143 | 0.014589 | 0.015737 |
| reign_tenure_months | 0.014281 | 0.016556 | 0.016349 | 0.015893 | 0.014248 | 0.014917 |
| time_since_greq_500_ged_best_os | 0.014245 | 0.014334 | 0.014347 | 0.015817 | 0.016403 | 0.016102 |
| reign_loss | 0.014106 | 0.015031 | 0.015632 | 0.016488 | 0.014242 | 0.014486 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.013996 | 0.014590 | 0.015897 | 0.017105 | 0.016104 | 0.017654 |
| time_since_greq_500_ged_best_ns | 0.013975 | 0.014902 | 0.015827 | 0.015907 | 0.016243 | 0.017016 |
| reign_couprisk | 0.013956 | 0.015612 | 0.015989 | 0.014992 | 0.016487 | 0.019175 |
| time_since_greq_100_ged_best_ns | 0.013886 | 0.015127 | 0.015007 | 0.015742 | 0.015132 | 0.015441 |
| time_since_greq_25_ged_best_ns | 0.013754 | 0.013814 | 0.013351 | 0.013756 | 0.014391 | 0.013845 |
| time_since_greq_5_ged_best_ns | 0.013423 | 0.013683 | 0.013377 | 0.012972 | 0.014481 | 0.013541 |
| time_since_greq_25_ged_best_sb | 0.013063 | 0.013896 | 0.014475 | 0.014130 | 0.014332 | 0.012618 |
| time_since_ged_dummy_os | 0.012968 | 0.012480 | 0.011634 | 0.012103 | 0.012847 | 0.013036 |
| time_since_greq_100_splag_1_1_ged_best_sb | 0.012833 | 0.013008 | 0.012962 | 0.013766 | 0.012524 | 0.012771 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.012801 | 0.013023 | 0.013101 | 0.014378 | 0.012989 | 0.013868 |
| time_since_ged_dummy_ns | 0.012679 | 0.012621 | 0.012177 | 0.013140 | 0.013292 | 0.013178 |

### Change history

The model has been in use since r.2020.02.01

## B-3.4 REIGN coups (reign_coups)

### Description

This model uses the predicted probability of coups from REIGN as the main feature for predicting conflict incidence. Coup d'états and civil wars tend to co-occur in many countries, and the two phenomena are

*Table B-8.* Feature importances for the Onset_24_25_all model, **ns**

| ns_onset24_25_all | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_25_ged_best_ns | 0.020169 | 0.015984 | 0.013881 | 0.014906 | 0.015085 | 0.014663 |
| time_since_greq_5_ged_best_ns | 0.016712 | 0.018606 | 0.014755 | 0.014466 | 0.013996 | 0.015589 |
| reign_precip | 0.015518 | 0.016781 | 0.018307 | 0.013461 | 0.014053 | 0.013960 |
| time_since_greq_500_ged_best_ns | 0.015226 | 0.014395 | 0.015300 | 0.014706 | 0.017029 | 0.017360 |
| reign_pctile_risk | 0.014793 | 0.014284 | 0.015798 | 0.013466 | 0.015985 | 0.017924 |
| reign_couprisk | 0.014748 | 0.013594 | 0.015551 | 0.012523 | 0.015930 | 0.018319 |
| time_since_ged_dummy_ns | 0.014742 | 0.014860 | 0.013900 | 0.013752 | 0.013896 | 0.013056 |
| time_since_greq_500_ged_best_os | 0.014654 | 0.015761 | 0.016379 | 0.015281 | 0.016148 | 0.016025 |
| time_since_greq_500_splag_1_1_ged_best_ns | 0.014484 | 0.015886 | 0.016636 | 0.016063 | 0.016791 | 0.016253 |
| time_since_greq_100_ged_best_ns | 0.014420 | 0.014977 | 0.014014 | 0.014258 | 0.015083 | 0.016240 |
| reign_tenure_months | 0.014283 | 0.013673 | 0.016176 | 0.017595 | 0.015817 | 0.014742 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.014042 | 0.013728 | 0.013023 | 0.015359 | 0.015296 | 0.015143 |
| time_since_greq_100_ged_best_os | 0.014013 | 0.014298 | 0.015427 | 0.017006 | 0.016305 | 0.015426 |
| time_since_greq_500_ged_best_sb | 0.013895 | 0.013840 | 0.014913 | 0.016454 | 0.014899 | 0.015497 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.013383 | 0.013297 | 0.014166 | 0.014854 | 0.013576 | 0.015242 |
| reign_irregular | 0.013309 | 0.014732 | 0.015341 | 0.015205 | 0.016493 | 0.015447 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.012824 | 0.013612 | 0.012908 | 0.012651 | 0.013620 | 0.013419 |
| time_since_greq_100_ged_best_sb | 0.012776 | 0.012377 | 0.013948 | 0.014139 | 0.014409 | 0.014442 |
| time_since_greq_25_ged_best_os | 0.012767 | 0.014560 | 0.014960 | 0.014105 | 0.014932 | 0.015694 |
| reign_lastelection | 0.012762 | 0.013870 | 0.016409 | 0.014535 | 0.015912 | 0.014772 |
| reign_loss | 0.012373 | 0.013901 | 0.015595 | 0.015547 | 0.015945 | 0.015166 |
| time_since_greq_5_ged_best_os | 0.011722 | 0.012853 | 0.013466 | 0.014729 | 0.013152 | 0.012758 |
| time_since_greq_25_ged_best_sb | 0.011573 | 0.011421 | 0.011716 | 0.013338 | 0.013316 | 0.014875 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.011427 | 0.012609 | 0.013155 | 0.013595 | 0.014566 | 0.014558 |
| time_since_greq_5_ged_best_sb | 0.011072 | 0.011220 | 0.011147 | 0.011217 | 0.010896 | 0.012725 |

inextricably linked (Bell and Sudduth, 2017). The main variable is the predicted monthly probability of a coup against a leader, drawn from CoupCast, a monthly forecast of coup risk in all UN member states. These data are updated each month by the One Earth Future Foundation and are available for download from `www.oefresearch.org` (Bell, 2016a).

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated using data for Africa only.

### Feature importances

The table below reports the feature importances by each type of violence. The importance of all features show a constant decrease over time (from step 1 to step 36), signaling the rapid influence of electoral transitions and political leadership changes on conflict.

*Table B-9.* Feature importances for the Reign_coups model, **sb**, **os**, **ns**

| sb_reign_coups | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| reign_pctile_risk | 0.557373 | 0.555865 | 0.555168 | 0.556641 | 0.551278 | 0.548092 |
| reign_couprisk | 0.442627 | 0.444135 | 0.444832 | 0.443359 | 0.448722 | 0.451908 |

| os_reign_coups | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| reign_pctile_risk | 0.54853 | 0.5499 | 0.55264 | 0.552382 | 0.549659 | 0.546929 |
| reign_couprisk | 0.45147 | 0.4501 | 0.44736 | 0.447618 | 0.450341 | 0.453071 |

| ns_reign_coups | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| reign_pctile_risk | 0.562262 | 0.568115 | 0.56833 | 0.567677 | 0.557052 | 0.558601 |
| reign_couprisk | 0.437738 | 0.431885 | 0.43167 | 0.432323 | 0.442948 | 0.441399 |

### Change history

The model has been in use since r.2020.02.01

## B-3.5 ACLED protest (acled_protest)

### Description

The model includes variables capturing the recent history of protests in each country, drawn from the ACLED dataset (Raleigh et al., 2010).

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

### Feature importances

The following table summarizes the feature importances for each predicted outcome. Similar to the patterns underlined above, the proximity to conflict episodes in time is more important than the spatial one to predict violence. Moreover, the importance of conflict history slightly increases over time (from step 1 to 36), unlike spatial lags' relative importance.

*Table B-10.* Feature importances for the Acled_protest model, **sb**, **os**, **ns**

| sb_acled_protest | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_acled_dummy_pr | 0.518372 | 0.524272 | 0.527706 | 0.538298 | 0.515977 | 0.551286 |
| splag_1_1_acled_count_pr | 0.350453 | 0.344519 | 0.338507 | 0.322718 | 0.354980 | 0.329880 |
| time_since_splag_1_1_acled_dummy_pr | 0.131175 | 0.131209 | 0.133786 | 0.138985 | 0.129043 | 0.118835 |
| **os_acled_protest** | **1** | **3** | **6** | **12** | **24** | **36** |
| time_since_acled_dummy_pr | 0.515508 | 0.531588 | 0.549461 | 0.542385 | 0.490420 | 0.532951 |
| splag_1_1_acled_count_pr | 0.364981 | 0.352770 | 0.341938 | 0.351283 | 0.406311 | 0.365970 |
| time_since_splag_1_1_acled_dummy_pr | 0.119511 | 0.115643 | 0.108602 | 0.106332 | 0.103269 | 0.101079 |
| **ns_acled_protest** | **1** | **3** | **6** | **12** | **24** | **36** |
| time_since_acled_dummy_pr | 0.497855 | 0.490727 | 0.495670 | 0.501213 | 0.499997 | 0.507667 |
| splag_1_1_acled_count_pr | 0.392286 | 0.397450 | 0.392958 | 0.379011 | 0.376732 | 0.363237 |
| time_since_splag_1_1_acled_dummy_pr | 0.109860 | 0.111822 | 0.111372 | 0.119776 | 0.123270 | 0.129096 |

### Change history

The model has been in use since r.2020.02.01

## B-3.6 Country dummies (cdummies)

### Description

A model with only country dummies, a random-forest variant of the random-effects model. The numbers refer to the ViEWS internal country numbering system. Some selected countries:

| | |
|---|---|
| 53 | Sierra Leone |
| 57 | Ethiopia |
| 67 | Algeria |
| 70 | Central African Republic |
| 79 | Nigeria |
| 120 | Somalia |
| 155 | Burundi |
| 163 | South Africa |
| 167 | Democratic Republic of Congo |
| 213 | Libya |
| 235 | Uganda |

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

### Feature importances

The tables below report the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$).

*Table B-11.* Feature importances for the Cdummies model, **sb**

| sb_cdummies | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| cdum_120 | 0.266935 | 0.265544 | 0.264838 | 0.264285 | 0.264430 | 0.267143 |
| cdum_67 | 0.213026 | 0.212320 | 0.213487 | 0.213790 | 0.213163 | 0.213687 |
| cdum_167 | 0.096939 | 0.097385 | 0.097363 | 0.097145 | 0.097253 | 0.097600 |
| cdum_155 | 0.060688 | 0.059616 | 0.060083 | 0.059396 | 0.060408 | 0.060271 |
| cdum_79 | 0.060195 | 0.061116 | 0.059581 | 0.060015 | 0.060965 | 0.060185 |
| cdum_235 | 0.056087 | 0.055595 | 0.056146 | 0.056459 | 0.056494 | 0.056073 |
| cdum_165 | 0.046653 | 0.047400 | 0.046400 | 0.046469 | 0.046840 | 0.046481 |
| cdum_53 | 0.022525 | 0.022378 | 0.022708 | 0.022723 | 0.022682 | 0.022672 |
| cdum_245 | 0.008878 | 0.008790 | 0.009013 | 0.008914 | 0.008876 | 0.008801 |
| cdum_214 | 0.008528 | 0.008573 | 0.008325 | 0.008437 | 0.008827 | 0.008506 |
| cdum_50 | 0.007597 | 0.007206 | 0.007043 | 0.006936 | 0.007126 | 0.007235 |
| cdum_222 | 0.005721 | 0.005591 | 0.005548 | 0.005562 | 0.005798 | 0.005523 |
| cdum_246 | 0.005654 | 0.005620 | 0.005795 | 0.005667 | 0.005710 | 0.005679 |
| cdum_213 | 0.005036 | 0.004850 | 0.005017 | 0.004784 | 0.004891 | 0.004820 |
| cdum_163 | 0.004506 | 0.004532 | 0.004313 | 0.004594 | 0.004293 | 0.004392 |
| cdum_244 | 0.004478 | 0.004425 | 0.004239 | 0.004520 | 0.004223 | 0.004394 |
| cdum_164 | 0.004444 | 0.004449 | 0.004238 | 0.004458 | 0.004371 | 0.004277 |
| cdum_57 | 0.004435 | 0.004206 | 0.004399 | 0.004057 | 0.004123 | 0.004197 |
| cdum_158 | 0.004413 | 0.004379 | 0.004315 | 0.004398 | 0.004293 | 0.004293 |
| cdum_42 | 0.004400 | 0.004420 | 0.004331 | 0.004451 | 0.004580 | 0.004375 |
| cdum_74 | 0.004393 | 0.004570 | 0.004407 | 0.004387 | 0.004330 | 0.004324 |
| cdum_242 | 0.004371 | 0.004419 | 0.004373 | 0.004371 | 0.004334 | 0.004361 |
| cdum_81 | 0.004344 | 0.004432 | 0.004280 | 0.004363 | 0.004281 | 0.004534 |
| cdum_154 | 0.004339 | 0.004299 | 0.004333 | 0.004487 | 0.004278 | 0.004376 |
| cdum_161 | 0.004313 | 0.004381 | 0.004227 | 0.004525 | 0.004474 | 0.004419 |

*Table B-12.* Feature importances for the Cdummies model, **os**

| os_cdummies | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| cdum_167 | 0.421143 | 0.425890 | 0.429658 | 0.439213 | 0.455289 | 0.477486 |
| cdum_163 | 0.215561 | 0.209713 | 0.199745 | 0.183001 | 0.150060 | 0.117827 |
| cdum_79 | 0.093096 | 0.093432 | 0.095760 | 0.096866 | 0.102317 | 0.105743 |
| cdum_53 | 0.039877 | 0.040296 | 0.040955 | 0.041726 | 0.043412 | 0.045874 |
| cdum_155 | 0.031001 | 0.031299 | 0.032166 | 0.032184 | 0.034447 | 0.035221 |
| cdum_70 | 0.029834 | 0.029845 | 0.030463 | 0.030909 | 0.032903 | 0.034239 |
| cdum_235 | 0.025539 | 0.026058 | 0.026812 | 0.026702 | 0.028349 | 0.029636 |
| cdum_43 | 0.012155 | 0.012302 | 0.012406 | 0.012698 | 0.013539 | 0.013890 |
| cdum_156 | 0.011517 | 0.011868 | 0.011531 | 0.011831 | 0.012267 | 0.012767 |
| cdum_165 | 0.003764 | 0.003964 | 0.004002 | 0.004241 | 0.004389 | 0.004265 |
| cdum_214 | 0.003757 | 0.003946 | 0.004134 | 0.004074 | 0.003973 | 0.004296 |
| cdum_162 | 0.003683 | 0.003848 | 0.003976 | 0.004163 | 0.004415 | 0.004388 |
| cdum_157 | 0.003520 | 0.003343 | 0.003480 | 0.003529 | 0.003789 | 0.003510 |
| cdum_164 | 0.003517 | 0.003293 | 0.003548 | 0.003463 | 0.003587 | 0.003580 |
| cdum_49 | 0.003517 | 0.003308 | 0.003341 | 0.003470 | 0.003591 | 0.003535 |
| cdum_174 | 0.003493 | 0.003270 | 0.003395 | 0.003435 | 0.003749 | 0.003617 |
| cdum_42 | 0.003477 | 0.003261 | 0.003650 | 0.003522 | 0.003360 | 0.003564 |
| cdum_56 | 0.003462 | 0.003381 | 0.003472 | 0.003649 | 0.003551 | 0.003545 |
| cdum_173 | 0.003460 | 0.003451 | 0.003452 | 0.003404 | 0.003648 | 0.003477 |
| cdum_154 | 0.003458 | 0.003291 | 0.003286 | 0.003447 | 0.003600 | 0.003413 |
| cdum_54 | 0.003456 | 0.003507 | 0.003526 | 0.003598 | 0.003575 | 0.003575 |
| cdum_170 | 0.003455 | 0.003768 | 0.003632 | 0.003491 | 0.003554 | 0.003527 |
| cdum_80 | 0.003452 | 0.003477 | 0.003354 | 0.003559 | 0.003610 | 0.003656 |
| cdum_55 | 0.003429 | 0.003338 | 0.003494 | 0.003667 | 0.003651 | 0.003735 |
| cdum_74 | 0.003405 | 0.003434 | 0.003274 | 0.003387 | 0.003894 | 0.003637 |

*Table B-13.* Feature importances for the Cdummies model, **ns**

| ns_cdummies | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| cdum_163 | 0.344979 | 0.340596 | 0.330492 | 0.311798 | 0.278586 | 0.244426 |
| cdum_79 | 0.274122 | 0.275546 | 0.281585 | 0.288281 | 0.303871 | 0.317834 |
| cdum_120 | 0.131738 | 0.133675 | 0.134461 | 0.137067 | 0.144943 | 0.151994 |
| cdum_167 | 0.088675 | 0.089070 | 0.088926 | 0.092553 | 0.097045 | 0.101545 |
| cdum_57 | 0.024640 | 0.024246 | 0.024889 | 0.026135 | 0.027692 | 0.028403 |
| cdum_213 | 0.022650 | 0.022654 | 0.022568 | 0.023381 | 0.025131 | 0.026246 |
| cdum_237 | 0.020733 | 0.020102 | 0.020824 | 0.021384 | 0.022593 | 0.023705 |
| cdum_245 | 0.010131 | 0.009852 | 0.010126 | 0.010818 | 0.011373 | 0.012007 |
| cdum_70 | 0.010046 | 0.010013 | 0.010163 | 0.010570 | 0.011452 | 0.011804 |
| cdum_246 | 0.007781 | 0.007484 | 0.007921 | 0.008045 | 0.008594 | 0.008561 |
| cdum_235 | 0.002039 | 0.002007 | 0.002126 | 0.002275 | 0.002090 | 0.002131 |
| cdum_50 | 0.001869 | 0.001998 | 0.002094 | 0.001967 | 0.002209 | 0.002210 |
| cdum_53 | 0.001863 | 0.002017 | 0.001885 | 0.001935 | 0.001828 | 0.001883 |
| cdum_54 | 0.001863 | 0.001843 | 0.001787 | 0.001865 | 0.001691 | 0.002013 |
| cdum_166 | 0.001848 | 0.001745 | 0.001734 | 0.001783 | 0.001721 | 0.001938 |
| cdum_74 | 0.001830 | 0.001673 | 0.001725 | 0.001985 | 0.001824 | 0.001932 |
| cdum_170 | 0.001802 | 0.001727 | 0.001800 | 0.001780 | 0.001721 | 0.002005 |
| cdum_165 | 0.001767 | 0.001813 | 0.001766 | 0.001755 | 0.001817 | 0.001925 |
| cdum_174 | 0.001756 | 0.001642 | 0.001667 | 0.001753 | 0.001732 | 0.001880 |
| cdum_78 | 0.001734 | 0.001725 | 0.001811 | 0.001794 | 0.001832 | 0.001979 |
| cdum_76 | 0.001730 | 0.001938 | 0.001789 | 0.001754 | 0.001742 | 0.001986 |
| cdum_243 | 0.001716 | 0.001820 | 0.001748 | 0.001840 | 0.001928 | 0.001849 |
| cdum_156 | 0.001715 | 0.001747 | 0.001634 | 0.001848 | 0.001792 | 0.001913 |
| cdum_73 | 0.001705 | 0.001756 | 0.001758 | 0.001779 | 0.001711 | 0.001983 |
| cdum_161 | 0.001682 | 0.001672 | 0.001758 | 0.001761 | 0.001855 | 0.001842 |

#### Change history

The model has been in use since r.2020.02.01

### B-3.7 Demography (demog)

#### Description

A demographic model identical to the one in Hegre et al. (2019).

#### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

#### Feature importances

The following table summarizes the feature importances by outcome and for different time steps. Population size is the most relevant feature in predicting all violence types. The share of urban population is especially influential in predicting **sb** and **ns** conflicts, while education is slightly more important in predicting one-sided violence.

#### Change history

The model has been in use since r.2018.07.01

### B-3.8 International Crisis Group: Crisis Watch (icgcw)

#### Description

This model makes use of the warnings issued monthly by the International Crisis Group's Crisis Watch (`https://www.crisisgroup.org/crisiswatch`).

*Table B-14.* Feature importances for the Demog model, **sb**, **os**, **ns**

| sb_demog | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| fvp_population200 | 0.219894 | 0.220324 | 0.218509 | 0.218396 | 0.225867 | 0.228540 |
| fvp_ssp2_urban_share_iiasa | 0.179895 | 0.178867 | 0.181906 | 0.175815 | 0.173086 | 0.169568 |
| fvp_ssp2_edu_sec_15_24_prop | 0.161832 | 0.163793 | 0.161504 | 0.163617 | 0.160836 | 0.157831 |
| wdi_sp_dyn_imrt_in | 0.152002 | 0.148188 | 0.148372 | 0.145445 | 0.144429 | 0.143186 |
| fvp_grpop200 | 0.149710 | 0.151041 | 0.153467 | 0.161165 | 0.159468 | 0.160229 |
| wdi_sp_dyn_tfrt_in | 0.136668 | 0.137787 | 0.136242 | 0.135562 | 0.136313 | 0.140646 |

| os_demog | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| fvp_population200 | 0.249454 | 0.254186 | 0.250858 | 0.249368 | 0.243952 | 0.240763 |
| fvp_ssp2_edu_sec_15_24_prop | 0.168157 | 0.169320 | 0.164902 | 0.161208 | 0.146488 | 0.150925 |
| fvp_ssp2_urban_share_iiasa | 0.167606 | 0.160710 | 0.167648 | 0.164116 | 0.165587 | 0.158352 |
| wdi_sp_dyn_tfrt_in | 0.159228 | 0.158385 | 0.149055 | 0.151917 | 0.151946 | 0.150024 |
| wdi_sp_dyn_imrt_in | 0.130064 | 0.129945 | 0.133557 | 0.133994 | 0.144374 | 0.151877 |
| fvp_grpop200 | 0.125490 | 0.127454 | 0.133979 | 0.139396 | 0.147654 | 0.148059 |

| ns_demog | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| fvp_population200 | 0.283978 | 0.280631 | 0.288608 | 0.282503 | 0.295437 | 0.293703 |
| fvp_ssp2_urban_share_iiasa | 0.163848 | 0.165097 | 0.163748 | 0.168963 | 0.161231 | 0.157347 |
| fvp_grpop200 | 0.149020 | 0.152063 | 0.151692 | 0.150851 | 0.143195 | 0.130046 |
| fvp_ssp2_edu_sec_15_24_prop | 0.138737 | 0.137126 | 0.137894 | 0.138222 | 0.137714 | 0.141107 |
| wdi_sp_dyn_imrt_in | 0.133961 | 0.131202 | 0.129395 | 0.124002 | 0.121110 | 0.125100 |
| wdi_sp_dyn_tfrt_in | 0.130457 | 0.133881 | 0.128662 | 0.135459 | 0.141313 | 0.152698 |

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

### Feature importances

The table below reports the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). As expected, the importance of these rapidly moving features decreases over time, which explains why they may be especially relevant in capturing early signs of violence.

*Table B-15.* Feature importances for the Icgw model, **sb**, **os**, **ns**

| sb_icgcw | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| icgcw_alerts | 0.410876 | 0.328994 | 0.413471 | 0.238607 | 0.318129 | 0.181533 |
| icgcw_unobserved | 0.259145 | 0.269542 | 0.251464 | 0.356391 | 0.322079 | 0.363788 |
| icgcw_improved | 0.249332 | 0.277421 | 0.255746 | 0.335005 | 0.311468 | 0.408719 |
| icgcw_deteriorated | 0.052536 | 0.051155 | 0.055869 | 0.039632 | 0.023919 | 0.028193 |
| icgcw_opportunities | 0.028111 | 0.072888 | 0.023450 | 0.030366 | 0.024404 | 0.017766 |

| os_icgcw | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| icgcw_alerts | 0.342203 | 0.296928 | 0.249408 | 0.147917 | 0.125652 | 0.112650 |
| icgcw_unobserved | 0.285304 | 0.341172 | 0.363936 | 0.422891 | 0.550268 | 0.556464 |
| icgcw_deteriorated | 0.206761 | 0.161404 | 0.190461 | 0.175204 | 0.097504 | 0.093204 |
| icgcw_improved | 0.135719 | 0.142596 | 0.173903 | 0.200493 | 0.206836 | 0.221738 |
| icgcw_opportunities | 0.030013 | 0.057901 | 0.022292 | 0.053496 | 0.019740 | 0.015944 |

| ns_icgcw | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| icgcw_deteriorated | 0.434510 | 0.477854 | 0.392844 | 0.472816 | 0.410794 | 0.393297 |
| icgcw_alerts | 0.282320 | 0.190265 | 0.286608 | 0.176645 | 0.143078 | 0.113024 |
| icgcw_improved | 0.144947 | 0.148498 | 0.159494 | 0.197637 | 0.242158 | 0.314608 |
| icgcw_unobserved | 0.109981 | 0.113704 | 0.103307 | 0.134287 | 0.166764 | 0.152908 |
| icgcw_opportunities | 0.028242 | 0.069678 | 0.057748 | 0.018614 | 0.037205 | 0.026163 |

### Change history

The model has been in use since r.2020.02.01

## B-3.9 Neighbor history (neibhist)

### Description

A model capturing the conflict history in neighboring countries, using a subset of the features in the *cflong* model. As conflict tends to cluster both in space and time, areas close to countries in conflict are more likely to experience violence (Buhaug and Gleditsch, 2008).

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

### Feature importances

The tables below report the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). Please note that in broad models like *neibhist*, the tables report only the first 25 most important features, which not surprisingly coincide with conflict history variables. Time since any of the types of violence exceeded 500 deaths is the most important feature in predicting all violence types, suggesting that high intensity conflicts, which are likely to have more disruptive effects in a country, are especially influential in predicting violence in neighboring areas. This may suggest that migratory flows in the aftermath of a major conflict would be a worth exploring avenue for future research.

*Table B-16.* Feature importances for the Neibhist model, **sb**, **os**, **ns**

| sb_neibhist | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_500_splag_1_1_ged_best_ns | 0.076841 | 0.076832 | 0.076111 | 0.074613 | 0.073476 | 0.073593 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.075501 | 0.076080 | 0.073790 | 0.071938 | 0.066879 | 0.064614 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.072749 | 0.073570 | 0.072877 | 0.072662 | 0.075975 | 0.075340 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.061189 | 0.061273 | 0.062534 | 0.062490 | 0.061371 | 0.060594 |
| splag_1_1_acled_count_os | 0.058712 | 0.060632 | 0.062985 | 0.064393 | 0.072850 | 0.078522 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.056703 | 0.057463 | 0.057135 | 0.057121 | 0.057184 | 0.057491 |
| time_since_greq_100_splag_1_1_ged_best_sb | 0.056575 | 0.055827 | 0.055034 | 0.056697 | 0.058403 | 0.058144 |
| splag_1_1_acled_count_ns | 0.042850 | 0.042162 | 0.042840 | 0.041951 | 0.044177 | 0.045202 |
| splag_1_1_acled_count_sb | 0.042408 | 0.041881 | 0.041650 | 0.042628 | 0.046996 | 0.047805 |
| time_since_greq_5_splag_1_1_ged_best_ns | 0.042121 | 0.040240 | 0.040841 | 0.040658 | 0.040373 | 0.042686 |
| time_since_greq_25_splag_1_1_ged_best_ns | 0.040012 | 0.038241 | 0.038270 | 0.038099 | 0.040879 | 0.040658 |
| time_since_greq_25_splag_1_1_ged_best_os | 0.039585 | 0.040298 | 0.039546 | 0.038677 | 0.039083 | 0.038770 |
| time_since_greq_25_splag_1_1_ged_best_sb | 0.039094 | 0.039013 | 0.039185 | 0.038500 | 0.035794 | 0.036869 |
| splag_1_1_ged_best_sb | 0.036108 | 0.035692 | 0.034379 | 0.036499 | 0.034974 | 0.032654 |
| splag_1_1_ged_best_os | 0.034776 | 0.035412 | 0.034452 | 0.035094 | 0.032930 | 0.031510 |
| time_since_greq_5_splag_1_1_ged_best_sb | 0.034365 | 0.035153 | 0.036012 | 0.035884 | 0.036081 | 0.034081 |
| splag_1_1_tlag_1_ged_best_sb | 0.034125 | 0.033339 | 0.032932 | 0.033359 | 0.032035 | 0.031193 |
| splag_1_1_tlag_1_ged_best_os | 0.033810 | 0.032380 | 0.033664 | 0.032885 | 0.030535 | 0.028372 |
| splag_1_1_ged_best_ns | 0.028081 | 0.027182 | 0.026847 | 0.027731 | 0.022664 | 0.023123 |
| time_since_greq_5_splag_1_1_ged_best_os | 0.026749 | 0.027815 | 0.028167 | 0.028097 | 0.029082 | 0.029261 |
| splag_1_1_tlag_1_ged_best_ns | 0.025354 | 0.026062 | 0.026273 | 0.023443 | 0.021521 | 0.020407 |
| time_since_splag_1_1_acled_dummy_ns | 0.018850 | 0.018899 | 0.019642 | 0.020475 | 0.020999 | 0.021330 |
| time_since_splag_1_1_acled_dummy_sb | 0.013749 | 0.013763 | 0.013823 | 0.014557 | 0.015337 | 0.016399 |
| time_since_splag_1_1_acled_dummy_os | 0.009694 | 0.010793 | 0.011009 | 0.011550 | 0.010400 | 0.011381 |

*Table B-17.* Feature importances for the Neibhist model, **os**

| os_neibhist | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_500_splag_1_1_ged_best_ns | 0.082015 | 0.082010 | 0.080501 | 0.080713 | 0.080371 | 0.078190 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.071508 | 0.071852 | 0.070455 | 0.067823 | 0.065866 | 0.062419 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.069335 | 0.069491 | 0.068805 | 0.067968 | 0.068655 | 0.066761 |
| splag_1_1_acled_count_os | 0.068320 | 0.068225 | 0.067154 | 0.074163 | 0.076464 | 0.076879 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.058578 | 0.059387 | 0.060364 | 0.058275 | 0.060471 | 0.061200 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.048565 | 0.049100 | 0.050210 | 0.049619 | 0.050040 | 0.051252 |
| splag_1_1_acled_count_sb | 0.047319 | 0.046761 | 0.045018 | 0.044185 | 0.050458 | 0.053303 |
| time_since_greq_25_splag_1_1_ged_best_ns | 0.046477 | 0.045033 | 0.043807 | 0.042162 | 0.047613 | 0.051959 |
| splag_1_1_ged_best_os | 0.045528 | 0.048846 | 0.046166 | 0.049533 | 0.039203 | 0.039602 |
| time_since_greq_100_splag_1_1_ged_best_sb | 0.045260 | 0.045218 | 0.045242 | 0.045187 | 0.048539 | 0.048629 |
| splag_1_1_acled_count_ns | 0.043312 | 0.041533 | 0.042147 | 0.044177 | 0.043878 | 0.042302 |
| splag_1_1_tlag_1_ged_best_os | 0.043020 | 0.042607 | 0.043955 | 0.040617 | 0.038796 | 0.035012 |
| splag_1_1_ged_best_sb | 0.040345 | 0.042046 | 0.041785 | 0.041218 | 0.037906 | 0.036806 |
| splag_1_1_tlag_1_ged_best_sb | 0.039267 | 0.038365 | 0.040161 | 0.039473 | 0.036327 | 0.034206 |
| time_since_greq_25_splag_1_1_ged_best_sb | 0.037289 | 0.035767 | 0.035530 | 0.035014 | 0.033042 | 0.032824 |
| time_since_greq_5_splag_1_1_ged_best_ns | 0.036695 | 0.037224 | 0.039654 | 0.040738 | 0.041412 | 0.047040 |
| time_since_greq_5_splag_1_1_ged_best_sb | 0.030406 | 0.030117 | 0.031332 | 0.030208 | 0.034611 | 0.035579 |
| time_since_greq_25_splag_1_1_ged_best_os | 0.029782 | 0.029988 | 0.029344 | 0.030978 | 0.033017 | 0.035185 |
| splag_1_1_ged_best_ns | 0.027766 | 0.027020 | 0.027481 | 0.026192 | 0.025460 | 0.023244 |
| splag_1_1_tlag_1_ged_best_ns | 0.025607 | 0.026331 | 0.026771 | 0.025500 | 0.022298 | 0.021235 |
| time_since_greq_5_splag_1_1_ged_best_os | 0.019927 | 0.019850 | 0.019718 | 0.021097 | 0.021343 | 0.024485 |
| time_since_splag_1_1_acled_dummy_ns | 0.016888 | 0.015321 | 0.016052 | 0.015167 | 0.015773 | 0.014424 |
| time_since_splag_1_1_acled_dummy_sb | 0.014145 | 0.015379 | 0.014995 | 0.015790 | 0.015324 | 0.014692 |
| time_since_splag_1_1_acled_dummy_os | 0.012646 | 0.012531 | 0.013352 | 0.014202 | 0.013133 | 0.012772 |

*Table B-18.* Feature importances for the Neibhist model, **ns**

| ns_neibhist | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| time_since_greq_500_splag_1_1_ged_best_ns | 0.076208 | 0.076289 | 0.076965 | 0.075081 | 0.070862 | 0.065755 |
| time_since_greq_500_splag_1_1_ged_best_sb | 0.073533 | 0.071795 | 0.070271 | 0.068532 | 0.061912 | 0.059963 |
| time_since_greq_500_splag_1_1_ged_best_os | 0.071990 | 0.070728 | 0.072560 | 0.070901 | 0.071379 | 0.066311 |
| splag_1_1_acled_count_os | 0.063203 | 0.066210 | 0.063490 | 0.068858 | 0.071395 | 0.071655 |
| time_since_greq_100_splag_1_1_ged_best_ns | 0.062124 | 0.060442 | 0.060726 | 0.060757 | 0.061730 | 0.060812 |
| time_since_greq_100_splag_1_1_ged_best_os | 0.062092 | 0.062029 | 0.062884 | 0.062881 | 0.065431 | 0.065302 |
| time_since_greq_100_splag_1_1_ged_best_sb | 0.052416 | 0.054161 | 0.053675 | 0.054214 | 0.054711 | 0.056800 |
| splag_1_1_acled_count_ns | 0.045059 | 0.044861 | 0.045833 | 0.049133 | 0.043529 | 0.040821 |
| splag_1_1_acled_count_sb | 0.041756 | 0.045190 | 0.042687 | 0.042981 | 0.042931 | 0.043505 |
| splag_1_1_ged_best_sb | 0.040321 | 0.040061 | 0.041111 | 0.039177 | 0.040238 | 0.039515 |
| time_since_greq_25_splag_1_1_ged_best_os | 0.039112 | 0.037867 | 0.039225 | 0.040640 | 0.043320 | 0.045834 |
| splag_1_1_tlag_1_ged_best_sb | 0.038849 | 0.038355 | 0.037566 | 0.037285 | 0.037841 | 0.039160 |
| splag_1_1_ged_best_os | 0.037369 | 0.036311 | 0.035553 | 0.036966 | 0.035785 | 0.036134 |
| splag_1_1_tlag_1_ged_best_os | 0.035629 | 0.033457 | 0.034520 | 0.034160 | 0.032823 | 0.030430 |
| time_since_greq_25_splag_1_1_ged_best_ns | 0.035490 | 0.034769 | 0.035666 | 0.036100 | 0.039613 | 0.041778 |
| time_since_greq_5_splag_1_1_ged_best_ns | 0.032816 | 0.033681 | 0.033991 | 0.033742 | 0.035464 | 0.041362 |
| splag_1_1_ged_best_ns | 0.032525 | 0.033757 | 0.032611 | 0.029116 | 0.030266 | 0.029231 |
| splag_1_1_tlag_1_ged_best_ns | 0.031072 | 0.030128 | 0.031644 | 0.028423 | 0.026364 | 0.025618 |
| time_since_greq_25_splag_1_1_ged_best_sb | 0.029329 | 0.028436 | 0.029295 | 0.029836 | 0.029772 | 0.029716 |
| time_since_greq_5_splag_1_1_ged_best_os | 0.029224 | 0.029233 | 0.029358 | 0.030807 | 0.031778 | 0.033375 |
| time_since_greq_5_splag_1_1_ged_best_sb | 0.021729 | 0.022179 | 0.022901 | 0.022710 | 0.023489 | 0.025295 |
| time_since_splag_1_1_acled_dummy_ns | 0.019477 | 0.021270 | 0.020331 | 0.020792 | 0.022247 | 0.025422 |
| time_since_splag_1_1_acled_dummy_os | 0.014589 | 0.014208 | 0.013201 | 0.012645 | 0.012187 | 0.011563 |
| time_since_splag_1_1_acled_dummy_sb | 0.014088 | 0.014584 | 0.013938 | 0.014264 | 0.014933 | 0.014643 |

### Change history

The model has been in use since r.2020.02.01

## B-3.10 REIGN drought (reign_drought)

### Description

This model integrates the precipitation variable built in REIGN (Bell, 2016b). This variable was created by converting PREC/L gridded precipitation data (`https://www.esrl.noaa.gov/psd/data/gridded/data.precl.html`) at the 2.5 degree x 2.5 degree level to create a spatially-weighted standard precipitation index (SPI) at the country-month unit-of-analysis. The United States National Oceanic and Atmospheric Administration (NOAA) updates PREC/L each month. New scores are converted and distributed by the REIGN project of the One Earth Future Foundation. Precipitation exhibits a parabolic association to violence, whereby both extremes (drought and flood) are linked to a higher risk of coups. This is in line with recent empirical research which has shown that drought occurrence is likely to increase the risk of conflict especially in vulnerable communities, which may lack the means to cope with environmental shocks (von Uexkull et al., 2016).

#### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for Africa only.

#### Feature importances

The table below reports the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). The model includes only one feature, which is therefore assigned a relative importance of 1 (maximum).

*Table B-19.* Feature importances for the Reign_drought model, **sb**, **os**, **ns**

| sb_reign_drought | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| reign_precip | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| os_reign_drought | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| reign_precip | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| ns_reign_drought | 1 | 3 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|---|
| reign_precip | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

#### Change history

The model has been in use since r.2020.02.01

### B-3.11 Dynasim (ds_25; ds_dummy; ds_all)

#### Description

Two 'dynamic simulation' models (Hegre et al., 2019) are trained. One using the incidence of conflict with at least one BRD (*ds_dummy*), the other using the incidence of at least 25 BRDs (*ds_25*). Both models are quite similar to the 'canonical' model in Hegre et al. (2019). They are simulated together, so that their simulations inform each other. In addition, we also simulate a model with incidence of **sb**, **os**, and **ns** conflict together causing 500 deaths in a month as the outcome and include it as part of the conflict history. We do not use the final model directly in the ensemble.

#### Estimation details

The model uses the 'dynamic simulation' approach (see Section B-2) and is estimated on data for Africa only.

#### Change history

The model has been in use since r.2018.10.01

### B-3.12 All global (all_glob)

#### Description

This model includes all features used across all models. As the effect of a change occurring in one location may spatially and temporally propagate and thus lead to implications even in a remote location, global models enable ViEWS to integrate global patterns and trends into predictions, in order to better encapsulate possible non-linearities and cascade effects that occur globally. To this end, the main purpose of adding this comprehensive model to the ensemble lies in its ability to capture interactions and non-linearities between the predictors. Interactive and conditional effects may in fact be more relevant to conflict than stand-alone factors. For instance, adverse climatic conditions may be a 'threat-multiplier' (Schwartz and Randall, 2003), exacerbating pre-existing structural conditions akin poverty or ethnic fractionalization (Koubi, 2019).

#### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for all countries in the world.

#### Feature importances

The tables below report the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). Please note that in broad models like *all_glob*, the tables report only the first 25 most important features, which not surprisingly coincide with conflict history variables. Among all features at the global level, the intensity of past violent, outcome-specific, events is the most influential in predictions.

*Table B-20.* Feature importances for the all_glob model, **sb**

| sb_all_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| ged_best_sb_tlag2 | 0.022087 | 0.018478 | 0.010086 | 0.009263 | 0.011573 |
| ged_best_sb_tlag1 | 0.022042 | 0.015214 | 0.011699 | 0.013451 | 0.013395 |
| ged_months_since_last_sb_tx25 | 0.017486 | 0.015903 | 0.016580 | 0.012158 | 0.009737 |
| ged_tx_sb_25_tlag1 | 0.014977 | 0.013738 | 0.008328 | 0.012193 | 0.009972 |
| ged_months_since_last_sb_tx100 | 0.013982 | 0.013348 | 0.009985 | 0.007311 | 0.007070 |
| ged_best_sb_tlag4 | 0.013953 | 0.015877 | 0.012621 | 0.010511 | 0.008861 |
| ged_tx_sb_25_tlag2 | 0.013527 | 0.016671 | 0.012391 | 0.006794 | 0.006500 |
| ged_tx_sb_25_tlag6 | 0.013471 | 0.008958 | 0.006778 | 0.004870 | 0.004967 |
| ged_tx_sb_25_tlag5 | 0.013415 | 0.006877 | 0.008293 | 0.008755 | 0.004982 |
| ged_best_sb_tlag5 | 0.013049 | 0.012781 | 0.009093 | 0.011687 | 0.007290 |
| ged_months_since_last_sb_tx5 | 0.012905 | 0.015544 | 0.010745 | 0.006464 | 0.009037 |
| ged_best_sb_tlag7 | 0.012831 | 0.015090 | 0.009041 | 0.010707 | 0.007612 |
| ged_best_sb_tlag3 | 0.012501 | 0.015791 | 0.013683 | 0.013489 | 0.010606 |
| ged_best_sb_tlag10 | 0.011335 | 0.010032 | 0.010000 | 0.008389 | 0.008426 |
| ged_best_sb_tlag9 | 0.011027 | 0.006404 | 0.007656 | 0.008286 | 0.008147 |
| ged_tx_sb_5_tlag3 | 0.010610 | 0.006814 | 0.007674 | 0.005676 | 0.005925 |
| ged_tx_sb_25_tlag7 | 0.010532 | 0.012721 | 0.004815 | 0.005916 | 0.005200 |
| ged_tx_sb_25_tlag4 | 0.009769 | 0.008205 | 0.010669 | 0.010717 | 0.007790 |
| ged_tx_sb_5_tlag1 | 0.009582 | 0.006016 | 0.005139 | 0.010287 | 0.007367 |
| ged_tx_sb_5_tlag2 | 0.009456 | 0.006882 | 0.008577 | 0.009621 | 0.005645 |
| ged_tx_sb_5_tlag5 | 0.009014 | 0.005258 | 0.009750 | 0.004833 | 0.005521 |
| ged_best_sb_tlag6 | 0.008973 | 0.014630 | 0.011709 | 0.008657 | 0.012058 |
| ged_tx_sb_25_tlag3 | 0.008881 | 0.011904 | 0.009694 | 0.006427 | 0.006292 |
| ged_best_sb_tlag11 | 0.008375 | 0.006652 | 0.010247 | 0.006277 | 0.007016 |
| ged_tx_sb_5_tlag10 | 0.008213 | 0.003105 | 0.004366 | 0.005317 | 0.004881 |

*Table B-21.* Feature importances for the all_glob model, **os**

| os_all_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| ged_best_os_tlag1 | 0.017036 | 0.012680 | 0.011166 | 0.009043 | 0.007223 |
| ged_months_since_last_os_tx25 | 0.015480 | 0.009897 | 0.007947 | 0.008634 | 0.008492 |
| ged_best_os_tlag2 | 0.014195 | 0.009482 | 0.008189 | 0.006979 | 0.006162 |
| ged_best_os_tlag3 | 0.011676 | 0.008552 | 0.008725 | 0.007004 | 0.006908 |
| ged_months_since_last_os_tx5 | 0.010657 | 0.006722 | 0.007556 | 0.008388 | 0.007803 |
| ged_best_os_tlag5 | 0.010307 | 0.009601 | 0.007034 | 0.005417 | 0.005523 |
| ged_best_os_tlag4 | 0.010152 | 0.009927 | 0.008227 | 0.006596 | 0.005062 |
| ged_tx_os_5_tlag2 | 0.009941 | 0.004337 | 0.005856 | 0.002314 | 0.003222 |
| ged_tx_os_5_tlag1 | 0.009693 | 0.005961 | 0.004122 | 0.004054 | 0.003574 |
| ged_tx_os_25_tlag1 | 0.009303 | 0.005177 | 0.003877 | 0.002028 | 0.002292 |
| ged_best_os_tlag11 | 0.009272 | 0.006268 | 0.005989 | 0.005907 | 0.004785 |
| ged_best_os_tlag8 | 0.009238 | 0.005737 | 0.006410 | 0.007143 | 0.005095 |
| ged_best_os_tlag6 | 0.008553 | 0.009770 | 0.008693 | 0.006330 | 0.006216 |
| ged_best_os_tlag7 | 0.008369 | 0.007162 | 0.008918 | 0.005678 | 0.005408 |
| ged_best_os_tlag12 | 0.008237 | 0.006177 | 0.006616 | 0.005158 | 0.005068 |
| ged_tx_os_5_tlag3 | 0.007899 | 0.004434 | 0.006433 | 0.003978 | 0.003000 |
| ged_best_os_tlag9 | 0.007710 | 0.006977 | 0.006234 | 0.005462 | 0.005122 |
| ged_tx_os_25_tlag2 | 0.007663 | 0.004865 | 0.001851 | 0.002356 | 0.001681 |
| ged_months_since_last_os_tx100 | 0.007151 | 0.007082 | 0.008333 | 0.007131 | 0.005601 |
| ged_tx_os_25_tlag4 | 0.007049 | 0.005085 | 0.003111 | 0.002297 | 0.001332 |
| ged_best_os_tlag10 | 0.006898 | 0.006985 | 0.006174 | 0.005063 | 0.004197 |
| ged_tx_os_1_tlag1 | 0.006543 | 0.002634 | 0.002715 | 0.002951 | 0.002863 |
| ged_tx_os_25_tlag3 | 0.006095 | 0.004047 | 0.003279 | 0.002690 | 0.001444 |
| ged_tx_os_5_tlag4 | 0.005552 | 0.006271 | 0.005870 | 0.002960 | 0.002587 |
| ged_best_sb_tlag1 | 0.005274 | 0.005443 | 0.004486 | 0.005190 | 0.004228 |

*Table B-22.* Feature importances for the all_glob model, **ns**

| ns_all_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| ged_best_ns_tlag1 | 0.012406 | 0.009645 | 0.009691 | 0.005836 | 0.005152 |
| ged_months_since_last_ns_tx25 | 0.011698 | 0.008446 | 0.011169 | 0.010355 | 0.008953 |
| ged_months_since_last_ns_tx100 | 0.010978 | 0.011984 | 0.012737 | 0.008676 | 0.007867 |
| ged_best_ns_tlag2 | 0.010120 | 0.007360 | 0.009901 | 0.005824 | 0.005068 |
| ged_tx_ns_25_tlag1 | 0.009361 | 0.004305 | 0.004407 | 0.003015 | 0.002467 |
| ged_months_since_last_ns_tx5 | 0.008600 | 0.008298 | 0.007955 | 0.007928 | 0.008323 |
| ged_months_since_last_ns | 0.008504 | 0.007299 | 0.008058 | 0.007396 | 0.007417 |
| ged_tx_ns_5_tlag2 | 0.008389 | 0.004273 | 0.004247 | 0.003892 | 0.003285 |
| ged_best_ns_tlag3 | 0.008124 | 0.007365 | 0.007657 | 0.007042 | 0.005597 |
| ged_tx_ns_5_tlag1 | 0.007788 | 0.004986 | 0.005179 | 0.002687 | 0.002714 |
| ged_best_ns_tlag4 | 0.007787 | 0.006942 | 0.007209 | 0.006238 | 0.005625 |
| ged_best_ns_tlag5 | 0.007745 | 0.006064 | 0.006997 | 0.005530 | 0.005025 |
| ged_best_ns_tlag11 | 0.007450 | 0.007133 | 0.004387 | 0.004884 | 0.003908 |
| ged_tx_ns_1_tlag1 | 0.007434 | 0.005446 | 0.004515 | 0.002357 | 0.002932 |
| ged_best_ns_tlag7 | 0.007412 | 0.007203 | 0.005353 | 0.004819 | 0.005083 |
| ged_tx_ns_25_tlag2 | 0.007162 | 0.003968 | 0.006183 | 0.003992 | 0.001768 |
| ged_best_ns_tlag12 | 0.006973 | 0.005251 | 0.005122 | 0.004933 | 0.003835 |
| ged_best_ns_tlag8 | 0.006924 | 0.008834 | 0.004221 | 0.005180 | 0.004711 |
| ged_best_ns_tlag6 | 0.006636 | 0.009722 | 0.006591 | 0.005878 | 0.004709 |
| ged_best_ns_tlag9 | 0.006285 | 0.006875 | 0.004579 | 0.005833 | 0.004222 |
| ged_best_ns_tlag10 | 0.006240 | 0.006320 | 0.005766 | 0.004752 | 0.003658 |
| ged_tx_ns_25_tlag3 | 0.006088 | 0.004330 | 0.003657 | 0.003356 | 0.001959 |
| ged_tx_ns_5_tlag9 | 0.005191 | 0.002816 | 0.003012 | 0.002655 | 0.002380 |
| ged_tx_ns_5_tlag6 | 0.005089 | 0.004934 | 0.003989 | 0.003085 | 0.003110 |
| ged_tx_ns_5_tlag11 | 0.005062 | 0.002996 | 0.002286 | 0.002159 | 0.002731 |

### Change history

The model has been in use since r.2020.02.01

## B-3.13 REIGN (reign_glob)

### Description

This model makes use of features derived from the Rulers, Election, and Irregular Governance dataset, which can be found at `www.oefresearch.org` (Bell, 2016b). This dataset provides political, socioeconomic, and conflict data at the leader-month unit-of-analysis for all state members of the United Nations from January 1950 to the Present. The One Earth Future Foundation updates and distributes REIGN every month. Data relevant to this project include information on elections, leader traits, political regime tenures, and coups. Political regime has been repeatedly connected to violence by empirical studies (Vreeland, 2008). This model builds on the evidence of the literature that focuses on violence escalation (Davenport and Lichbach, 2005) and more broadly investigates the temporal context or transitional settings in which conflicts are more likely to erupt, such as political transition induced by coups (Belkin and Schofer, 2003) or in the proximity of elections (Burchard, 2015).

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on data for all countries in the world.

### Feature importances

The tables below report the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). Please note that for broad models like *reign*, the tables report only the first 25 most important features. Irregularity in elections is the most important feature in prediction violence of any type, followed by other elements ascribing to the elections (e.g. the loss of the incumbent). Personal characteristics of the political leader, including their age, also display a high relative importance in predicting violence, especially **sb** one.

*Table B-23.* Feature importances for the Reign_glob model, **sb**

| sb_reign_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| reign_irregular | 0.208641 | 0.208297 | 0.208288 | 0.207164 | 0.210730 |
| reign_age | 0.178890 | 0.181832 | 0.185165 | 0.186732 | 0.185910 |
| reign_loss | 0.153299 | 0.154663 | 0.155090 | 0.155401 | 0.156180 |
| reign_tenure_months | 0.139317 | 0.140344 | 0.139382 | 0.140608 | 0.140476 |
| reign_lastelection | 0.107261 | 0.107320 | 0.107190 | 0.107740 | 0.106991 |
| in_africa | 0.022224 | 0.022485 | 0.020909 | 0.020778 | 0.019729 |
| reign_elected | 0.022194 | 0.021695 | 0.021788 | 0.021604 | 0.021678 |
| reign_gov_foreign_occupied | 0.016611 | 0.016745 | 0.016833 | 0.018321 | 0.017901 |
| reign_militarycareer | 0.014008 | 0.013730 | 0.014417 | 0.014216 | 0.013157 |
| reign_gov_warlordism | 0.013945 | 0.012048 | 0.010800 | 0.008607 | 0.006743 |
| reign_gov_parliamentary_democracy | 0.012664 | 0.012834 | 0.012586 | 0.012440 | 0.012008 |
| reign_gov_military | 0.012261 | 0.012314 | 0.011770 | 0.011164 | 0.009875 |
| reign_male | 0.012138 | 0.012377 | 0.011660 | 0.010164 | 0.009663 |
| reign_gov_presidential_democracy | 0.009912 | 0.009994 | 0.009883 | 0.009994 | 0.009642 |
| reign_gov_personal_dictatorship | 0.009241 | 0.009084 | 0.009105 | 0.009008 | 0.009282 |
| reign_gov_dominant_party | 0.006209 | 0.006364 | 0.006425 | 0.007018 | 0.007487 |
| reign_anticipation | 0.005159 | 0.005312 | 0.005336 | 0.004943 | 0.005183 |
| reign_gov_military_personal | 0.004482 | 0.004178 | 0.004774 | 0.004480 | 0.005159 |
| reign_exec_ant | 0.004466 | 0.004092 | 0.004307 | 0.004136 | 0.004230 |
| reign_delayed | 0.004176 | 0.004109 | 0.003118 | 0.003231 | 0.003991 |
| reign_irreg_lead_ant | 0.004001 | 0.003248 | 0.003303 | 0.004164 | 0.003370 |
| reign_gov_provisional_civilian | 0.003656 | 0.003112 | 0.002750 | 0.002363 | 0.002125 |
| reign_gov_monarchy | 0.003021 | 0.002997 | 0.003360 | 0.002908 | 0.002552 |
| reign_gov_party_personal | 0.002647 | 0.002770 | 0.002901 | 0.003299 | 0.003632 |
| reign_direct_recent | 0.002342 | 0.002253 | 0.002413 | 0.002044 | 0.001990 |

*Table B-24.* Feature importances for the Reign_glob model, **os**

| os_reign_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| reign_irregular | 0.213691 | 0.211442 | 0.210742 | 0.209952 | 0.210172 |
| reign_loss | 0.180192 | 0.180848 | 0.179966 | 0.180735 | 0.180594 |
| reign_tenure_months | 0.160809 | 0.163599 | 0.163671 | 0.164865 | 0.164027 |
| reign_age | 0.156930 | 0.157695 | 0.161969 | 0.162319 | 0.158592 |
| reign_lastelection | 0.135470 | 0.133956 | 0.134785 | 0.132562 | 0.136518 |
| in_africa | 0.015470 | 0.015526 | 0.016116 | 0.015534 | 0.014978 |
| reign_militarycareer | 0.014685 | 0.013762 | 0.013352 | 0.013824 | 0.012963 |
| reign_elected | 0.013276 | 0.012928 | 0.013929 | 0.014047 | 0.015185 |
| reign_gov_personal_dictatorship | 0.009630 | 0.010197 | 0.010634 | 0.009605 | 0.009369 |
| reign_gov_presidential_democracy | 0.008094 | 0.008115 | 0.008777 | 0.009365 | 0.008810 |
| reign_delayed | 0.006321 | 0.005837 | 0.004704 | 0.004240 | 0.005092 |
| reign_gov_parliamentary_democracy | 0.006291 | 0.006615 | 0.005948 | 0.005819 | 0.005764 |
| reign_gov_military_personal | 0.006044 | 0.005834 | 0.005194 | 0.004311 | 0.004712 |
| reign_anticipation | 0.005972 | 0.005671 | 0.005551 | 0.005367 | 0.005139 |
| reign_gov_dominant_party | 0.004895 | 0.004914 | 0.005039 | 0.004821 | 0.005482 |
| reign_gov_foreign_occupied | 0.004889 | 0.005393 | 0.005906 | 0.007227 | 0.007796 |
| reign_exec_ant | 0.004728 | 0.004424 | 0.004129 | 0.004012 | 0.004297 |
| reign_male | 0.004314 | 0.003309 | 0.002968 | 0.002961 | 0.003490 |
| reign_gov_monarchy | 0.004249 | 0.004265 | 0.004282 | 0.002436 | 0.001798 |
| reign_gov_military | 0.003972 | 0.003873 | 0.004152 | 0.003932 | 0.004415 |
| reign_irreg_lead_ant | 0.003297 | 0.002765 | 0.003064 | 0.003841 | 0.003415 |
| reign_gov_warlordism | 0.003095 | 0.002658 | 0.002430 | 0.001934 | 0.002306 |
| reign_ref_ant | 0.002877 | 0.002113 | 0.002744 | 0.002695 | 0.002498 |
| reign_gov_provisional_civilian | 0.002610 | 0.002258 | 0.002197 | 0.001733 | 0.001752 |
| reign_pt_attempt | 0.002406 | 0.001798 | 0.001371 | 0.001214 | 0.001511 |

*Table B-25.* Feature importances for the Reign_glob model, **ns**

| ns_reign_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| reign_irregular | 0.192352 | 0.191660 | 0.190843 | 0.193407 | 0.193789 |
| reign_loss | 0.170381 | 0.173078 | 0.172364 | 0.179598 | 0.183808 |
| reign_tenure_months | 0.166850 | 0.168046 | 0.170760 | 0.174177 | 0.180197 |
| reign_age | 0.150410 | 0.149035 | 0.150902 | 0.143843 | 0.142593 |
| reign_lastelection | 0.137853 | 0.140840 | 0.140846 | 0.140274 | 0.141517 |
| reign_gov_oligarchy | 0.026595 | 0.022337 | 0.019282 | 0.010370 | 0.003594 |
| reign_militarycareer | 0.018147 | 0.017430 | 0.017488 | 0.016344 | 0.015527 |
| in_africa | 0.017522 | 0.018547 | 0.017913 | 0.017361 | 0.019392 |
| reign_gov_warlordism | 0.015352 | 0.014624 | 0.013374 | 0.012295 | 0.010706 |
| reign_elected | 0.015090 | 0.015316 | 0.015544 | 0.016829 | 0.016537 |
| reign_gov_personal_dictatorship | 0.010203 | 0.010772 | 0.010351 | 0.011471 | 0.011439 |
| reign_gov_presidential_democracy | 0.008592 | 0.008742 | 0.009698 | 0.010503 | 0.010276 |
| reign_delayed | 0.006506 | 0.005988 | 0.006633 | 0.005541 | 0.005881 |
| reign_anticipation | 0.005081 | 0.005176 | 0.005579 | 0.005631 | 0.005118 |
| reign_gov_parliamentary_democracy | 0.005073 | 0.005003 | 0.004823 | 0.005154 | 0.005216 |
| reign_exec_ant | 0.004861 | 0.005228 | 0.005324 | 0.004747 | 0.004887 |
| reign_male | 0.004761 | 0.004091 | 0.004103 | 0.003965 | 0.004503 |
| reign_gov_dominant_party | 0.004412 | 0.004632 | 0.004260 | 0.004555 | 0.004868 |
| reign_irreg_lead_ant | 0.003515 | 0.003720 | 0.003818 | 0.003643 | 0.004047 |
| reign_gov_provisional_civilian | 0.002666 | 0.002798 | 0.002126 | 0.002385 | 0.002188 |
| reign_ref_ant | 0.002367 | 0.001552 | 0.001631 | 0.002408 | 0.001813 |
| reign_gov_party_personal_military_hybrid | 0.002359 | 0.002600 | 0.002566 | 0.003015 | 0.002059 |
| reign_direct_recent | 0.002324 | 0.001983 | 0.001639 | 0.001839 | 0.001490 |
| reign_gov_military | 0.001925 | 0.002070 | 0.001906 | 0.002037 | 0.001632 |
| reign_gov_monarchy | 0.001872 | 0.001683 | 0.001612 | 0.001498 | 0.001174 |

### Change history

The model has been in use since r.2020.02.01

## B-3.14 V-DEM (vdem_glob)

### Description

A model based on all the mid-level indices from the Varieties of Democracy (V-Dem) dataset describing the political institutions of the country (Coppedge et al., 2019; Pemstein et al., 2019). V-Dem introduces a multidimensional index that reflects the complexity of the concept of democracy as a system of rule that goes beyond the simple presence of elections. Democracy exhibits a U-shape effect on conflict, whereby the two extremes of the democratic scale feature a higher probability of conflict (Hegre et al., 2001).

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on global data.

### Feature importances

The tables below report the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). Please note that in broad models like *Vdem*, the tables report only the first 25 most important features. Not surprisingly, the index of physical violence computed by V-Dem is the most important feature in predicting violence of any type. Freedom of movement and the regularity of elections are among the most important features in explaining **sb** violence. Freedom from forced labour is especially relevant in predicting one-sided violence, while the dynamics of the elections have considerable importance in predicting non-state conflicts.

*Table B-26.* Feature importances for the V-Dem model, **sb**

| sb_vdem_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| vdem_v2x_clphy | 0.010409 | 0.009486 | 0.007102 | 0.008419 | 0.007369 |
| vdem_v2xcl_dmove | 0.010281 | 0.010253 | 0.009846 | 0.009782 | 0.009394 |
| tlag_12_vdem_v2xcl_dmove | 0.008273 | 0.008927 | 0.008476 | 0.009493 | 0.007569 |
| tlag_12_vdem_v2xel_regelec | 0.008099 | 0.007464 | 0.009039 | 0.008306 | 0.007912 |
| tlag_60_tlag_12_vdem_v2xel_regelec | 0.007935 | 0.006767 | 0.007093 | 0.006836 | 0.006442 |
| tlag_12_vdem_v2x_clphy | 0.007516 | 0.008111 | 0.008251 | 0.007429 | 0.007080 |
| vdem_v2xcl_rol | 0.007450 | 0.007591 | 0.007145 | 0.007634 | 0.008181 |
| tlag_60_tlag_12_vdem_v2xcl_dmove | 0.007317 | 0.007400 | 0.008208 | 0.007663 | 0.006717 |
| vdem_v2xel_regelec | 0.007214 | 0.008346 | 0.008644 | 0.008624 | 0.009097 |
| tlag_60_vdem_v2xcl_dmove | 0.007162 | 0.007038 | 0.007272 | 0.007079 | 0.007332 |
| vdem_v2x_civlib | 0.007131 | 0.006562 | 0.005557 | 0.007136 | 0.005469 |
| vdem_v2x_gencl | 0.006957 | 0.005559 | 0.007438 | 0.007583 | 0.008400 |
| tlag_12_vdem_v2xcl_slave | 0.006934 | 0.006412 | 0.005884 | 0.005654 | 0.006469 |
| tlag_12_vdem_v2x_gencl | 0.006771 | 0.006221 | 0.006397 | 0.006232 | 0.006246 |
| tlag_60_vdem_v2xel_regelec | 0.006563 | 0.008206 | 0.006850 | 0.006282 | 0.006449 |
| vdem_v2x_clpriv | 0.006500 | 0.006413 | 0.006948 | 0.005309 | 0.006259 |
| tlag_12_vdem_v2xcl_rol | 0.006321 | 0.006399 | 0.006639 | 0.005534 | 0.006375 |
| vdem_v2xcl_slave | 0.006106 | 0.005758 | 0.007753 | 0.007373 | 0.005643 |
| tlag_60_tlag_12_vdem_v2xme_altinf | 0.005680 | 0.005071 | 0.005590 | 0.005543 | 0.004977 |
| tlag_60_vdem_v2xcl_slave | 0.005548 | 0.004286 | 0.004658 | 0.003572 | 0.004496 |
| tlag_60_tlag_12_vdem_v2xcl_rol | 0.005495 | 0.004377 | 0.004586 | 0.004583 | 0.004747 |
| tlag_60_vdem_v2xdl_delib | 0.005234 | 0.003879 | 0.004773 | 0.004912 | 0.004335 |
| tlag_60_vdem_v2x_clphy | 0.005183 | 0.005299 | 0.005632 | 0.004905 | 0.005429 |
| tlag_60_vdem_v2xme_altinf | 0.005161 | 0.004861 | 0.005709 | 0.004103 | 0.005475 |
| vdem_v2xme_altinf | 0.005094 | 0.005221 | 0.004608 | 0.005784 | 0.004885 |

*Table B-27.* Feature importances for the V-Dem model, **os**

| os_vdem_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| vdem_v2xcl_dmove | 0.007946 | 0.009199 | 0.009129 | 0.008349 | 0.007315 |
| vdem_v2xcl_slave | 0.007537 | 0.008226 | 0.006554 | 0.005790 | 0.005848 |
| tlag_12_vdem_v2xcl_slave | 0.007094 | 0.007479 | 0.005385 | 0.005959 | 0.005029 |
| vdem_v2xcl_prpty | 0.007071 | 0.006698 | 0.005962 | 0.006549 | 0.006184 |
| tlag_12_vdem_v2x_clphy | 0.007042 | 0.005834 | 0.007693 | 0.006508 | 0.006366 |
| vdem_v2x_clphy | 0.007008 | 0.007179 | 0.008405 | 0.007383 | 0.006875 |
| tlag_12_vdem_v2xme_altinf | 0.006669 | 0.005538 | 0.005768 | 0.005342 | 0.006030 |
| vdem_v2xcl_rol | 0.006528 | 0.005875 | 0.006549 | 0.005442 | 0.006197 |
| tlag_12_vdem_v2xcl_prpty | 0.006373 | 0.005926 | 0.005797 | 0.005629 | 0.005674 |
| tlag_12_vdem_v2xcl_dmove | 0.006347 | 0.007499 | 0.007835 | 0.007607 | 0.006646 |
| tlag_12_vdem_v2xel_regelec | 0.006320 | 0.004557 | 0.006972 | 0.006040 | 0.004481 |
| vdem_v2xme_altinf | 0.006012 | 0.007229 | 0.005850 | 0.005309 | 0.005511 |
| vdem_v2xel_regelec | 0.005968 | 0.007952 | 0.005737 | 0.006263 | 0.004896 |
| vdem_v2xeg_eqdr | 0.005776 | 0.004770 | 0.005833 | 0.005946 | 0.004073 |
| tlag_12_vdem_v2x_gencl | 0.005745 | 0.005400 | 0.005441 | 0.005239 | 0.005452 |
| tlag_12_vdem_v2x_clpriv | 0.005710 | 0.004855 | 0.005461 | 0.005100 | 0.004281 |
| vdem_v2x_freexp_altinf | 0.005689 | 0.004930 | 0.005021 | 0.005501 | 0.004628 |
| tlag_60_vdem_v2xeg_eqdr | 0.005643 | 0.005250 | 0.005840 | 0.004507 | 0.004542 |
| vdem_v2x_clpriv | 0.005611 | 0.005563 | 0.005622 | 0.005243 | 0.005352 |
| tlag_12_vdem_v2x_freexp_altinf | 0.005547 | 0.004733 | 0.004739 | 0.004399 | 0.004352 |
| tlag_12_vdem_v2xcl_rol | 0.005517 | 0.004994 | 0.005015 | 0.005245 | 0.005244 |
| vdem_v2x_frassoc_thick | 0.005457 | 0.005394 | 0.004757 | 0.005177 | 0.003563 |
| tlag_60_vdem_v2xme_altinf | 0.005397 | 0.004755 | 0.004698 | 0.004960 | 0.005055 |
| tlag_60_vdem_v2xcl_dmove | 0.005357 | 0.004754 | 0.006313 | 0.005020 | 0.005513 |
| vdem_v2x_civlib | 0.005339 | 0.005419 | 0.005104 | 0.005090 | 0.004677 |

*Table B-28.* Feature importances for the V-Dem model, **ns**

| ns_vdem_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| vdem_v2xel_regelec | 0.008199 | 0.007379 | 0.006131 | 0.007321 | 0.005601 |
| tlag_12_vdem_v2xel_regelec | 0.008081 | 0.005684 | 0.007889 | 0.005665 | 0.005667 |
| vdem_v2xpe_exlecon | 0.006939 | 0.006117 | 0.005958 | 0.004858 | 0.004896 |
| tlag_60_vdem_v2xel_regelec | 0.006772 | 0.007083 | 0.007874 | 0.006935 | 0.007811 |
| tlag_60_tlag_12_vdem_v2xel_regelec | 0.006772 | 0.007277 | 0.008235 | 0.007136 | 0.009181 |
| tlag_12_vdem_v2x_clphy | 0.006580 | 0.006579 | 0.006126 | 0.004922 | 0.005386 |
| vdem_v2x_clphy | 0.006100 | 0.006705 | 0.006863 | 0.005925 | 0.007264 |
| vdem_v2x_civlib | 0.005995 | 0.005617 | 0.004856 | 0.005398 | 0.004758 |
| tlag_60_vdem_v2xme_altinf | 0.005855 | 0.004374 | 0.005003 | 0.004955 | 0.003989 |
| vdem_v2xeg_eqdr | 0.005781 | 0.004741 | 0.004764 | 0.004893 | 0.004720 |
| tlag_60_tlag_12_vdem_v2xeg_eqdr | 0.005472 | 0.004841 | 0.004208 | 0.003711 | 0.004605 |
| tlag_12_vdem_v2xpe_exlecon | 0.005434 | 0.005493 | 0.004766 | 0.004780 | 0.005290 |
| vdem_v2x_partip | 0.005303 | 0.004247 | 0.004012 | 0.003771 | 0.004045 |
| tlag_60_vdem_v2x_clphy | 0.005264 | 0.005502 | 0.006210 | 0.005184 | 0.005097 |
| tlag_60_vdem_v2xeg_eqdr | 0.005193 | 0.004339 | 0.004367 | 0.004695 | 0.004037 |
| vdem_v2xme_altinf | 0.005174 | 0.004109 | 0.004753 | 0.005062 | 0.004162 |
| vdem_v2xcl_acjst | 0.005166 | 0.004106 | 0.003656 | 0.003898 | 0.003257 |
| vdem_v2xcl_rol | 0.005162 | 0.005002 | 0.005092 | 0.004339 | 0.004600 |
| tlag_12_vdem_v2xdl_delib | 0.005151 | 0.003843 | 0.004141 | 0.003976 | 0.005028 |
| tlag_12_vdem_v2xcl_dmove | 0.005120 | 0.003617 | 0.004136 | 0.004772 | 0.004168 |
| vdem_v2x_clpriv | 0.005084 | 0.004817 | 0.004189 | 0.003827 | 0.004349 |
| tlag_60_tlag_12_vdem_v2x_clphy | 0.005025 | 0.005345 | 0.005596 | 0.005558 | 0.004463 |
| tlag_60_tlag_12_vdem_v2xpe_exlecon | 0.004975 | 0.006449 | 0.005714 | 0.004536 | 0.005099 |
| tlag_12_vdem_v2xme_altinf | 0.004967 | 0.004734 | 0.005384 | 0.006049 | 0.005579 |
| tlag_12_vdem_v2x_horacc_osp | 0.004920 | 0.004611 | 0.004203 | 0.003302 | 0.004165 |

### Change history

The model has been in use since r.2020.02.01

## B-3.15 WDI global (wdi_all_glob)

### Description

This model makes use of uses a whole set of World Development Indicators (World Bank, 1999) at the global level. The set of indicators broadly captures the level of development by country, including the quality of infrastructure, economic growth and the amount of national debt, education and gender equality, health care and provision, agricultural dependence and migration flows. Development is considered one of the best predictors of violence, to the point that scholars have defined conflict as 'development in reverse' (Collier et al., 2003).

### Estimation details

The model uses the 'one-step-ahead' approach (see Section B-2) and is estimated on global data.

### Feature importances

The tables below report the feature importances for each type of violence and different time steps ($s = 1, 3, 6, 12, 24, 36$). Please note that in broad models like *wdi*, the tables report only the first 25 most important features. At the global level, the number of refugees by country is the most important feature in predicting **sb** violence of any type. Demographic patterns and the amount of arable land are also considerably important in predicting violence, as suggested by empirical research on the causes of conflict (Buhaug and Gates, 2002).

*Table B-29.* Feature importances for the WDI model, **sb**

| sb_wdi_all_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| wdi_sm_pop_refg_or | 0.043940 | 0.047610 | 0.047649 | 0.039179 | 0.034368 |
| wdi_sp_pop_totl | 0.020928 | 0.019520 | 0.022193 | 0.021442 | 0.018922 |
| wdi_ag_lnd_totl_k2 | 0.017526 | 0.014938 | 0.016743 | 0.016395 | 0.018735 |
| wdi_ag_srf_totl_k2 | 0.015768 | 0.017599 | 0.015072 | 0.016832 | 0.014001 |
| wdi_sl_uem_advn_zs | 0.015278 | 0.013320 | 0.012481 | 0.011274 | 0.013791 |
| wdi_sm_pop_netm | 0.013522 | 0.014170 | 0.014317 | 0.011240 | 0.008703 |
| wdi_sm_pop_totl_zs | 0.013509 | 0.012435 | 0.013033 | 0.013958 | 0.014023 |
| wdi_ag_lnd_totl_ru_k2 | 0.012414 | 0.013469 | 0.014979 | 0.014695 | 0.012432 |
| wdi_sh_sta_maln_zs | 0.011259 | 0.010171 | 0.008640 | 0.010398 | 0.009715 |
| wdi_sl_tlf_totl_fe_zs | 0.011163 | 0.010590 | 0.012649 | 0.013276 | 0.012689 |
| wdi_nv_agr_totl_kd | 0.010118 | 0.009683 | 0.011207 | 0.010823 | 0.012881 |
| wdi_vc_idp_nwds | 0.009025 | 0.009478 | 0.008301 | 0.007395 | 0.006508 |
| wdi_vc_idp_nwcv | 0.008584 | 0.010641 | 0.007828 | 0.007692 | 0.007474 |
| wdi_sl_uem_neet_zs | 0.008477 | 0.007368 | 0.010582 | 0.008036 | 0.008334 |
| wdi_nv_agr_totl_cn | 0.008462 | 0.009254 | 0.007609 | 0.008798 | 0.009502 |
| wdi_sl_uem_neet_fe_zs | 0.008394 | 0.007675 | 0.007403 | 0.006019 | 0.006976 |
| wdi_nv_agr_totl_kn | 0.008017 | 0.006971 | 0.009070 | 0.011350 | 0.008457 |
| wdi_sl_uem_advn_fe_zs | 0.007937 | 0.007716 | 0.006736 | 0.007289 | 0.009086 |
| wdi_sl_uem_advn_ma_zs | 0.007911 | 0.006661 | 0.006998 | 0.007155 | 0.007189 |
| wdi_se_sec_nenr | 0.007479 | 0.005658 | 0.006467 | 0.005604 | 0.005090 |
| wdi_ag_lnd_frst_k2 | 0.006907 | 0.007203 | 0.006924 | 0.006360 | 0.006877 |
| wdi_se_enr_prim_fm_zs | 0.006430 | 0.006990 | 0.005490 | 0.006366 | 0.007612 |
| wdi_ny_gnp_mktp_pp_kd | 0.006403 | 0.004089 | 0.004999 | 0.004884 | 0.006438 |
| wdi_ny_gdp_mktp_kd | 0.006210 | 0.005229 | 0.005562 | 0.005603 | 0.005320 |
| wdi_se_enr_prsc_fm_zs | 0.006173 | 0.006555 | 0.006089 | 0.008686 | 0.010037 |

### Change history

The model has been in use since r.2020.02.01

*Table B-30.* Feature importances for the WDI model, **os**

| os_wdi_all_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| wdi_sm_pop_refg_or | 0.033814 | 0.029615 | 0.029063 | 0.027335 | 0.025151 |
| wdi_ag_lnd_totl_k2 | 0.013870 | 0.012723 | 0.013372 | 0.013248 | 0.015456 |
| wdi_sp_pop_totl | 0.012963 | 0.012967 | 0.013872 | 0.013465 | 0.012947 |
| wdi_ag_srf_totl_k2 | 0.011859 | 0.010836 | 0.014689 | 0.013504 | 0.014005 |
| wdi_ag_lnd_totl_ru_k2 | 0.011839 | 0.012915 | 0.014049 | 0.013313 | 0.013638 |
| wdi_nv_agr_totl_kn | 0.010434 | 0.009203 | 0.010695 | 0.012349 | 0.011364 |
| wdi_sm_pop_netm | 0.010406 | 0.011838 | 0.011434 | 0.011563 | 0.012668 |
| wdi_ag_lnd_frst_k2 | 0.009226 | 0.009949 | 0.009091 | 0.010017 | 0.010734 |
| wdi_ny_adj_dfor_cd | 0.008869 | 0.010070 | 0.009552 | 0.009487 | 0.008447 |
| wdi_sh_sta_maln_zs | 0.008781 | 0.009001 | 0.010058 | 0.009342 | 0.009811 |
| wdi_sh_h2o_basw_zs | 0.008450 | 0.008989 | 0.008487 | 0.007116 | 0.006873 |
| wdi_sp_dyn_imrt_fe_in | 0.008432 | 0.008173 | 0.007899 | 0.007888 | 0.008790 |
| wdi_ny_gdp_mktp_kn | 0.008257 | 0.009144 | 0.007993 | 0.007838 | 0.007767 |
| wdi_sh_sta_stnt_zs | 0.007836 | 0.005529 | 0.005291 | 0.005883 | 0.004955 |
| wdi_se_enr_prim_fm_zs | 0.007792 | 0.009065 | 0.009038 | 0.008933 | 0.009343 |
| wdi_se_enr_prsc_fm_zs | 0.007484 | 0.009033 | 0.006463 | 0.008496 | 0.008699 |
| wdi_ny_gdp_pcap_pp_kd | 0.007429 | 0.005627 | 0.005833 | 0.005636 | 0.005699 |
| wdi_nv_agr_totl_cn | 0.006803 | 0.006444 | 0.005904 | 0.005984 | 0.005245 |
| wdi_sp_pop_grow | 0.006739 | 0.006317 | 0.007145 | 0.007213 | 0.006895 |
| wdi_ny_gdp_mktp_cd | 0.006446 | 0.005384 | 0.006325 | 0.006507 | 0.007211 |
| wdi_nv_agr_totl_kd | 0.006357 | 0.007352 | 0.008007 | 0.007169 | 0.007694 |
| wdi_ny_gdp_mktp_kd | 0.006270 | 0.006399 | 0.005315 | 0.004691 | 0.004671 |
| wdi_sp_dyn_imrt_in | 0.006236 | 0.006320 | 0.006559 | 0.005670 | 0.007153 |
| wdi_sh_dyn_mort_fe | 0.006218 | 0.006933 | 0.006770 | 0.008312 | 0.007804 |
| wdi_ny_gdp_mktp_cn_ad | 0.006074 | 0.008801 | 0.006157 | 0.005831 | 0.004953 |

*Table B-31.* Feature importances for the WDI model, **ns**

| ns_wdi_all_glob | 1 | 6 | 12 | 24 | 36 |
|---|---|---|---|---|---|
| wdi_sm_pop_refg_or | 0.020467 | 0.020840 | 0.018242 | 0.017822 | 0.015669 |
| wdi_ny_adj_dfor_cd | 0.017834 | 0.019719 | 0.016775 | 0.017964 | 0.015996 |
| wdi_ag_lnd_totl_k2 | 0.016836 | 0.016825 | 0.014785 | 0.018350 | 0.014397 |
| wdi_sp_pop_totl | 0.015363 | 0.014837 | 0.013579 | 0.013186 | 0.016007 |
| wdi_ag_srf_totl_k2 | 0.014801 | 0.014135 | 0.014818 | 0.014442 | 0.012119 |
| wdi_ms_mil_xpnd_gd_zs | 0.012741 | 0.011757 | 0.012460 | 0.014282 | 0.012788 |
| wdi_ag_lnd_totl_ru_k2 | 0.011893 | 0.011930 | 0.012680 | 0.009569 | 0.012158 |
| wdi_dt_ixa_offt_cd | 0.011651 | 0.013646 | 0.013537 | 0.012616 | 0.013852 |
| wdi_vc_idp_tocv | 0.009432 | 0.008404 | 0.008051 | 0.006054 | 0.005757 |
| wdi_sh_sta_wash_p5 | 0.008316 | 0.006579 | 0.006031 | 0.006388 | 0.007928 |
| wdi_bg_gsr_nfsv_gd_zs | 0.008280 | 0.007608 | 0.009337 | 0.007375 | 0.004096 |
| wdi_ny_gdp_fcst_kd | 0.008251 | 0.007068 | 0.008304 | 0.006584 | 0.007511 |
| wdi_ny_gdp_mktp_kd | 0.008130 | 0.008188 | 0.008263 | 0.010983 | 0.007799 |
| wdi_nv_agr_totl_kd | 0.007332 | 0.007293 | 0.007783 | 0.007754 | 0.008291 |
| wdi_sh_h2o_basw_zs | 0.007181 | 0.006288 | 0.006983 | 0.006679 | 0.007669 |
| wdi_ag_lnd_agri_zs | 0.006814 | 0.007128 | 0.005942 | 0.007779 | 0.009439 |
| wdi_dt_ixa_dppg_cd | 0.006747 | 0.006095 | 0.008274 | 0.007779 | 0.007379 |
| wdi_ms_mil_xpnd_zs | 0.006461 | 0.006554 | 0.005376 | 0.006854 | 0.008142 |
| wdi_sp_urb_totl_in_zs | 0.006337 | 0.005902 | 0.005699 | 0.005418 | 0.005817 |
| wdi_ny_gdp_mktp_pp_kd | 0.006220 | 0.007069 | 0.006913 | 0.005994 | 0.005826 |
| wdi_ny_gdp_mktp_pp_cd | 0.006206 | 0.004472 | 0.007359 | 0.005351 | 0.006086 |
| wdi_tx_val_agri_zs_un | 0.006148 | 0.004653 | 0.005564 | 0.003679 | 0.004076 |
| wdi_se_prm_nenr | 0.006115 | 0.004396 | 0.005680 | 0.003712 | 0.004787 |
| wdi_ny_gdp_defl_zs_ad | 0.006030 | 0.007250 | 0.006502 | 0.005750 | 0.006189 |
| wdi_ic_bus_ease_xq | 0.005906 | 0.007127 | 0.007020 | 0.008923 | 0.007924 |

## B-4 CHANGES IN *CM* MODELS

The new *cm* level ensemble include 16 models. All the models are trained through a random forest algorithm, with the exception of two 'dynamic simulation' models (see Hegre et al., 2019). In the following, we discuss the changes to the constituent models. All the other models presented in Table B-1 are identical to the ones described in Hegre et al. (2019).

We have pursued two avenues to improve performance for new conflicts. The first is to strengthen the model of underlying latent risk of violence through 'structural' factors. The *vdem_glob* model includes an extensive set of variables describing countries' political institutions from the V-Dem dataset (Coppedge et al., 2017). This 'institutional' model is grounded on decades of empirical literature showing that the effect of democracy on the likelihood of conflict follows a U-shaped curve (Benoit, 1996; Acemoglu and Robinson, 2001; Hegre et al., 2001; Cederman, Wimmer, and Min, 2010). The new *wdi_all_glob* model contains a wide range of socio-economic indicators from the World Development Indicators (World Bank, 2019). These range from poverty and health measures through various inequality indicators to assessments of countries' ability to implement good policies. These two 'structural' models replace the former *inst* and *econ* models, both of which performed poorly.

The second avenue is to add models that monitor variables that could function as early signals of increasing tensions. The new models *acled_violence, icgcw, reign_glob, reign_coups*, are particularly useful, supplementing the existing *acled_protest* and conflict history models. The data in these models are updated monthly. The *acled_protest* and *acled_violence* models include the recent history of protest and violence from ACLED (Raleigh et al., 2010). The *icgcw* model makes use of the warnings issued monthly by the International Crisis Group's Crisis Watch (`https://www.crisisgroup.org/crisiswatch`). The *reign_glob* model incorporates information on recent elections, coups, and other leader changes sourced from the REIGN dataset (Bell, 2016b, `www.oefresearch.org`). The *reign_coups* model contains the predicted risk of military coups from this project. These models build on evidence from the literature on escalation and dynamics of political violence, broadly concerned with the context in which conflicts are more likely to erupt, such as in the context of political transitions induced by coups (Belkin and Schofer, 2003) or in the proximity of elections (Birch, Daxecker, and Höglund, 2020). The *reign_drought* model also taps into early signals of tensions by including predictions for drought/precipitation data. The model reflects the recent empirical findings connecting droughts to increased likelihood of conflict (von Uexkull, 2014; von Uexkull et al., 2016).

We have also amended our conflict history models in order to reduce the system's tendency to overestimate risk in countries with recent conflict. The new conflict history model (*cflong*) is more extensive than the previous one. In particular, it now contains detailed information of the severity of past violence in terms of number of people killed and how much time has elapsed since earlier violence. We also have included a *neibhist* model that extensively describes the conflict history of both a country and its neighboring countries, building from the evidence that violence tends to spatially 'cluster' (Buhaug and Gleditsch, 2008).

As before, the models include two broad 'dynamic simulation' models, one trained on the incidence of conflict with at least one BRD as the outcome variable, the other using the incidence of at least 25 BRDs. A new attribute is that we now simulate these models together, so that the forecasted histories of each of these models inform each other. In addition, we simulate a model with incidence of state-based (**sb**), one-sided (**os**) violence, and non-state (**ns**) conflict together causing 500 deaths in a month as the outcome, so that we can also include more serious events as part of the conflict history.

The remaining models presented in Table B-1 are identical to the ones described in Hegre et al. (2019).

In the next sections, we delve deeper into the different types of violence that are predicted and provide the prediction maps for each outcome.

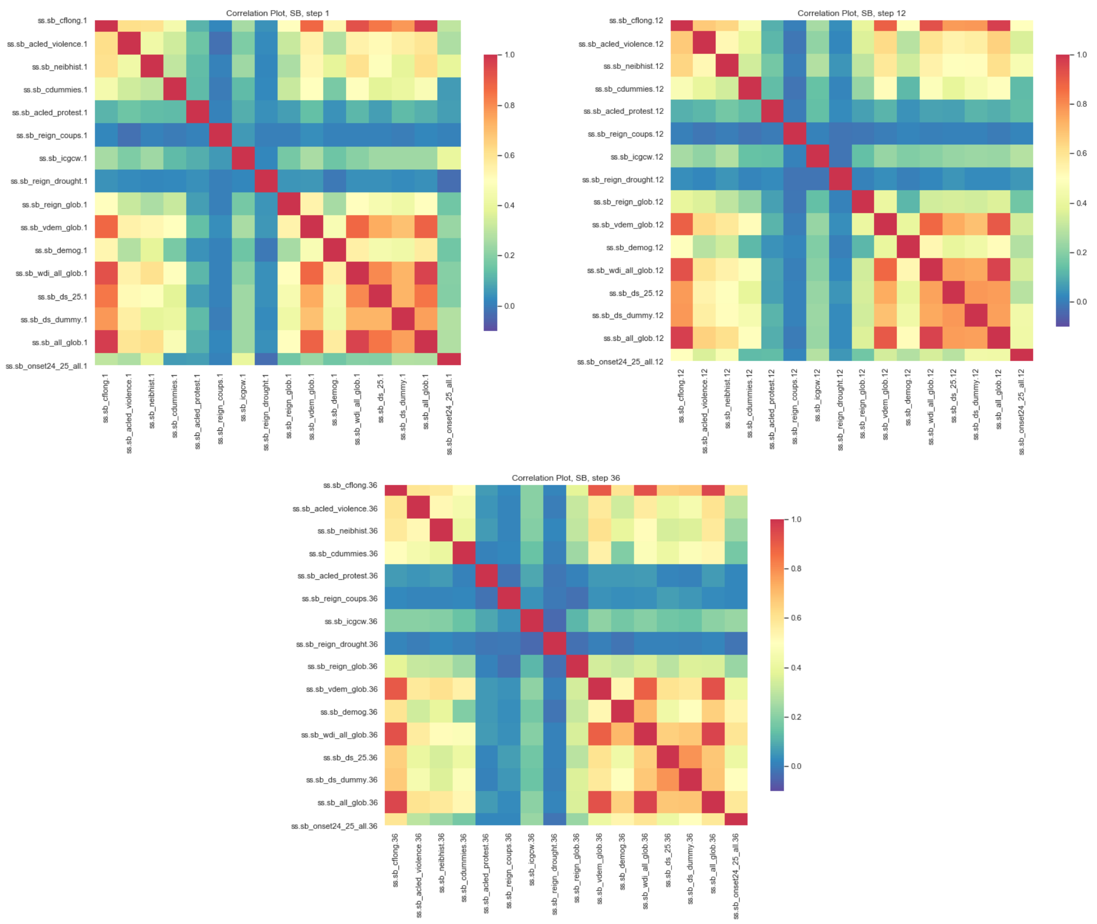
## B-5 DETAILED DESCRIPTIONS, CONSTITUENT MODELS

The following sections delve deeper into the three different types of violence (**sb**, **os**, **ns**) and present the results of predictions thereof.

### B-5.1 State-based conflict (sb)

State-based violence is defined according to the standard UCPD definition, as 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 (Pettersson and Eck, 2018). State-based conflicts are thus characterized by the use of the force by the government as an active side of conflict, as opposed to non-state conflicts where none of the active sides is a government. Predictions of these type of violence may be of interest for all scholars of civil wars and those concerned with the main drivers of conflict against the state. By contrast, the following sections may be more relevant for researchers of other forms of violence which do not involve the government as an active part.

*Figure B-1.* Correlation between predictions from constituent models, *cm*, **sb**




Not surprisingly, global models tend to be highly correlated among each other. The V-Dem global model especially shows correlation with a number of other models, including the other global ones, but also the *cflong* and the *Dynasim* models. The high correlation between models based on political regime (V-Dem), development (WDI) and conflict history (*cflong*) reflects the ability of these models to capture structural, slow-moving features and patterns that characterize countries relatively long-term. Models including more rapidly moving features, akin political turbulence (*reign_coups*, *acled_protest* and *acled_Violence*) or weather-related shocks (*reign_drought*) show less correlation with the other models. These trends in correlation increase over time, as shown by the correlation plots for steps 12 and 36, displaying a even higher correlation between slow-moving structural models.

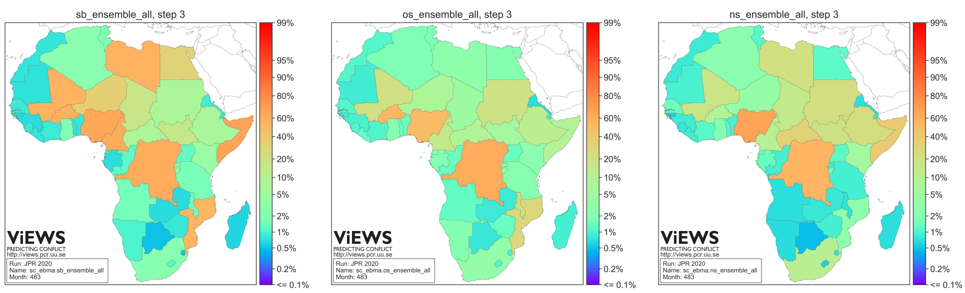
Presenting the predictions for different types of violence may help observe the relative influence of various sets of features in driving state-based, one-sided and non-state violence. Although the results of predictions cannot be considered as an orthodox causality test, they can still give useful indications to understand how a model is able to predict a specific type of violence and less so another one. Also, it may be interesting to observe how different models correlate in predicting each type of outcome. To have a closer inspection of how the predictors of **sb** conflict correlate, Figure B-1 shows the correlations between the predictions from each of the constituent models at three different steps ($s = 1, 12, 36$).

As follows, we present the prediction maps for **sb**.

### Prediction maps, forecasts (sb)

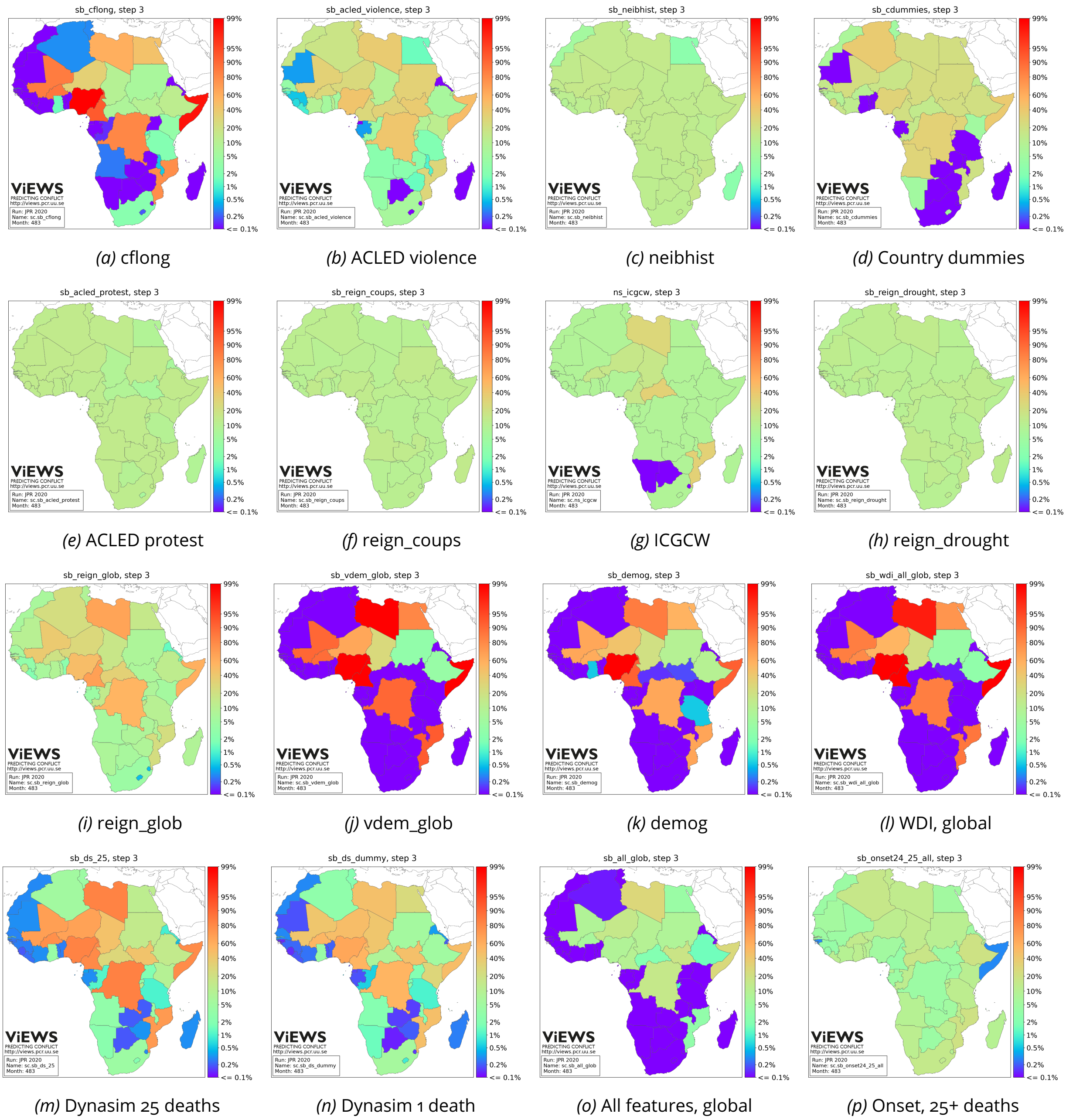
Here we present the forecasts for each of the constituent models in the ensembles for steps $s = (1, 3, 12, 36)$ at the *cm* level. Figure B-3 shows the forecasts for March 2020 based on the $s = 3$ model. Figure B-2 shows the forecast for the ensemble for the same month. All models are calibrated, i.e. re-scaled using a logit procedure so that mean predicted probability in the calibration period is similar to the observed proportion of conflict. Models that separate well between conflict and non-conflict have multi-colored maps. Weak models have green colors, since predicted probabilities deviate little from the mean probability of 0.11. Figure B-4 show the same for December 2022, i.e. $s = 36$ months into the future based on data up to and including December 2019.

*Figure B-2.* Predicted probabilities, ensemble models, **sb**, **os** and **ns**, $s = 3$ (month 483, March 2020), based on data up to December 2019




*(a)* Ensemble, **sb** *(b)* Ensemble, **os** *(c)* Ensemble, **ns**

*Figure B-3.* Predicted probabilities, constituent models, **sb**, $s = 3$ (month 483, March 2020), based on data up to December 2019



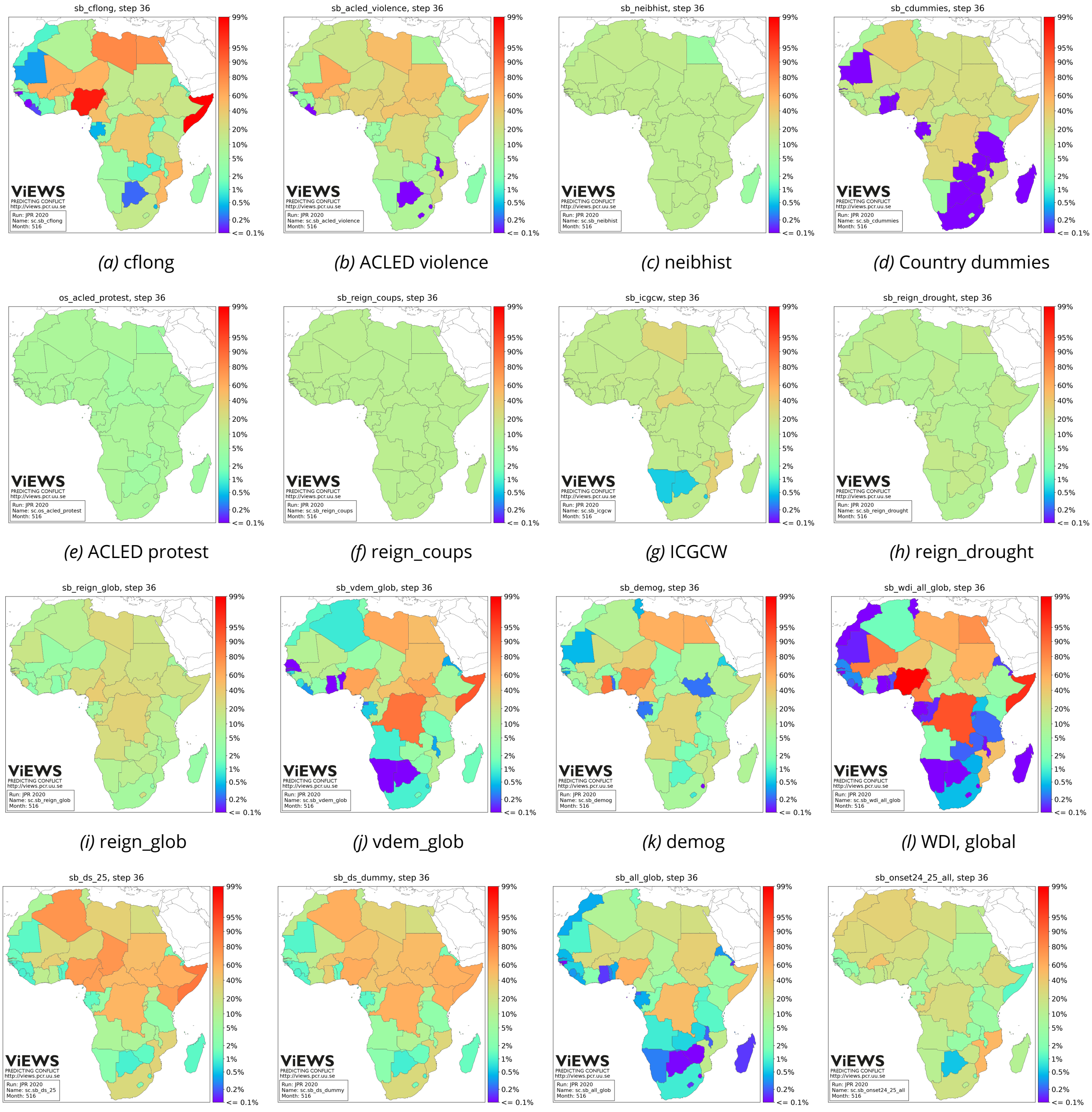

*(a)* cflong *(b)* ACLED violence *(c)* neibhist *(d)* Country dummies

*(e)* ACLED protest *(f)* reign_coups *(g)* ICGCW *(h)* reign_drought

*(i)* reign_glob *(j)* vdem_glob *(k)* demog *(l)* WDI, global

*(m)* Dynasim 25 deaths *(n)* Dynasim 1 death *(o)* All features, global *(p)* Onset, 25+ deaths

*Figure B-4.* Predicted probabilities, constituent models, **sb**, $s = 36$ (month 516, December 2022), based on data up to December 2019




*(a)* cflong *(b)* ACLED violence *(c)* neibhist *(d)* Country dummies

*(e)* ACLED protest *(f)* reign_coups *(g)* ICGCW *(h)* reign_drought

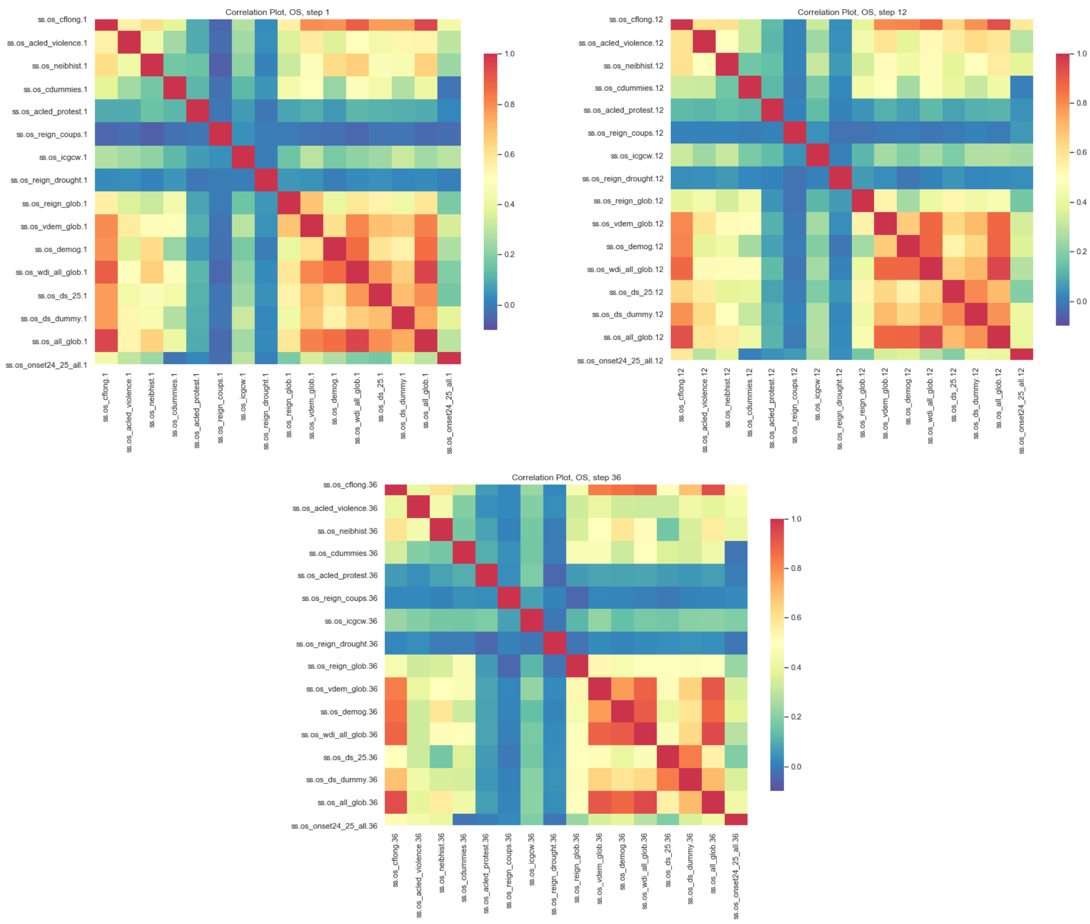
*(i)* reign_glob *(j)* vdem_glob *(k)* demog *(l)* WDI, global

*(m)* Dynasim 25 deaths *(n)* Dynasim 1 death *(o)* All features, global *(p)* Onset, 25+ deaths

### B-5.2 One-sided violence (os)

One-sided violence is defined as the deliberate use of armed force by the government of a state or by a formally organized group against civilians which result in at least 25 BRDs in a year (Pettersson and Eck, 2018). Unlike state-based conflict, *os* violence involves the use of the armed force by one actor against the civilian population. Predictions of **os** can be of particular interest to scholars of atrocities and violence against civilians.

*Figure B-5.* Correlation between predictions from constituent models, *cm*, **os**



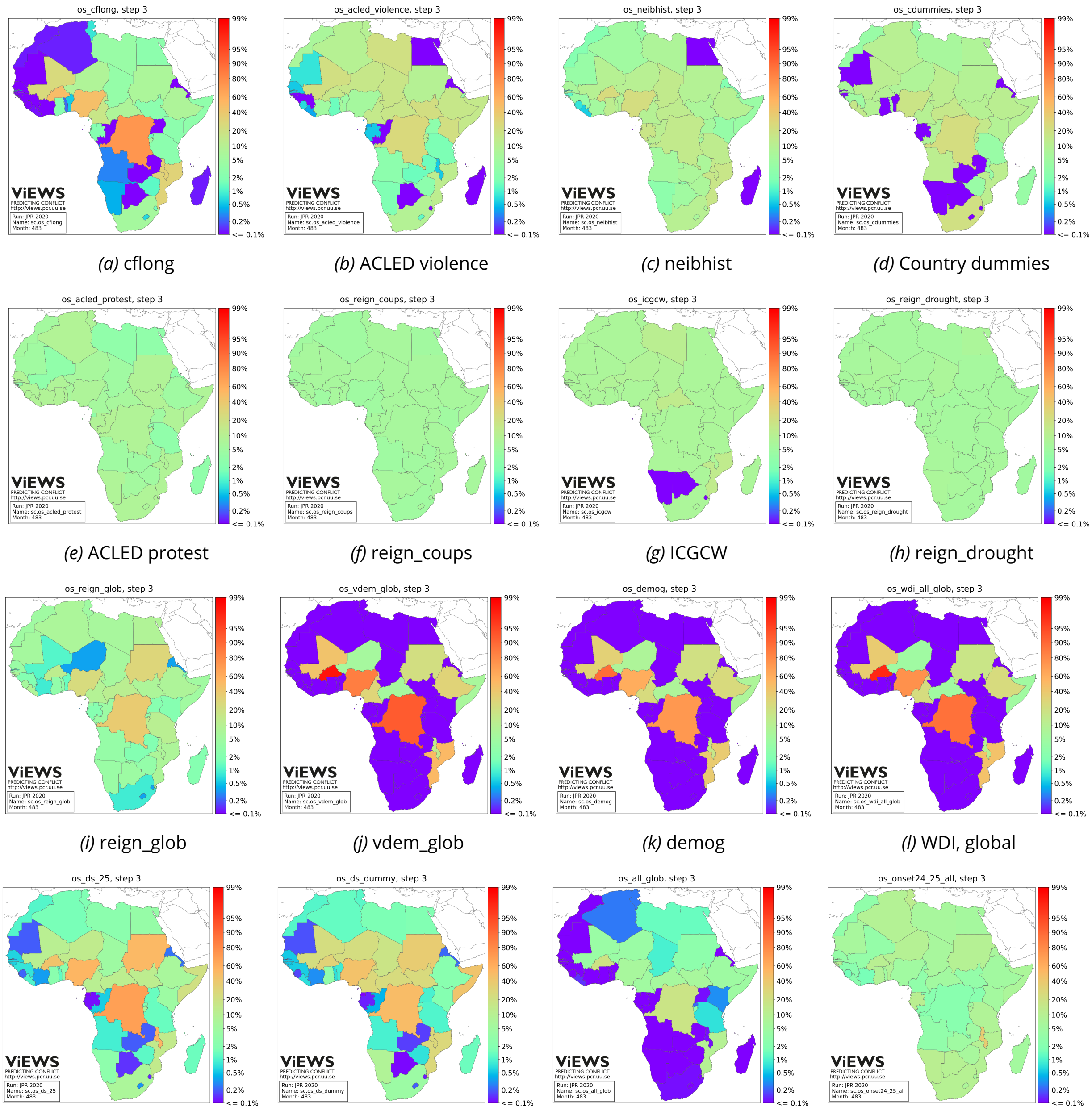

As we observed for **sb** violence, also for **os** global models tend to be highly correlated among each other. The *V-Dem* global model especially shows correlation with a number of other models, including the other global ones, but also the *cflong* and the *Dynasim* models. The high correlation between models based on political regime (V-Dem), development (WDI) and conflict history (*cflong*) reflects the ability of these models to capture structural, slow-moving features and patterns that characterize countries relatively long-term. Models including more rapidly moving features, akin political turbulence (*reign_coups*, *acled_protest* and *acled_violence*) or weather-related shocks (*reign_drought*) show less correlation with the other models. These trends in correlation increase over time, as shown by the correlation plots for steps 12 and 36, displaying a even higher correlation between slow-moving structural models.

As follows, we present the prediction maps for **os**.

#### Prediction maps, forecasts (os)

Here we present the forecasts for each of the constituent models in the ensembles for steps $s = (1, 3, 12, 36)$ at the *cm* level. Figure B-6 shows the forecasts for March 2020 based on the $s = 3$ model. Figure B-2 shows the forecast for the ensemble for the same month. All models are calibrated, i.e. re-scaled using a logit procedure so that mean predicted probability in the calibration period is similar to the observed proportion of conflict. Models that separate well between conflict and non-conflict have multi-colored maps. Weak models have green colors, since predicted probabilities deviate little from the mean probability of 0.11. Figure B-7 show the same for December 2022, i.e. $s = 36$ months into the future based on data up to and including December 2019.

*Figure B-6.* Predicted probabilities, constituent models, **os**, $s = 3$ (month 483, March 2020), based on data up to December 2019



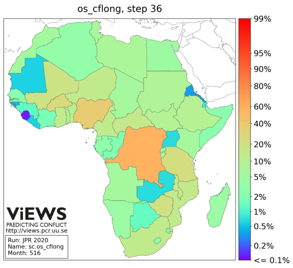

*(a)* cflong *(b)* ACLED violence *(c)* neibhist *(d)* Country dummies

*(e)* ACLED protest *(f)* reign_coups *(g)* ICGCW *(h)* reign_drought

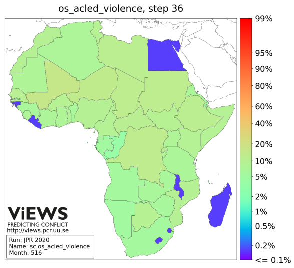
*(i)* reign_glob *(j)* vdem_glob *(k)* demog *(l)* WDI, global

*(m)* Dynasim 25 deaths *(n)* Dynasim 1 death *(o)* All features, global *(p)* Onset, 25+ deaths

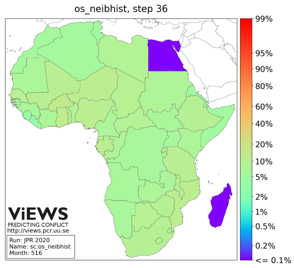
*Figure B-7.* Predicted probabilities, constituent models, **os**, $s = 36$ (month 516, December 2022), based on data up to December 2019



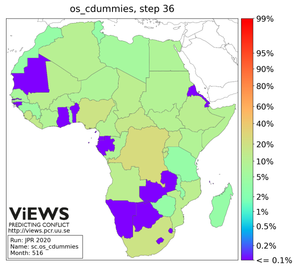

*(a)* cflong



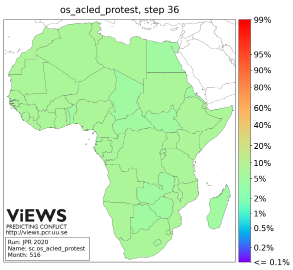

*(b)* ACLED violence



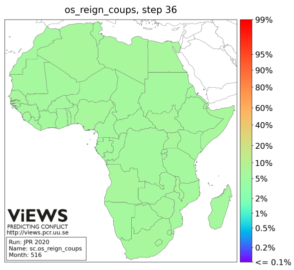

*(c)* neibhist



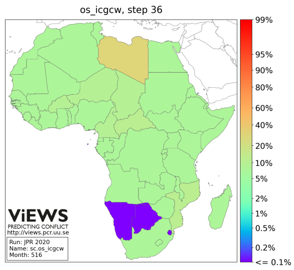

*(d)* Country dummies



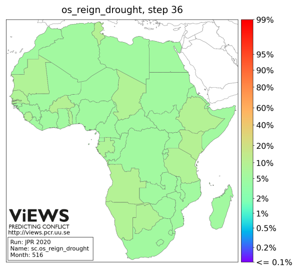

*(e)* ACLED protest



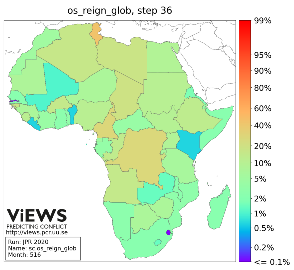

*(f)* reign_coups



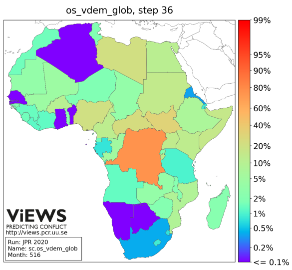

*(g)* ICGCW



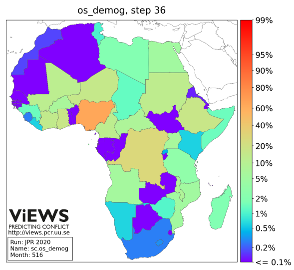

*(h)* reign_drought



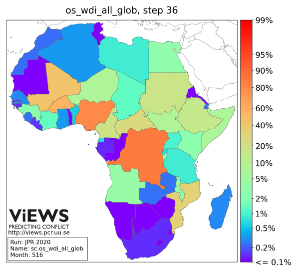

*(i)* reign_glob



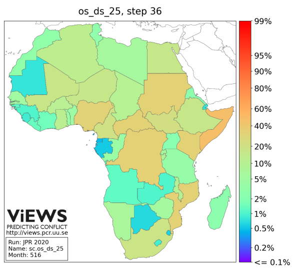

*(j)* vdem_glob



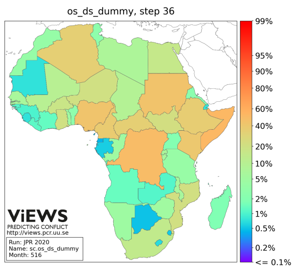

*(k)* demog



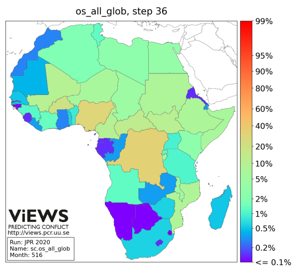

*(l)* WDI, global



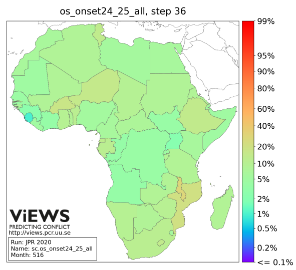

*(m)* Dynasim 25 deaths




*(n)* Dynasim 1 death




*(o)* All features, global



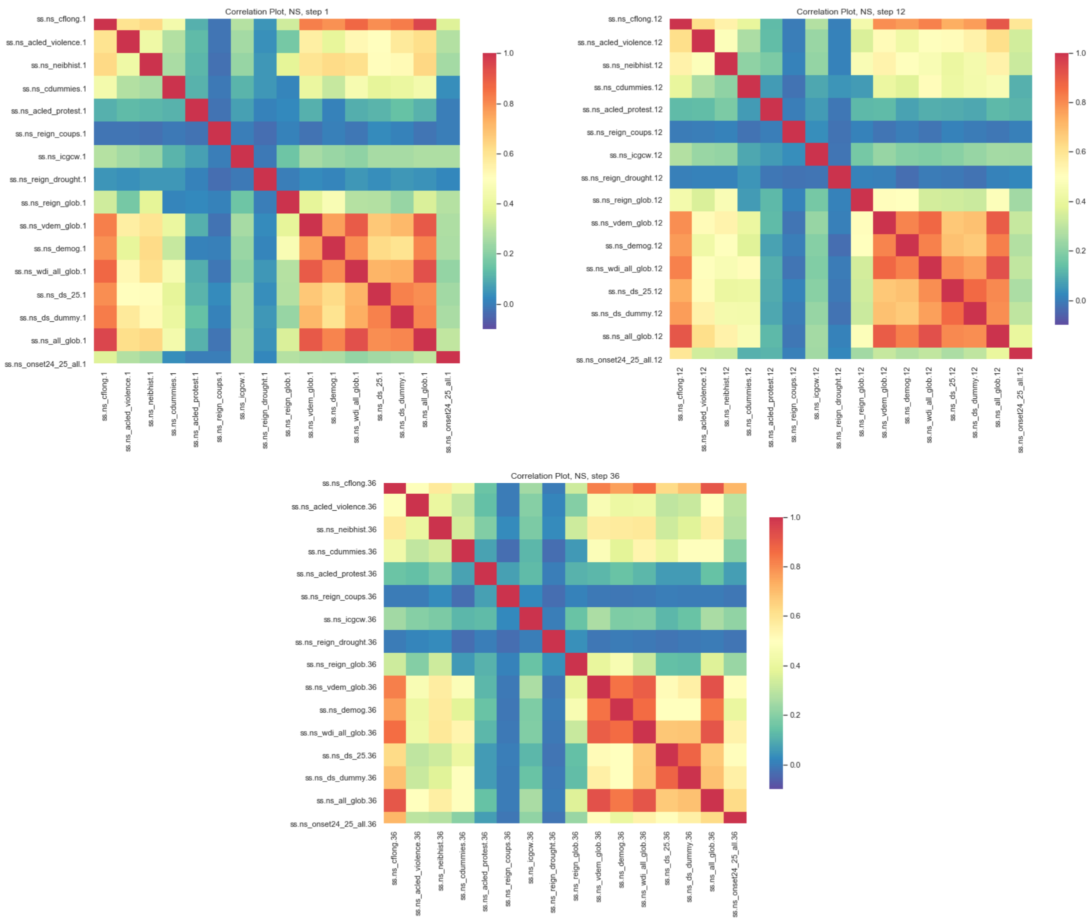

*(p)* Onset, 25+ deaths

### B-5.3 Non-state conflict (ns)

Non-state conflict is defined as the use of armed force by two formally organized group, none of which is the government of a state, which result in at least 25 BRDs in a year (Pettersson and Eck, 2018). Conflicts between rebel or ethnic groups fall in this category.

Figure B-8 shows the correlations between the predictions from each of the constituent models at three different steps ($s = 1, 12, 36$).

*Figure B-8.* Correlation between predictions from constituent models, *cm*, **ns**



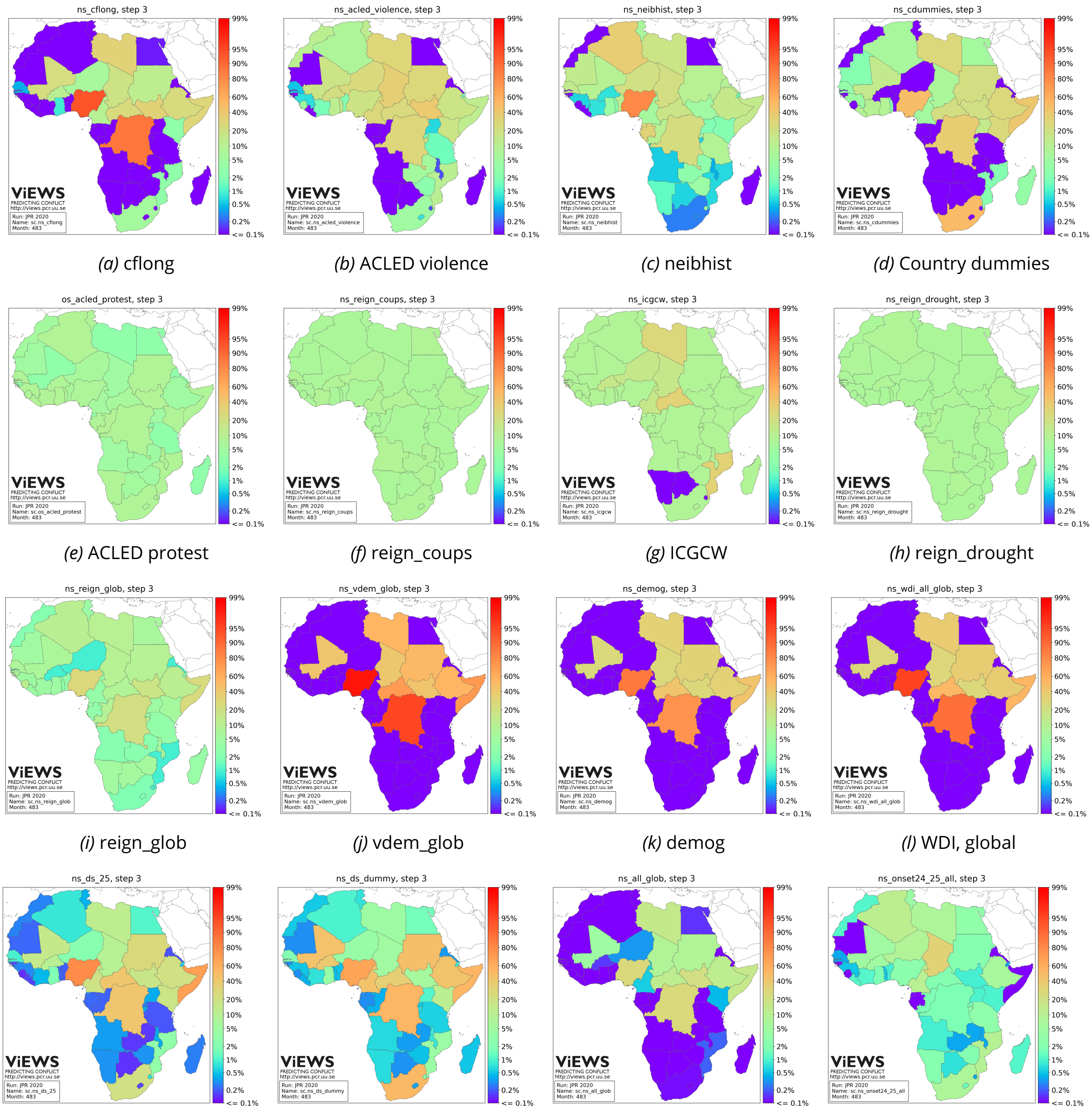

As we observed for the other types of violence, also for **ns** global models tend to be highly correlated among each other. The *V-Dem* global model especially shows correlation with a number of other models, including the other global ones, but also the *cflong* and the *Dynasim* models. The high correlation between models based on political regime (V-Dem), development (WDI) and conflict history (Cflong) reflects the ability of these models to capture structural, slow-moving features and patterns that characterize countries relatively long-term. Models including more rapidly moving features, akin political turbulence (*reign_coups*, *ACLED_protest* and *ACLED_violence*) or weather-related shocks (*reign_drought*) show less correlation with the other models. These trends in correlation increase over time, as shown by the correlation plots for steps 12 and 36, displaying a even higher correlation between slow-moving structural models.

As follows, we present the prediction maps for **ns**.

#### Prediction maps, forecasts (ns)

Here we present the forecasts for each of the constituent models in the ensembles for steps $s = (1, 3, 12, 36)$ at the *cm* level. Figure B-9 shows the forecasts for March 2020 based on the $s = 3$ model. Figure B-2 shows the forecast for the ensemble for the same month. All models are calibrated, i.e. rescaled using a logit procedure so that mean predicted probability in the calibration period is similar to the observed proportion of conflict. Models that separate well between conflict and non-conflict have multi-colored maps. Weak models have green colors, since predicted probabilities deviate little from the mean probability of 0.11. Figure B-10 show the same for December 2022, i.e. $s = 36$ months into the future based on data up to and including December 2019.

*Figure B-9.* Predicted probabilities, constituent models, **ns**, $s = 3$ (month 483, March 2020), based on data up to December 2019



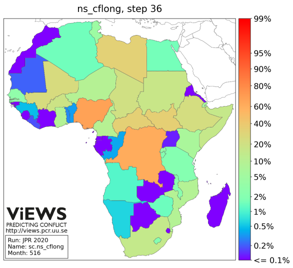

*(a)* cflong *(b)* ACLED violence *(c)* neibhist *(d)* Country dummies

*(e)* ACLED protest *(f)* reign_coups *(g)* ICGCW *(h)* reign_drought

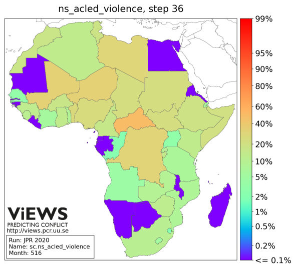
*(i)* reign_glob *(j)* vdem_glob *(k)* demog *(l)* WDI, global

*(m)* Dynasim 25 deaths *(n)* Dynasim 1 death *(o)* All features, global *(p)* Onset, 25+ deaths

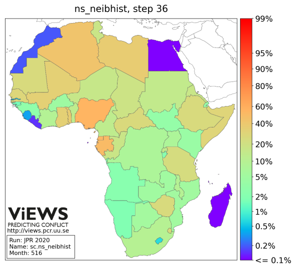
*Figure B-10.* Predicted probabilities, constituent models, **ns**, $s = 36$ (month 516, December 2022), based on data up to December 2019



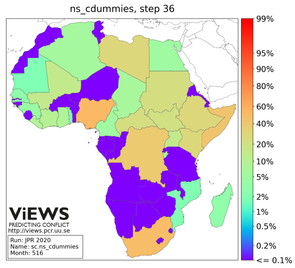

*(a)* cflong



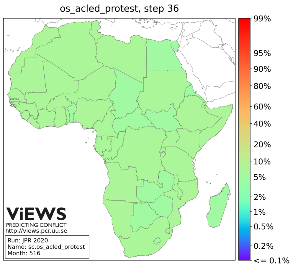

*(b)* ACLED violence



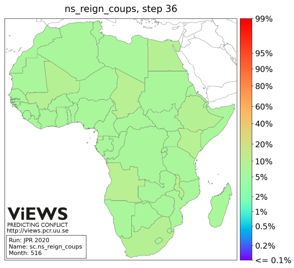

*(c)* neibhist



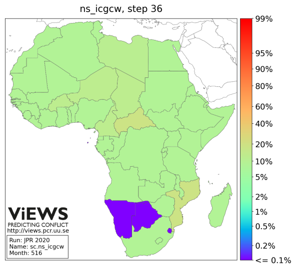

*(d)* Country dummies



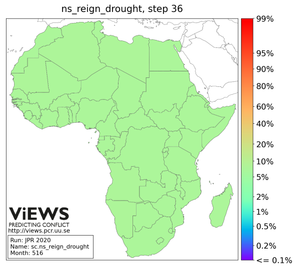

*(e)* ACLED protest



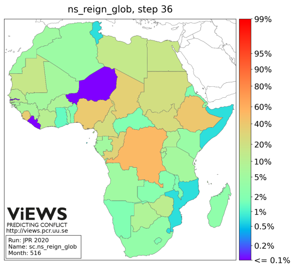

*(f)* reign_coups



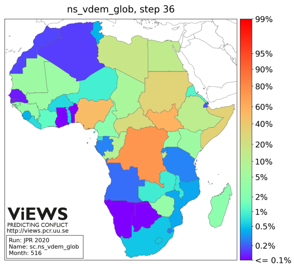

*(g)* ICGCW



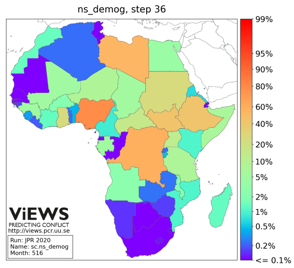

*(h)* reign_drought



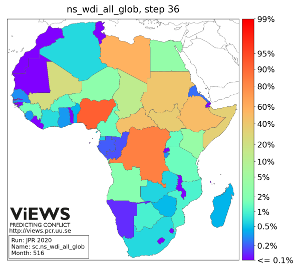

*(i)* reign_glob



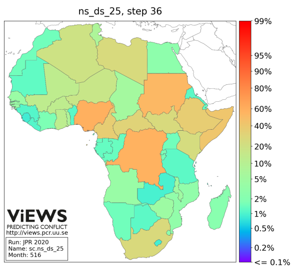

*(j)* vdem_glob



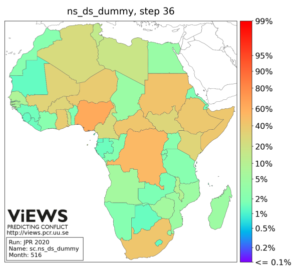

*(k)* demog



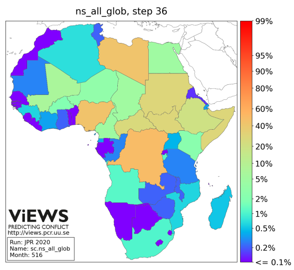

*(l)* WDI, global



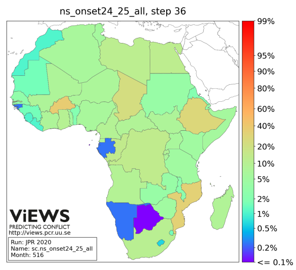

*(m)* Dynasim 25 deaths




*(n)* Dynasim 1 death




*(o)* All features, global




*(p)* Onset, 25+ deaths

## REFERENCES


Acemoglu, Daron and James A. Robinson (2001). "A Theory of Political Transitions". In: *American Economic Review* 91.4, pp. 938–963.

Belkin, Aaron and Evan Schofer (2003). "Toward a Structural Understanding of Coup Risk". In: *Journal of Conflict Resolution* 47.5, pp. 594–620.

Bell, Curtis (2016a). *The Rulers, Elections, and Irregular Governance Dataset (REIGN)*. Broomfield, CO: OEF Research. Available at oefresearch.org. URL: oefresearch.org.

— (2016b). *The Rulers, Elections, and Irregular Governance Dataset (REIGN)*. Broomfield, CO: OEF Research. Available at oefresearch.org. URL: oefresearch.org.

Bell, Curtis and Jun Koga Sudduth (2017). "The Causes and Outcomes of Coup during Civil War". In: *Journal of Conflict Resolution* 61.7, pp. 1432–1455. DOI: 10.1177/0022002715603098. URL: https://doi.org/10.1177/0022002715603098.

Benoit, Kenneth (1996). "Democracies Really are more Pacific (in General): Reexamining Regime Type and War Involvement". In: *Journal of Conflict Resolution* 40.4, pp. 636–657.

Besley, Timothy and Marta Reynal-Querol (May 2014). "The Legacy of Historical Conflict: Evidence from Africa". en. In: *American Political Science Review* 108.2, pp. 319–336. DOI: 10.1017/S0003055414000161. URL: https://www.cambridge.org/core/product/identifier/S0003055414000161/type/journal_article (visited on 02/04/2020).

Birch, Sarah, Ursula Daxecker, and Kristine Höglund (2020). "Electoral violence: An introduction". In: *Journal of Peace Research* 57.1, pp. 3–14. DOI: 10.1177/0022343319889657. URL: https://doi.org/10.1177/0022343319889657.

Breiman, Leo (2001). "Random Forests". In: *Machine learning* 45.1, pp. 5–32.

Buhaug, Halvard and Scott Gates (2002). "The Geography of Civil War". In: *Journal of Peace Research* 39.4, pp. 417–433. DOI: 10.1177/0022343302039004003. URL: https://doi.org/10.1177/0022343302039004003.

Buhaug, Halvard and Kristian Skrede Gleditsch (2008). "Contagion or Confusion? Why Conflicts Cluster in Space". In: *International Studies Quarterly* 52, pp. 215–233.

Burchard, Stephanie M. (2015). *Electoral Violence in Sub-Saharan Africa*. Boulder, CO: Lynne Rienner.

Cederman, Lars-Erik, Andreas Wimmer, and Brian Min (2010). "Why do ethnic groups rebel? New data and analysis". In: *World Politics* 62.1, pp. 87–119.

Chen, Tianqi and Carlos Guestrin (2016). "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794.

Collier, Paul, Lani Elliot, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol, and Nicholas Sambanis (2003). *Breaking the Conflict Trap. Civil War and Development Policy*. Oxford: Oxford University Press. URL: https://openknowledge.worldbank.org/handle/10986/13938.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Steven Wilson, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Nina Ilchenko, Joshua Krusell, Laura Maxwell, Valeriya Mechkova, Juraj Medzihorsky, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, and Daniel Ziblatt (2019). *V-Dem Country-Year Dataset v9*. URL: https://doi.org/10.23696/vdemcy19.

Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Steven M. Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Constanza Sanhueza Petrarca, Johannes von Römer, Laura Saxer, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova, and Steven Wilson (2017). *V-Dem Country-Year Dataset v7.1*. Varieties of Democracy (V-Dem) Project.

Davenport Christian, David A. Armstrong II and Mark Lichbach (2005). *Conflict Escalation and the Origins of Civil War*. Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, Illinois, USA.

Friedman, Jerome H. (2001). "Greedy function approximation: a gradient boosting machine". In: *Annals of statistics*, pp. 1189–1232.

Géron, Aurélien (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, Inc.

Hegre, Håvard, Marie Allansson, Matthias Basedau, Mike Colaresi, Mihai Croicu, Hanne Fjelde, Frederick Hoyles, Lisa Hultman, Stina Högbladh, Remco Jansen, Naima Mouhleb, Sayeed Auwn Muhammad, Desirée Nilsson, Håvard Mokleiv Nygård, Gudlaug Olafsdottir, Kristina Petrova, David Randahl, Espen Geelmuyden Rød, Gerald Schneider, Nina von Uexkull, and Jonas Vestby (2019). "ViEWS: A political Violence Early Warning System". In: *Journal of Peace Research* 56.2, pp. 155–174. URL: `https://doi.org/10.1177/0022343319823860`.

Hegre, Håvard, Halvard Buhaug, Katherine V. Calvin, Jonas Nordkvelle, Stephanie T. Waldhoff, and Elisabeth Gilmore (2016). "Forecasting civil conflict along the shared socioeconomic pathways". In: *Environmental Research Letters* 11.5, p. 054002. DOI: `10.188/1748-9326/11/5/054002`.

Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch (2001). "Toward a democratic civil peace? Democracy, political change, and civil war, 1816–1992". In: *American Political Science Review* 95.1, pp. 33–48.

Hegre, Håvard, Joakim Karlsen, Håvard Mokleiv Nygård, Håvard Strand, and Henrik Urdal (2013). "Predicting Armed Conflict 2010–2050". In: *International Studies Quarterly* 57.2, pp. 250–270. DOI: `10.1111/isqu.12007`.

Koubi, Vally (2019). "Climate Change and Conflict". In: *Annual Review of Political Science* 22.1, pp. 343–360.

Muchlinski, David, David Siroky, Jingrui He, and Matthew Kocher (2016). "Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data". In: *Political Analysis* 24.1, pp. 87–103. DOI: `10.1093/pan/mpv024`.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer (2019). *The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data*. V-Dem Working Paper No. 21. 4th edition. University of Gothenburg: Varieties of Democracy Institute.

Pettersson, Therése and Kristine Eck (2018). "Organized violence, 1989–2017". In: *Journal of Peace Research* 55.4, pp. 535–547. DOI: `10.1177/0022343318784101`. URL: `https://doi.org/10.1177/0022343318784101`.

Raleigh, Clionadh, Håvard Hegre, Joakim Karlsen, and Andrew Linke (2010). "Introducing ACLED: An Armed Conflict Location and Event Dataset". In: *Journal of Peace Research* 47.5, pp. 651–660. URL: `https://doi.org/10.1177/0022343310378914`.

Schwartz, Peter and Doug Randall (2003). *An Abrupt Climate Change Scenario and its Implications for United States National Security*. Report prepared for the US Department of Defense. US Department of Defense.

Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin, and Achim Zeileis (2008). "Conditional variable importance for random forests". In: *BMC Bioinformatics* 9.1, p. 307. DOI: `10.1186/1471-2105-9-307`. URL: `https://doi.org/10.1186/1471-2105-9-307`.

Strobl, Carolin, Anne-Laure Boulesteix, Achim Zeileis, and Torsten Hothorn (2007). "Bias in random forest variable importance measures: Illustrations, sources and a solution". In: *BMC Bioinformatics* 8.1, p. 25. DOI: `10.1186/1471-2105-8-25`. URL: `https://doi.org/10.1186/1471-2105-8-25`.

von Uexkull, Nina (2014). “Sustained drought, vulnerability and civil conflict in Sub-Saharan Africa”. In: *Political Geography* 43, pp. 16–26.
von Uexkull, Nina, Mihai Croicu, Hanne Fjelde, and Halvard Buhaug (2016). “Civil conflict sensitivity to growing-season drought”. In: *Proceedings of the National Academy of Sciences* 113.44, pp. 12391–12396.
Vreeland, James Raymond (2008). “The Effect of Political Regime on Civil War: Unpacking Anocracy”. In: *Journal of Conflict Resolution* 52.3, pp. 401–425.
World Bank (1999). *World Development Indicators 1999*. Washington, DC: Development Data Center, International Bank for Reconstruction and Development.
— (2019). *World Development Indicators*. Washington DC: The World Bank. URL: `https://datacatalog.worldbank.org/dataset/world-development-indicators`.

**FUNDING**

ViEWS is funded by the European Research Council, grant number 694640 and Uppsala University.

European Research Council
Established by the European Commission

**COLLABORATIONS**

ViEWS has an active interaction with other projects, including CLIMSEC, CAVE and CROP at PRIO (`https://prio.org/`), the MISTRA Geopolitics project, and most importantly the Uppsala Conflict Data Program (`https://ucdp.uu.se/`) at Uppsala University.

**CODEBASE & PUBLICATIONS**

ViEWS’ codebase is available at:

`https://github.com/UppsalaConflictDataProgram/OpenViEWS2`

The full pool of publications are accessible at:

`https://pcr.uu.se/research/views/publications/`