Models in VIEWS, version Fatalities002*

Håvard Hegre[1, 2], Sofia Nordenving[2], and James Dale[2]

[1]Peace Research Institute Oslo (PRIO)
[2]Department of Peace and Conflict Research, Uppsala University

March 12, 2024
VIEWS Version: Fatalities002

Abstract

In this paper, we describe the current models in the main ensemble and give an overview of the implemented algorithms.

*The research was funded by the European Research Council, project H2020-ERC-2015-AdG 694640 (ViEWS), the European Research Council Advanced Grant project *ANTICIPATE*, the Research Council of Norway, Riksbankens Jubileumsfond project *Societies at Risk*, Uppsala University, the Swedish Research Council project *DEMSCORE*, the UK Foreign, Commonwealth and Development Office, the United Nations High Commissioner for Refugees, and Peace Research Institute Oslo. The simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX). For more information on the project see `viewsforecasting.org`.

Contents

1 The dependent variable

The outcome that the model predicts is armed conflict as defined and compiled by the Uppsala Conflict Data Program (UCDP, Gleditsch et al., 2002; Sundberg and Melander, 2013; Pettersson et al., 2021; Hegre et al., 2020). The UCDP collects data on three types of conflict (see `https://www.pcr.uu.se/research/ucdp/definitions/`):

State-based (sb) conflict The use of armed conflict over either government or territory between armed actors in which at least one is a government of a state.

Non-state (ns) conflict The use of armed force between two or more organised armed groups, neither of which is a government of a state.

One sided (os) conflict The deliberate use of armed force by the government of a state or by a formally organised group against civilians.

The UCDP provides estimates for the number of persons killed in each of these three conflict types for each of the conflict events they can document. We aggregate the fatalities across events into monthly sums, for countries and for the PRIO-GRID cell structure (Tollefsen, Strand, and Buhaug, 2012a).

In the Fatalities 002 version all three types of conflict are used as input data but we only predict fatalities from State-based violence. The dependent variable will soon be extended to include the other two forms of violence. A more thorough description of the outcome variable can be found in *Levels of analysis and the dependent variables* (Hegre et al., 2023b).

2 Input Features

Fatalities002 is informed by data on hundreds of variables from data providers such as the Uppsala Conflict Data Program (UCDP), Varieties of Democracy (V-Dem)(Coppedge et al., 2022), ACLED (Raleigh, Kishi, and Linke, 2023), PRIO-GRID (Tollefsen, Strand, and Buhaug, 2012b), the World Development Indicators (The World Bank, 2022), IMF (World Economic Outlook Database, 2019), FAO (Food and Agriculture Organization of the United Nations, 1997), Mapspam (International Food Policy Research Institute, 2019), SPEI[1] (Vicente-Serrano, Beguería, and López-Moreno, 2010) and MIRCA (Portmann, Siebert, and Döll, 2010).

Based on these raw data variables, VIEWS also construct a suite of additional variables by applying data transformations such as time and space lags, imputations to fill in for missing data, and other common data processing techniques. Together, the raw and processed data variables informing the various VIEWS models are referred to as features, which are grouped into feature sets based on; the overall theme they relate to, the data provider(s) from which they are derived, and to optimize performance. An overview at the country level is presented in Table 1 and at PRIO-GRID level in Table 2. The full list of input variables and transforms can be found in the *viewsforecasting* repository, for country level see cm_querysets and for PRIO-GRID level see pgm_querysets. For further details on transform please consult the source code on VIEWS Transformations Library.

[1]SPEI data are drawn from the Global Drought Monitor (`spei.csic.es`)

Model	Feature set	Description
Baseline	qs_baseline	Baseline feature set with a few conflict history features as well as log population.
Conflict History	qs_conflict	A collection of variables that together map the conflict history of a country
Extended Conflict History	qs_conflict_long	A collection of variables that together map the extended conflict history of a country.
Democracy	qs_vdem_short	A collection of variables that capture democracy indices and the strength of political institutions in each country, such as liberal democracy, rule of law, equality, and the level of exclusion of social groups in politics.
Development Indicators	qs_wdi_short	Measures of development as provided by the World Bank Indicators, e.g., GDP per capita, infant mortality rate, and school enrollment.
News Topics	qs_topics	A collection of variables from the Mueller & Rauh (2018) topic model, capturing conflict risks based on the share of certain topics in news media.
Combined Extended	qs_joint_broad	A broad feature set that combines data from the topics data, varieties of democracy variables, world development indicators and on conflict history.
Combined Narrow	qs_joint_narrow	A narrow version of the joint broad set that combines a few variables from the topics model, varieties of democracy variables, world development indicators and on conflict history.
All Features	qs_all_features	A combination of baseline features from the conflict, topics, wdi, vdem and aquastat sets
Water and Agriculture	qs_aquastat	A feature set that combines a number of variables on water availability and agriculture from AQUSTAT
Food Security	qs_faostat	A collection of variables related to food security from FAOSTAT
Food Prices	qs_faoprices	A combination of food price indicators from FAOSTAT
GDP	qs_imfweo	A collection of features on past and predicted annual percentage change of Gross domestic product

Table 1. Feature set descriptions country level of analysis

Human name	Feature set	Description
Baseline	qs_baseline	Baseline features with a few conflict history variables as well as log population
Extended Conflict History	qs_conflict_long	A collection of variables that together map the sub-national conflict history.
Drought	qs_escwa_drought	A combination of baseline conflict history and drought vulnerability indicators
Natural and Social Geography	qs_natsoc	Basline conflict set together with natural and social geography features
Combined	qs_broad	A broad combination of important features from the other feature sets
Conflict History	qs_conf_history	A broad set mapping out detailed sub-national conflict history
Tree lag Conflict History	qs_treelag	A conflict history set using tree lag transforms
Space Time Conflict History	qs_sptime_dist	A conflict history set using spacetime transforms

Table 2. Feature set descriptions PRIO-GRID level of analysis

The input features can be divided into eight broad thematic categories of conflict drivers at the two levels of analysis (read more about VIEWS levels of analysis in Hegre et al., 2023b).

At both country and PRIO-GRID level, the *Conflict history* thematic category captures the history of conflict in each country and sub-national grid cell, e.g., the number of battle-related deaths per unit and level of analysis, and measures of the temporal and spatial distance to recent conflict events. Data providers are Uppsala Conflict Data Program (UCDP) and The Armed Conflict Location & Event Data Project (ACLED). The *Development* category include measures of development as provided by the World Bank (World Development Indicators) such as GDP per capita, infant mortality rate, and school enrollment. *Climate & societal vulnerability* feature sets captures climate hazards and societal vulnerability to these hazards, includingdrought indicators, reliance on agriculture, crop yields, precipitation, freshwater withdrawal, water management efficiency, and access to renewable resources. Data providers are the United Nations Food and Agriculture Organisation (FAO), FAO AQUASTAT, PRIO-GRID, MIRCA, MAPSPAM, SPEI Global Drought Monitor. *Food security and access to basic needs* include indicators of staple food prices along with measures of food security and access to basic human needs, such as mean food prices, food price inflation, undernourishment, access to clean water, and basic sanitation. These data are provided by United Nations Food and Agriculture Organisation (FAO), FAOSTAT.

At the country level, the *Political institutions, democracy* category capture democracy indices and the strength of political institutions in each country, such as liberal democracy, rule of law, equality, and the level of exclusion of social groups in politics. The main data provider is Varieties of Democracy (V-Dem). *Economic growth*, focuses specifically on historic and future economic growth, e.g., real GDP growth per year and growth forecasts for the coming years using data from The International Monetary Fund World Economic Outlook (IMF WEO). At the country level, VIEWS predictions as inform through *News monitoring* based on the Mueller & Rauh topic model (Mueller and Rauh, 2022), which captures conflict risks as drawn from a topic analysis of news media.

Exclusively on PRIO-GRID level, *Natural and social geography* captures terrain type, distance to natural resources, demography, proximity to cities and country borders provided by PRIO-GRID.

Categorizing input data variables into feature sets is part of the standard data organization routines in VIEWS, which greatly facilitates model development. Amongst other benefits, it allows us to call upon a pre-determined set of features, which is maintained in a single location, when training our models. This minimizes the risk of human error when compiling the input datasets and greatly facilitates maintenance of the model documentation.

3 Sub-models and machine learning algorithms

As a first step when training the model, each feature set is paired with an advanced machine learning algorithm (Tables 3 and 4). We have explored a number of algorithms to relate the feature sets to the outcome we seek to predict. Most of the models in Fatalities002 are tree-based models. With the exception of the GLM Markov model, none of the generalized linear models we tried yielded good performance. The Fatalities002 model uses two types of tree-based algorithms: Random forest (XGBRF Regressor, RandomForestRegressor) and Gradient boosting (GradientBoostingRegressor), including Light gradient boosting (LGBMRegressor) and Extreme gradient boosting (XGBRegressor). These algorithms are also used in two multi-stage models, the hurdle models (Hurdleregression) and Hierarchical Markov models (Markov_glm, Markov_rf).

For certain sub-models we use a PCA pre-processing function (Principle Component Analysis) to reduce the dimensionality in the data. For more details on algorithms and hyperparameters please consult Model Definitions.

Random forests XGBRFRegressor () or scikit's RandomForestRegressor (), implementing the Random Forest (Breiman, 2001) algorithm, an ensemble of decision trees, with each tree trained on a subset of features and bootstrapped data – with the aggregate ensemble reducing.

Gradient boosting models scikit's GradientBoostingRegressor(). Gradient Boosting Regressors (GBR) are another ensemble method improving decision trees sequentially by training each iteration on the residual of the past iteration. The algorithm starts by assigning equal weights to all data points. It then iteratively changes the weights by increasing the weight assigned to difficult observations that are misclassified, and lowering the weight for data points that are easy to classify or are correctly classified.

Extreme gradient boosting XGBRegressor (n_estimators=100, learning_rate = 0.05). The model is estimated using extreme gradient boosting (Chen and Guestrin, 2016). We use the XGBoost implementation, and performed an 'early-stopping' routine to identify the optimal number of estimators and learning rate parameters.

Light gradient boosting LightGBM (LGBMRegressor, n_estimators=100) is another gradient boosting method based on decision trees to increase the efficiency of the model and reduce memory usage. It uses novel techniques (One Side Sampling and Exclusive Feature Bundling) to overcome the limitations of histogram-based algorithms. The Light Gradient Boosting works by retaining instances that with larger gradients(those that contain more information but are under-trained) and randomly dropping data-points with small gradients. This leads to a more accurate estimation than uniformly random sampling.

Hurdle models As indicated by the study of the prediction outcome in Hegre et al. (2023b), most of the observations in our input data have no recorded fatalities, and the distribution of the non-zero observations are highly right-skewed. There are reasons to think that the data-generating process that leads to whether a country or a grid cell has any fatalities at all is quite different from the one that lead to subsequent fatalities (Fritz et al., 2021). Hurdle models take this into account by dividing the outcome into two variables, a dichotomous variable for whether there was non-zero fatalities or not, and the log count of fatalities if there was at least one fatality. The model then trains a classifier for the zero/non-zero distinction, and a regressor for the non-zero observations. At the predict stage, the predicted number of fatalities is the product of the probability of non-zero observations, and the expected number of fatalities given there is at least one, conditional on the predictors. We have explored a number of variants of the hurdle models, using classifier and regression versions of the tree-based algorithms described above.[2]

Markov_glm and Markov_rf Markov models are a more sophisticated formulation of the hurdle-model idea that different models do well in different situations. The models use an observed Markov modeling approach with four different latent states which produce fatalities, and where the forecast of fatalities is conditional on the likelihood of the conflict state. The four conflict states used are the same as in Randahl and Vegelius (2022a), that is, 'peace', 'escalation', 'de-escalation', and 'conflict'. Transitions between states are restricted such that each state only has two possible future states. The transitions allowed are from peace to peace and to escalation, from escalation to conflict and to deescalation, from deescalation to peace and to escalation, and from conflict to conflict and to deescalation. The escalation and deescalation states are thus transient, as they do not allow transitions to themselves. Transitions between states are modelled as a binary logistic regression model, and the log number of fatalities conditional on the latent states are modelled using OLS regression. The **Markov_glm** version use logistic and linear regression models, and the **Markov_rf** models the transition between states using a random forest classifier, and the log number of fatalities conditional on the latent states using a random forest regressor. For more details on the Markov modeling approach, see (Randahl and Vegelius, 2022b).

4 Final ensemble – using "the wisdom of the crowd"

Much like a crowd tends to be wiser than the individuals composing it, prediction models that are informed by a number of smaller and specialized sub-models are known to be more robust and generate stronger predictions than single models. As a second step in the model training procedures, the sub-models above are therefore combined into two groups or ensembles of models – one ensemble for each level of analysis.

Two different ensembling techniques were used for this purpose: The country-level ensemble model combines the predictions from each of the sub-models using a genetic algorithm that assigns different weights to the contribution from each model in order to maximise predictive performance. The sub-national ensemble model, on the other hand, uses a simple unweighted average of the sub-model results. The ensembling techniques above are motivated and described further in (Hegre et al., 2023a).

[2] Our hurdle model implementation is based on code developed by Geoff Hurdock: `https://geoffruddock.com/building-a-hurdle-regression-estimator-in-scikit-learn/`

Human name	Model name	Feature set	Algorithm	PCA
Baseline Random Forest	fatalities002_baseline_rf	qs_baseline	Random forest regressor, XGBoost implementation	False
Conflict History Random Forest	fatalities002_conflicthistory_rf	qs_conflict_history	Random forest regressor, XGBoost implementation	False
Conflict History Gradient Boosting	fatalities002_conflicthistory_gbm	qs_conflict	Gradient boosting regressor	False
Conflict History Hurdle	fatalities002_conflict-history_hurdle_lgb	qs_conflict	Hurdle model with Light gradient boost classifier and regressor	False
Conflict History XGBoost	fatalities002_conflicthistory_long_xgb	qs_conflict_long	Random forest regressor, XGBoost implementation	False
Democracy Hurdle	fatalities002_vdem_hurdle_xgb	qs_vdem_short	Hurdle model with XGBoost classifier and regressor	False
Development Indicators Random Forest	fatalities002_wdi_rf	qs_wdi_short	Random forest regressor, XGBoost implementation	False
News Topics Random Forest	fatalities002_topics_rf	qs_topics	Random forest regressor, XGBoost implementation	False
News Topics Gradient Boosting	fatalities002_topics_xgb	qs_topics	Gradient boosting regressor	False
News Topics Hurdle	fatalities002_topics_hurdle_lgb	qs_topics	Hurdle model with Light gradient boost classifier and regressor	False
Combined Extended Random Forest	fatalities002_joint_broad_rf	qs_joint_broad	Random forest regressor, XGBoost implementation	False
Combined Extended Hurdle	fatalities002_joint_broad_hurdle_rf	qs_joint_broad	Hurdle model with Random forest classifier and regressor	False
Combined Narrow Random Forest	fatalities002_joint_narrow_xgb	qs_joint_narrow	Random forest regressor, XGBoost implementation	False
Combined Narrow Hurdle	fatalities002_joint_narrow_hurdle_xgb	qs_joint_narrow	Hurdle model with XGBoost classifier and regressor	False
Combined Narrow Hurdle	fatalities002_joint_narrow_hurdle_lgb	qs_joint_narrow	Hurdle model with Light gradient boost classifier and regressor	False
All Features Random Forest	fatalities002_all_pca3_xgb	qs_all_features	Random forest regressor, XGBoost implementation	True
Water and Agriculture Random Forest	fatalities002_aquastat_rf	qs_aquastat	Random forest regressor, XGBoost implementation	False
Food Security Random Forest	fatalities002_faostat_rf	qs_faostat	Random forest regressor, XGBoost implementation	False
Food Prices Random Forest	fatalities002_faoprices_rf	qs_faoprices	Random forest regressor, XGBoost implementation	False
GDP Random Forest	fatalities002_imfweo_rf	qs_imfweo	Random forest regressor, XGBoost implementation	False
Combined Narrow Hurdle	fatalities002_Markov_glm	qs_joint_narrow	Markov model using logistic and linear regression	False
Combined Narrow Random Forest	fatalities002_Markov_rf	qs_joint_narrow	Markov model with Random forest classifier and regressor	False

Table 3. Models and Algorithms Country Level

Human name	Model name	Feature set	Algorithm	PCA
Baseline Light Gradient Boosting	fatalities002_pgm_baseline _lgbm	qs_baseline	Light gradient boosting regressor	False
Extended Conflict History Light Gradient Boosting	fatalities002_pgm_conflictlong _lgbm	qs_conflict_long	Light gradient boosting regressor	False
Extended Conflict History Hurdle	fatalities002_pgm_conflictlong _hurdle _lgbm	qs_conflict_long	Hurdle model with Light gradient boost classifier and regressor	False
Drought Hurdle	fatalities002_pgm_escwa_drought _hurdle_lgbm	qs_escwa_drought	Hurdle model with Light gradient boost classifier and regressor	False
Drough Light Gradient Boosting	fatalities002_pgm_escwa _drought_lgbm	qs_escwa_drought	Light gradient boosting regressor	False
Natural and Social Geography Hurdle	fatalities002_pgm_natsoc_hurdle _lgbm	qs_natsoc	Hurdle model with Light gradient boost classifier and regressor	False
Natural and Social Geography Light Gradient Boosting	fatalities002_pgm_natsoc_lgbm	qs_natsoc	Light gradient boosting regressor	False
Combined Extended Hurdle	fatalities002_pgm_broad_hurdle _lgbm	qs_broad	Hurdle model with Light gradient boost classifier and regressor	False
Combined Extended Light Gradient Boosting	fatalities002_pgm_broad_lgbm	qs_broad	Light gradient boosting regressor	False
Conflict History XGBoost	fatalities002_pgm_conflict_history _xgb	qs_conf_history	Random forest regressor, XGBoost implementation	False
Tree Lag Conflict History Hurdle	fatalities002_pgm_conflict _treelag_hurdle	qs_treelag	Hurdle regressor	False
Space Time Conflict History Hurdle	fatalities002_pgm_conflict_sptime _dist_hurdle	qs_sptime_dist	Hurdle regressor	False

Table 4. Models and Algorithms PRIO-GRID Level

5 Change history

5.1 Fatalities002

Only minor changes were made for model definitions transitioning to the Fatalities002 model version, limited to renaming and restructuring parts of the source code.

5.2 Fatalities001

The first version of the Fatalities model was introduced in (Hegre et al., 2022) and developed with funding support from UK FCDO. Model definitions for Fatalities001 can be found in the FCDO_predicting_fatalities repository

5.3 ViEWS-ESCWA

The VIEWS system was expanded to cover the Middle East (including Turkey and Iran) thanks to funding from the UN ESCWA Theisen et al. (2021).

5.4 ViEWS2020

ViEWS2020 was introduced (Hegre et al., 2021). This update introduced a new infrastructure for training, evaluating, and weighting models that allowed for more optimal combining of models into ensembles and a number of new forecasting models that contributed to improve overall performance, in particular with respect to effectively classifying high- and low-risk cases.

5.5 ViEWS2018

The first version of the ViEWS early warning system, the 'ViEWS2018' version launched in July 2018 (Hegre et al., 2019). This first version of VIEWS used a dynamic simulation procedure (for more details see (Hegre et al., 2016)) to predict conflict probability.

References

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

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.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I Lindberg, Jan Teorell, A Altman, Michael Bernhard, A Cornell, MS Fish, L Gastaldi, et al. (2022). "VDem [Country–Year/Country–Date] Dataset v12. Varieties of Democracy (V-Dem) Project". In: URL: `https://doi.org/10.23696/vdemds22`.

Food and Agriculture Organization of the United Nations (1997). *FAOSTAT statistical database.*

Fritz, Cornelius, Marius Mehrl, Paul W. Thurner, and Göran Kauermann (2021). "The role of governmental weapons procurements in forecasting monthly fatalities in intrastate conflicts: A semiparametric hierarchical hurdle model". In: *International Interactions* 48.x, pp. 1–22. eprint: `https://doi.org/10.1080/03050629.2022.1993210`. URL: `https://doi.org/10.1080/03050629.2022.1993210`.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand (2002). "Armed conflict 1946–2001: A new dataset". In: *Journal of peace research* 39.5, pp. 615–637.

Hegre, Håvard, Marie Allansson, Matthias Basedau, Mike Colaresi, Mihai Catalin 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, Curtis Bell, Michael Colaresi, Mihai Croicu, Frederick Hoyles, Remco Jansen, Maxine Ria Leis, Angelica Lindqvist-McGowan, David Randahl, Espen Geelmuyden Rød, and Paola Vesco (2021). "ViEWS$_{2020}$: Revising and evaluating the ViEWS political Violence Early-Warning System". In: *Journal of Peace Research* 58.3, pp. 599–611. eprint: https://doi.org/10.1177/0022343320962157. URL: https://doi.org/10.1177/0022343320962157.

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.

Hegre, Håvard, Michael Colaresi, Sofia Nordenving, Mihai Croicu, James Dale, and Paola Vesco (2023a). *Ensembling and calibration in VIEWS*. VIEWS documentation paper series, Fatalities002. Fatalities002 Documentation. URL: https://viewsforecasting.org/wp-content/uploads/VIEWS_documentation_Ensembling_Fatalities002.pdf.

Hegre, Håvard, Mihai Catalin Croicu, Kristine Eck, and Stina Högbladh (2020). "Introducing the UCDP Candidate Events Dataset". In: *Research & Politics* 7.3, p. 2053168020935257. eprint: https://doi.org/10.1177/2053168020935257. URL: https://doi.org/10.1177/2053168020935257.

Hegre, Håvard, James Dale, David Randahl, Mihai Croicu, Remco Jansen, Paola Vesco, Angelica Lindqvist-McGowan, Peder Landsverk, Maxine Leis, Hannes Mueller, Christopher Rauh, Espen Geelmuyden Rød, Forogh Akbari, Tim Gåsste, and Malika Rakhmankulova (2022). *Forecasting Fatalities*. Uppsala: working paper. URL: https://www.diva-portal.org/smash/get/diva2:1667048/FULLTEXT01.pdf.

Hegre, Håvard, Sofia Nordenving, Mihai Croicu, James Dale, and Paola Vesco (2023b). *Levels of analysis and the dependent variables*. VIEWS documentation paper series, Fatalities002. URL: https://viewsforecasting.org/wp-content/uploads/VIEWS_documentation_LevelsandOutcomes.pdf.

International Food Policy Research Institute (2019). *Global spatially-disaggregated crop production statistics data for 2010 version 2.0*.

Mueller, Hannes and Christopher Rauh (Apr. 2022). "The Hard Problem of Prediction for Conflict Prevention". In: *Journal of the European Economic Association*. jvac025. eprint: https://academic.oup.com/jeea/advance-article-pdf/doi/10.1093/jeea/jvac025/43465294/jvac025.pdf. URL: https://doi.org/10.1093/jeea/jvac025.

Pettersson, Therése, Shawn Davies, Amber Deniz, Garoun Engström, Nanar Hawach, Stina Högbladh, Margareta Sollenberg, and Magnus Öberg (2021). "Organized violence 1989–2020, with a special emphasis on Syria". In: *Journal of Peace Research* 58.4, pp. 809–825. eprint: https://doi.org/10.1177/00223433211026126. URL: https://doi.org/10.1177/00223433211026126.

Portmann, Felix T, Stefan Siebert, and Petra Döll (2010). "MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling". In: *Global biogeochemical cycles* 24.1.

Raleigh, Clionadh, Roudabeh Kishi, and Andrew Linke (2023). "Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices". In: *Humanities and Social Sciences Communications* 10.1, pp. 1–17.

Randahl, David and Johan Vegelius (2022a). "Predicting escalating and de-escalating violence in Africa using Markov models". In: *International Interactions* 48.4, pp. 597–613. eprint: https://doi.org/10.1080/03050629.2022.2049772. URL: https://doi.org/10.1080/03050629.2022.2049772.

— (2022b). "Predicting escalating and de-escalating violence in Africa using Markov Models". In: *International Interactions* 48.x, pp. 000–000.

Sundberg, Ralph and Erik Melander (2013). "Introducing the UCDP Georeferenced Event Dataset". In: *Journal of Peace Research* 50.4, pp. 523–532.

The World Bank (2022). *World Development Indicators*.

Theisen, Ole Magnus, Håvard Hegre, Halvard Buhaug, Stefan Döring, Remco Jansen, Angelica Lindqvist-McGowan, Ida Rudolfsen, Paoal Vesco, Joaquin Salido Marcos, Yara Acaf, Pattile Nahabedian, and Lubna Ismail (2021). *Understanding the Potential Linkages between Climate Change and Conflict in the Arab Region*. E/ESCWA/CL6.GCP/2021/TP.9. UN ESCWA, Beirut, Lebanon.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug (2012a). “PRIO-GRID: A unified spatial data structure”. In: *Journal of Peace Research* 49.2, pp. 363–374.

— (2012b). “PRIO-GRID: A unified spatial data structure”. In: *Journal of Peace Research* 49.2, pp. 363–374.

Vicente-Serrano, Sergio M, Santiago Beguería, and Juan I López-Moreno (2010). “A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index”. In: *Journal of climate* 23.7, pp. 1696–1718.

World Economic Outlook Database, October (2019). *International monetary fund.*