A review and comparison of conflict early warning systems☆

Espen Geelmuyden Rød [a,*], Tim Gåsste [a], Håvard Hegre [a,b]

[a] *Department of Peace and Conflict Research, Uppsala University, Sweden*
[b] *Peace Research Institute, Oslo, Norway*

ARTICLE INFO

ABSTRACT

We review and compare conflict early warning systems on three dimensions: transparency and accessibility, key parameters, and forecasts. The review reveals a need for improved transparency and accessibility of data and code, considerable variation in key parameters across systems, and significant overlaps in countries with the highest risk. We propose that developing standards and platforms that promote transparency, accessibility, and inter-system cooperation can improve knowledge proliferation and system development to mitigate and prevent political violence.

1. Introduction

This article reviews and compares conflict early warning systems.[1] A conflict early warning system (CEWS) is a risk analysis apparatus that provides forecasts of political violence to increase public awareness and prevent or mitigate conflict. These systems can greatly benefit decision-makers and crisis teams that work to reduce conflict-induced human suffering. The importance of these efforts can hardly be overstated. According to the Uppsala Conflict Data Program (UCDP), nearly 1,000,000 people were killed due to political violence between 2010 and 2020, and many more have been injured or forced to flee their homes.[2]

Many new forecasting systems have been created in recent decades, and international organizations and governments increasingly use CEWS in their internal decision-making processes. However, no systematic overviews exist. This article provides a comprehensive overview of existing CEWS, aiming to showcase their diversity and compare them on three important dimensions: (1) transparency and accessibility, (2) key parameters related to data and methods, and (3) forecasts.

Our summary highlights similarities and differences between systems and presents potential issues. In particular, we argue that increased transparency and cooperation will enable the production of more accurate and timely conflict forecasts that can guide policy and save lives. Further, our comparison of forecasts for African countries reveals agreement on high-risk countries. Our analysis also identifies a lack of dynamism in temporally fine-grained forecasts, indicating that more work is needed. The insights from our work can improve the development of future conflict forecasting systems and also be of interest to social scientists from other fields who aim to develop early warning systems.

[1] In the name of transparency, we note that the authors currently are associated with the Violence Early-Warning System (ViEWS), one of the systems under review. However, a commitment to a neutral assessment has been integral to the project's purpose of embracing methodological diversity, championing inter-system transparency, and collectively pushing the field forward. We make the code and data used in the article available for others to review independently.

[2] Source: https://ucdp.uu.se/exploratory.

Table 1
Overview of conflict early warning systems.

Conflict early warning system	Citation	Type
Atrocity Forecasting Project (AFP)	Goldsmith and Butcher (2018)	Academic
Conflict Forecast (CF)	Mueller and Rauh (2021)	Academic
Early Warnings Project (EWP)	EWP (2021)	Academic
Global Conflict Risk Index (GCRI)	Halkia, Ferri, Schellens, Papazoglou, and Thomakos (2020)	Operational
Peoples Under Threat (PUT)	PUT (2021)	Academic
Political Risk Services (PRS)	PRS (n.d.)	Commercial
PREVIEW	Manger, Mayer, and Fischer (2021)	Operational
ViEWS	Hegre et al. (2021, 2018)	Academic
Volatility Risk Index (VRI) (ACLED)	Raleigh, Hegre, Karlsen, and Linke (2010)	Academic
Water, Peace and Security (WPS)	Kuzma et al. (2020)	Academic

2. Review

In the following, we briefly review ten CEWS. The complete list is provided in Table 1.[3] We chose systems that forecast large-scale political violence, provide quantitative predictions,[4] and have a regular update schedule. Based on the inclusion criteria, we review a set of CEWS with diverse scope conditions, mandates, and outcomes. We believe that comprehensively showcasing the variety of CEWS and comparing them is helpful not only for the community of users (policymakers, peace researchers) and developers of CEWS but also for social scientists from other fields of expertise looking to generate early warning systems.

We categorize the systems according to their forecasting targets to structure the review. Four systems forecast political violence in a broad sense, three focus on mass killings, and two on armed conflict narrowly defined. The Political Risk Services (PRS) system is an exception. It provides industry-specific assessments (financial transfer, direct investment, and export market) and political risk forecasts. Correspondingly, the PRS's overarching goal is to make forecasts of political risks commercially relevant for firms and investors to make informed business decisions. The system is the only one in our review that is primarily commercially oriented. The other systems are either academic (produced by researchers employed in academic institutions or NGOs) or operational (produced by organizations that are also actively engaged in preventing conflict or mitigating its consequences).

The first group of CEWS, composed of ViEWS, PREVIEW, Conflict Forecast (CF), and The Volatility Risk Index (VRI), aim to forecast political violence broadly defined. ViEWS is a publicly available data-based system that provides early warnings for three forms of political violence: armed conflict between states and rebel groups, armed conflict between non-state actors, and violence against civilians. The project aims to integrate insights from the conflict research community into a theoretically and methodologically consistent CEWS.

The German Federal Foreign Office produces PREVIEW. The system gathers data and forecasts conflict violence with nearly global coverage and regional-specific forecasts depending on political priorities (e.g., in the Sahel Region). This is to guide German decision-makers and practitioners. In addition to quantitative modeling, PREVIEW provides regional qualitative assessments compiled by drawing on expertise at the federal foreign office's regional and country desks and involving stakeholders and experts across the government. The final integrated country assessments result in concrete policy recommendations for the German government (Manger et al., 2021).

CF is an academic project that develops and employs machine learning methods to predict conflict accurately. The project aims to improve forecasts of conflict outbreaks in previously peaceful countries. Better predictions of these cases are essential since existing CEWS do well at separating violent places from peaceful ones but typically perform poorly for new onsets (Mueller & Rauh, 2021).

VRI aims to produce accurate early warning signals and risk assessments in the short term and emphasizes a dynamic approach to forecasting. Instead of measuring the number of fatalities, the VRI studies volatility and the risk of violent surges relative to case-specific violence baselines. Like CF, the VRI has developed this dynamic approach to forecast the intensity and frequency of violence under challenging circumstances more effectively.[5]

The second group of CEWS, composed of the Atrocity Forecasting Project (AFP), Early Warning Project (EWP), and Peoples Under Threat (PUT), studies genocide, politicide, and mass killings. While the AFP and EWP are separate systems that utilize different methods and data, both projects share similar overarching goals. The two projects

---

[3] CoupCast from One Earth Future (https://oneearthfuture.org/activities/coup-cast) and Adverse Regime Transitions from the Varieties of Democracy Institute (Morgan, Beger, & Glynn, 2019) are related forecasting systems. However, since most coups and regime changes occur without organized violence, we did not include them in our overview. iCAST (https://www.lockheedmartin.com/en-us/capabilities/research-labs/advanced-technology-labs/icews.html) is a commercial forecasting system developed by Lockheed Martin's Integrated Crisis Early Warning System (ICEWS) program to forecast instability events. It is another influential system that falls outside the scope of our comparison because we could not find information on a regular update schedule for forecasts. We have not included forecasting projects based on qualitative expert opinion surveys (e.g., the Conflict Cartographer project, https://www.prio.org/projects/1900), which are promising complements to CEWS relying on the analysis of observational data.

[4] To facilitate comparison against what happened in actuality, forecasts need to have an element of quantitative scoring. If not, there is a risk that predictions are vague and consistent with any future outcome (Gleditsch, 2022).

[5] At the time of writing, The VRI is under development (beta version available).

strive to protect vulnerable communities by communicating the drivers of mass killings and using quantitative forecasting models to prevent or mitigate violence (Goldsmith & Butcher, 2018).

Like AFP and EWP, the PUT index aims to provide information, risk assessments, and early warnings of genocide and mass killings to protect vulnerable communities. The PUT index also forecasts the risk of systematic violent repression. To this end, PUT weighs and aggregates ten theoretically motivated indicators, and PUT differs from other CEWS by the absence of statistical predictive modeling.

Finally, the Global Conflict Risk Index (GCRI) and the Water, Peace, and Security (WPS) CEWS exclusively study armed conflict. However, the projects' perspectives and goals differ. The GCRI was developed by the European Union's Disaster Risk Management Knowledge Centre (DRMKC) to enhance the EU's conflict prevention capacities and understanding of conflict risk by generating robust assessments based on open-source quantitative evidence (Halkia et al., 2020). The WPS partnership was formed to mitigate and ameliorate security risks caused by water scarcity and aims to achieve this objective by pursuing two primary goals. First, to provide knowledge about the linkages between water, security, and conflict. Second, to utilize that knowledge to provide tools and services that can effectively help communities, societies, and countries address water scarcities, as well as their consequences, such as violence and conflict (Kuzma et al., 2020).

3. Comparison dimensions

We base our comparison of the CEWS on three groups of relevant features: (1) transparency and accessibility, (2) key parameters, and (3) forecasts.

Transparency and accessibility refer to system documentation, data, and code availability that enables researchers and decision-makers to evaluate and replicate the system and the forecasts or warnings themselves. Transparency is essential for early-warning systems for the same reasons it is for all scientific production. Today, most top-ranked international journals publishing peace and conflict research demand transparency and accessibility. In the context of a CEWS, transparency is vital for their credibility. For warnings to be taken seriously, decision-makers and observers need to understand how they were constructed. A CEWS is a complex chain of procedures from collecting and transforming input data through modeling and evaluation to producing the final forecasts. At all stages, there is always a risk of errors and data leakage, and difficult decisions and trade-offs must be made. With transparency, all these decisions can be evaluated, and the system's strengths and weaknesses will be apparent to everyone. Such transparency also facilitates interpretation. Even though black-box systems can be understood through interpretation tools that work outside the system (see Molnar, 2021, for a review), a detailed understanding of how the system works internally is beneficial.

Finally, just as in research, transparency is necessary to accumulate knowledge. Developing a CEWS is a complex endeavor that requires lots of resources, primarily from public funding. Transparent sharing of data collection procedures, the input data themselves, methods, procedures, and code, ensures that the conflict early-warning community can jointly develop the field as quickly and competently as possible.

We acknowledge that the challenges to openness are more significant for the CEWS than other academic efforts for two reasons. First, several of the producers of the CEWS reviewed in this article are also operational actors. A Ministry of Foreign Affairs or the UN World Food Program cannot publish assessments of the risk of armed conflict or genocides without risking diplomatic tension that threatens their primary operations. Transparency, in those cases, would have to be limited to more generic descriptions of methods, goals, and usage. Such CEWS, however, would also be more robust and more credible if they share anything that does not harm operations, such as summaries of the evaluation of predictive performance, for instance.

Second, academic providers of CEWS ace some challenges not common for standard academic output, such as research articles or books. In particular, CEWS are often complex systems requiring specialized knowledge or access to high-performance computing, and the level of replicability expected for other academic work may not be feasible. Moreover, many CEWS are developed using input data that are not publicly available and may be sensitive. While acknowledging some practical concerns, we maintain that any non-operational producers of CEWS should make data, methods, code, and predictions freely available. Such transparency is also beneficial to operational producers of CEWS, as highly transparent systems can function as valuable benchmarks for the less transparent systems' internal quality assurance procedures. Operational producers should share as much as possible without harming their primary operations.

The key parameters of a CEWS refer to the specifics of the data and methods used. What is the prediction target, and how is it operationalized? What is the unit of analysis and population of interest? What is the forecasting horizon? What estimation techniques and performance metrics are used? A comprehensive overview of these parameters is of great use to users and developers since it sheds light on the commonalities of existing CEWS, reflects perceptions of user demand, and identifies opportunities for data and methods cross-fertilization.

We do not rank choices on key parameters. Rather, we view the diversity of the CEWS implementations we review in this article as desirable and necessary. Resources and the provider's mandate largely determine key parameters of the CEWS. Important trade-offs need to be made in this regard. For example, whether it is wise to invest in global geographic scope, spatio-temporal granularity, and regular data updates depends on mandate and resources. The same applies to the choice of estimation technique, many of which require specialized skills, extensive coding and are often computationally expensive. Moreover, a broad set of parameter combinations is helpful for innovation. For example, smaller, more focused CEWS can draw on richer data and develop methods that can later be scaled up.

Table 2
Overview of public access to forecast, actuals and input data and code.

Conflict early warning system	Data			Code
	Forecast	Actuals	Input	
Atrocity Forecasting Project (AFP)	Yes[a]	Yes		
Conflict Forecast (CF)	Yes	Yes		
Early Warnings Project (EWP)	Yes	Yes	Yes	Yes
Global Conflict Risk Index (GCRI)[b]				
Peoples Under Threat (PUT)	Yes[a]		Yes[a]	
Political Risk Services[c] (PRS)				
PREVIEW[b]				
ViEWS	Yes	Yes	Yes	Yes
Volatility Risk Index (VRI) (ACLED)	Yes[a]			
Water, Peace and Security (WPS)	Yes[a]	Yes[a]	Yes[a]	

[a]Data not available in convenient, downloadable format.
[b]Operational system.
[c]Commercial system.

Finally, since CEWS ultimately are judged by the forecasting performance, we also compare forecasts for the CEWS that provide data access. We are primarily interested in country rankings but also compare overall system performance using a suite of performance metrics and visualization tools.

4. Comparison

4.1. Transparency and accessibility

Most systems provide thorough documentation on websites or in journal articles and technical notes. The level of transparency and accessibility of these documents is impressive overall. However, our survey of CEWS still reveals room for improvement. As shown in Table 2, most CEWS do not provide code or data on actuals (dependent variable) and input features (independent variables) to reproduce the forecasts. As discussed, code and data used in some systems may be private or sensitive and can not be made publicly available. Systems that use publicly available data or disseminate forecasts to the public, on the other hand, should strive to make code and data as easily accessible as possible. EWP and ViEWS are the only systems to provide complete access to all data and code.

Moreover, while many systems provide online tools to interact with the most recent forecasts (e.g., WPS, VRI, CF), downloadable access to the forecast data (predictions) is only provided by ViEWS, CF, and EWP. However, we did gain access to forecasts from two additional CEWS (AFP, PUT) by reaching out to project teams, which indicates a willingness to share. Access to code and data in downloadable format on project websites would be preferable. We purchased access to the commercial PRS system but were (as expected) not granted access to the operational systems GCRI and PREVIEW. VRI and WPS also did not provide data or forecasts that we could use to assess the systems.

We would also like to highlight that a CEWS needs to be transparent about internal data collection procedures or consider how the choice of externally collected input data sources affects system transparency (Eck, 2012; Weidmann & Rød, 2015). After all, forecasts ultimately rest on the reliability and validity of their underlying input data.

The political violence data provided by the EWP and the UCDP are available without any restrictions on scrutiny or comparison with other data sources.[6] ACLED has terms of use that are much more ambiguous in terms of the extent to which they permit comparison with other sources or criticizing their coding decisions. For example, prohibiting users from 'using ACLED's data or analysis in any manner that may harm, target, oppress, or defame ACLED'.[7] When transparency of core underlying data sources is incomplete or lacking, it is clear that the transparency of the CEWS using the data is also affected.

Our review shows that most CEWS provide detailed documentation on key parameters, but access to data and code for replication is lacking. Improved transparency and accessibility of data collection procedures, data, and code, at least for CEWS producers without operational constraints, would allow for external, rigorous assessment of existing CEWS and secure the accumulation of best practices and knowledge.

4.2. Key features

Table 3 summarizes key features of the systems we review, displaying outcomes, data, estimation, spatiotemporal resolution, country coverage, and forecasting periods. The table highlights several similarities and differences between existing early-warning systems.

Table 3 shows that most CEWS focus on one outcome/dependent variable, such as mass killings (e.g., EWP) or conflict incidence (e.g., GCRI). However, CF, PREVIEW, and ViEWS forecast multiple types of political violence. Moreover, while the prediction targets of the ten CEWS listed in Table 3 vary, the underlying data overlap considerably. Three systems rely on UCDP (Gleditsch, Wallensteen, Eriksson, Sollenberg, & Strand, 2002; Pettersson et al., 2021), and three use data from ACLED (Raleigh et al., 2010) to operationalize the outcomes. Four systems code their data, for example, AFP (The Targeted Mass Killing

---

[6] See https://earlywarningproject.ushmm.org/downloads, https://ucdp.uu.se/apidocs/, and https://ucdp.uu.se/downloads/ for details.

[7] See https://acleddata.com/terms-of-use/. In an earlier version of ACLED's terms of use, there was an explicit prohibition against 'benchmarking.' This clause has now been removed, but it is unclear how much scrutiny is possible before ACLED considers it 'defamation.'

Table 3
Key features of the conflict early warning systems.

	Dependent variable (DV)	DV threshold	DV Data	Estimation	Spatial	Temporal	Coverage	Forecast
Atrocity Forecasting Project	Genocide or politicide onset	25 civilians Year	TMK	Generalized additive model (GAM)	Country	Year	Global	3 years
Conflict Forecast	Violence onset Conflict onset	0 < BRD/m 4 < BRD/m per 1mil pop	UCDP	Primarily topic model and random forest	Country	Month	Global	12 months
Early Warnings Project	Mass killing onset	1000 civilians Year	Own coding	Logit model with elastic net regularization	Country	Year	Global	2 years
Global Conflict Risk Index	Conflict incidence	25 BRD Year	UCDP	Logistic regression	Country	Year	Global	4 years
Peoples Under Threat	Geno., mass kill, and repr. risk index	NA	Various (Index)	Aggregation and weighing of indicators	Country	Year	Global	1 year
Political Risk Services	Political/country risk index	NA	NA	Aggregation and weighing of indicators	Country	Month Year	Global	16 months 5 years
Preview	Armed conflict deaths Security related violent incidents Protests and riots	Counts	ACLED	Random Forest XGBoost Average Model Stacking	ADM1	Month Quarter Year	Global	2 years
ViEWS	State-based violence One-sided violence Non-state violence	25 and 1/ (log) counts Month	UCDP	Random Forest XGBoost Average EBMA	Country and prio-grid Planned: actor	Month	Africa	36 months
Volatility Risk Index (ACLED)	Violence risk (relative)	SD ≥ 2 events over baseline	ACLED	Calculation of increases in the frequency & intensity of violence relative to baseline levels	ADM1	Week	Global	Short-term NA
Water, Peace and Security	Conflict incidence	12 fatalities Year	ACLED	Random forest	ADM1	Month	Africa, Asia, and Middle East	12 months

Table 4
Evaluation metrics by conflict early warning system.

	Are metrics public?	AUROC	AUPR	Brier	Informal tools e.g., maps, tiles, or rankings
Atrocity Forecasting Project	Yes	Yes			Yes
Conflict Forecast	Yes	Yes	Yes		Yes
Early Warnings Project	Yes	Yes	Yes		Yes
Global Conflict Risk Index[a]	Yes	Yes		Yes	Yes
People Under Threat	Yes				Yes
Political Risk Services[b]					
PREVIEW[a]					Yes
ViEWS	Yes	Yes	Yes	Yes	Yes
Volatility Risk Index	Yes				Yes
Water, Peace and Security	Yes	Yes	Yes	Yes	Yes

[a]Operational system.
[b]Commercial system.

Dataset (TMK), Butcher, Goldsmith, Nanlohy, Sowmya, & Muchlinski, 2020).

Table 3 further reveals variation in estimation methods. Machine learning algorithms have made their way into the armed conflict forecasting literature and are used in five systems we surveyed. The techniques employed are logistic regressions with elastic net regularization, lasso or GAM, random forest, k-nearest neighbor, and neural networks. Two systems aggregate variables to an index without training weights to produce forecasts, and one uses standard logistic regression. Three systems also use ensemble methods to combine the results from multiple forecasting models (ViEWS, CF, and PREVIEW).

The most frequent spatial resolution is at the country level, but four projects are forecast at the subnational level (ViEWS, WPS, VRI, and PREVIEW). ViEWS forecasts at both the country and the subnational level. The temporal resolution is more diverse than the spatial, with four projects forecasting at the year level, four at the month level, one at both (PRS), one at the quarterly level (PREVIEW), and one at the week level (VRI).

The country coverage is global for most projects except for ViEWS (Africa and the Middle East) and WPS (Africa, Asia, and the Middle East). These projects are likely limited to specific regions because the spatial and temporal resolution restrict data availability. The forecasting horizon also varies. Some projects make forecasts for one year (PUT, CF, WPS, PREVIEW), but most cover multiple years (2–5).

Table 4 shows variations in the metrics used to evaluate CEWS forecasts. Most systems rely on many informal tools to discuss and show forecasts. The most common informal tools are prediction maps and rankings. The EWP further visualizes how much higher than average the predicted risk is in the years immediately preceding the start of mass killings and the proportion of mass killing onsets included in the list of 30 countries with the highest predicted risk. Most systems also report confusion matrices (true/false positives, true/false negatives). Further, five systems publish reports in which they qualitatively assess the forecasts. AFP releases reports every three years, EWP and PUT every year, WPS every quarter, and ViEWS every month.

Formal metrics are often reported. As shown in Table 4, six projects report AUROC, four AUPR, and three Brier scores. These metrics, especially AUROC and AUPR, are

standard in the conflict forecasting community. Whereas AUROC and AUPR reward forecasts that can order observations correctly, Brier reveals information about the sharpness of predictions (close to 0 or 1) and is also helpful in gauging calibration errors. Many projects also use the individual components of AUROC (sensitivity, specificity, precision, recall) to produce metrics and graphs that highlight aspects of the results not captured by AUROC and AUPR. PREVIEW forecasts change in violence and therefore relies on other metrics, namely the mean squared error (MSE), mean absolute error, and 'TADDA' (targeted absolute distance with direction augmentation, see Vesco et al., 2022). From 2022, ViEWS made available forecasts of the number of fatalities evaluated using MSE. Finally, GCRI shares more information than the other operational and commercial actors.

Our comparison of key features shows diversity in outcomes, data, estimation, spatio-temporal coverage, and evaluation. At the same time, there is convergence on critical parts of the systems. In particular, most CEWS use machine learning for estimation and evaluate results using similar tools and metrics.

4.3. Forecasts

In this section, we systematically compare forecasts from six CEWS, namely AFP, EWP, CF, PUT, PRS, and VIEWS.[8] Due to a lack of access to forecast data, GCRI, PREVIEW, VRI, and WPS are not included in the comparison. Others have also performed similar comparisons in the past; see Kuzma et al. (2020) and Halkia et al. (2020). However, the comparison we undertake here is more comprehensive, including more systems, metrics, and outcomes than previous efforts. We limit our focus to Africa to include all relevant CEWS in the comparison.

Fig. 1 displays each system's top 10 ranked African countries in 2020. In the maps, the color graded red represents the top 10 country rankings; the color grey represents the 11–25 risk ranking bracket; the color light grey represents the 26-n risk ranking bracket; missing risk rankings are represented by the white and black shade. Although the prediction targets and forecasting periods differ, the top-risk countries are similar across systems. Sudan, Nigeria, and Somalia are included in the

top 10 for all six CEWS, whereas South Sudan, Ethiopia, DRC, Libya, Mali, CAR, and Cameroon are also on most systems' top lists.

Several countries are in the top 10 for only 1–2 CEWS, for example, Egypt (AFP) and Niger (ViEWS). For some of these, such as Egypt, the ranking is relatively uniform across CEWS, 16 (CF), 18 (EWP), 14 (PUT), and 16 (ViEWS). However, some also vary widely in their ranking by the other CEWS. Niger, which is ranked ten by ViEWS, is ranked 30 (AFP), 12 (CF), 23 (EWP), and 16 (PUT). Nonetheless, the overall ranking patterns show many agreements across systems.

As discussed, the country rankings are overall similar across CEWS. To spot differences, we display bi-separation plots comparing ViEWS to each of the other systems in Fig. 2.[9] The figure shows predictions of large-scale political violence in 2020 for African countries. Bi-separation plots help single out individual observations that have been ranked differently by two models (Colaresi & Mahmood, 2017). Red dots indicate a country that experienced political violence in 2020, while blue dots display peaceful observations. The model on the *y*-axis (vertically) shows an improvement over the model on the *x*-axis (horizontally) when the red dots are located above the 45°line (to the left) and when the blue dots lie below the line (to the right). The opposite is the case for improvements of the model on the *x*-axis.

Figs. 2(a) and 2(b) compare ViEWS and AFP with each system's outcome variable. First, note that there is only one TMK onset in 2020 (Ethiopia in Fig. 2(b)). Further, Ethiopia is "on the line", indicating that the systems predicted a similar risk of TMK in the country in 2020. In Fig. 2(a), we see that AFP does better for countries such as South Sudan and Libya, while ViEWS does better for, among others, DRC and Cameroon. Overall, the distance in predicted probability between the systems is higher for the five highlighted cases that ViEWS predict better.

The comparisons between ViEWS and CF in Fig. 2(c) highlight the agreement between the two CEWS. Most of the countries with political violence lie close to the vertical line. Indeed, the largest discrepancies are seen in the highlighted blue dots. In Fig. 2(c), CF has much lower predicted probabilities for true negatives in South Africa and Ghana, whereas ViEWS does better for Botswana and Djibouti.

Fig. 3 shows radar plots comparing the systems' predictive performance using predictions and actuals for 2020. Each radar plot compares all CEWS on one performance metric (AUPR and AUROC) for each system's outcome. In Fig. 3(a), we can see that the predictions from ViEWS have the highest AUPR score of four out of five outcomes. CF performs better for the PRS outcome, and AFP performs much poorer than the other systems for this metric. In Fig. 3(b), we see that ViEWS outperforms the other systems for the AUROC metric. For the remaining CEWS, the picture is mixed. CF is second best for the PRS and CF outcomes but last for EWP; AFP is on par with ViEWS for the EWP outcome but last for CF.

---

[8] We use FPV as the outcome for ViEWS. FPV combines the three forms of organized violence covered by the UCDP – state-based, one-sided, and non-state conflict (Melander, Pettersson, & Themnér, 2016). The variable would take the value 1 if the UCDP recorded at least 25 fatalities in a country in either of these types of conflict. FPV is our preferred outcome since the different types of organized political violence defined by UCDP tend to occur in similar places; see "UCDP GED map: fatal events in 2020 by type of violence, world map" on https://ucdp.uu.se/downloads/charts/. For CF, we use the armed conflict outbreak outcome. PRS does not provide predicted probabilities. To work with the PRS data, we used the Political Risk Rating to generate predicted probabilities for the FPV outcome. Aggregating components, such as government stability, socioeconomic conditions, investment profile, internal conflict, external conflict, corruption, military in politics, religious tensions, law and order, ethnic tensions, democratic accountability, and bureaucracy quality, generate the Political Risk Rating. We chose it because it is a powerful predictor of the 18 monthly forecasts provided by PRS. In Appendix E, Table A.6, we show that the variable explains 72% of the variation in the 18-month forecasts.

[9] To reduce the complexity of the comparison and discussion, we chose ViEWS as the baseline in Fig. 2. We provide all replication data and code so readers can easily replicate the figure with a different baseline.
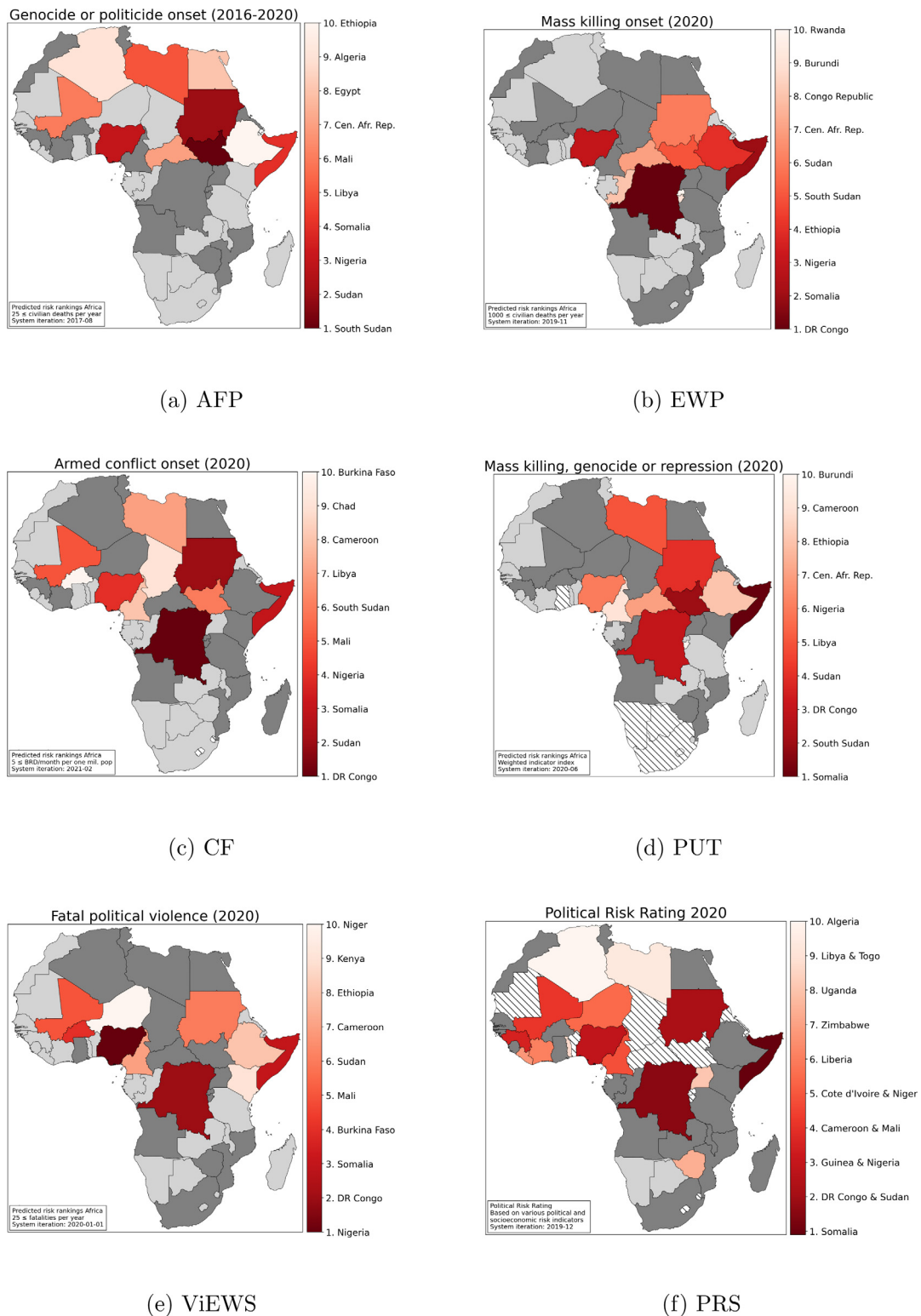
Fig. 1. Maps of country rankings for conflict early warning systems in 2020 (country-year predictions). The graded color red represents the top 10 country rankings; grey represents the 11–25 risk ranking bracket; light grey represents the 26-n risk ranking bracket; and missing risk rankings are represented by the white and black shades.
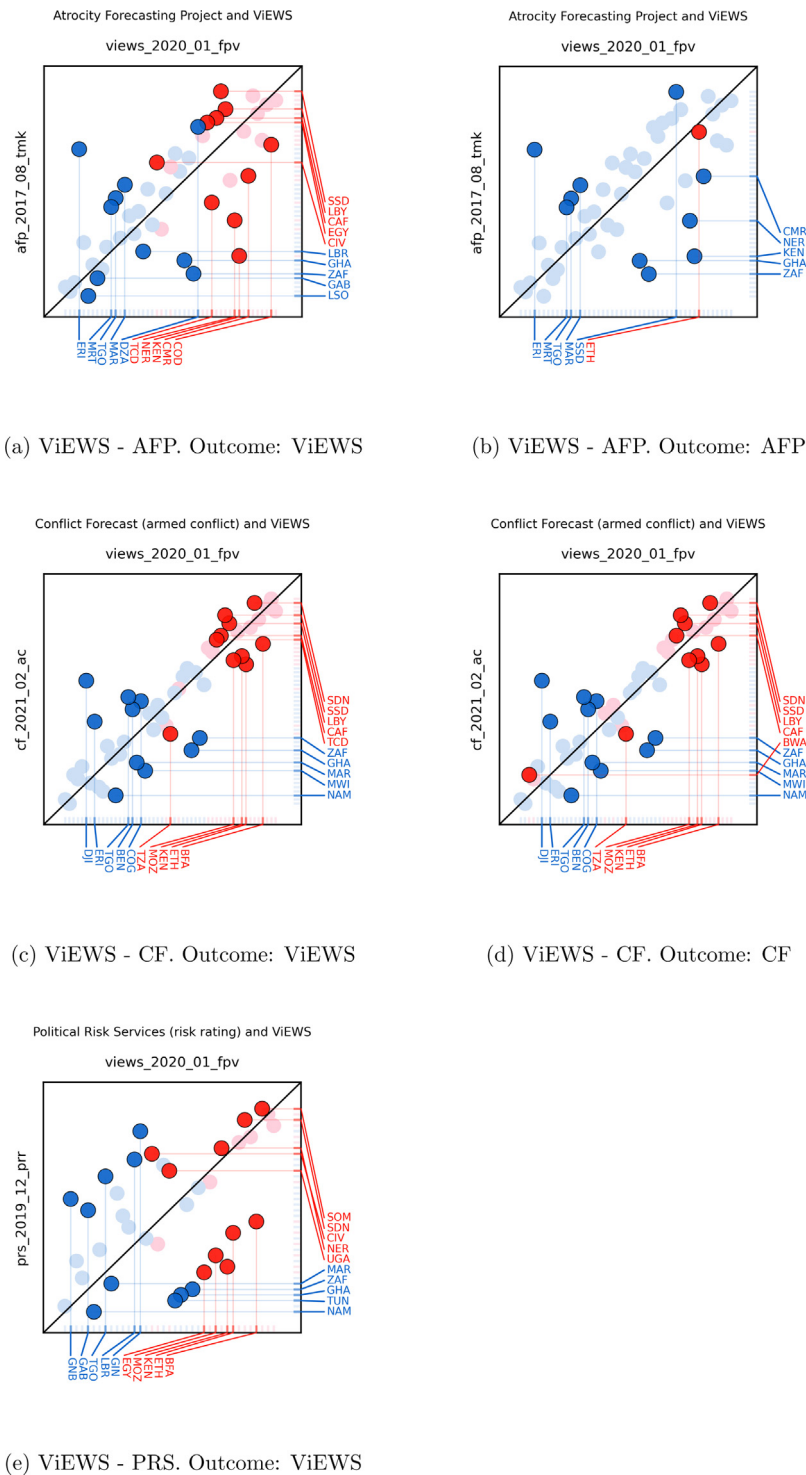
(a) ViEWS - AFP. Outcome: ViEWS

(b) ViEWS - AFP. Outcome: AFP

(c) ViEWS - CF. Outcome: ViEWS

(d) ViEWS - CF. Outcome: CF

(e) ViEWS - PRS. Outcome: ViEWS

Fig. 2. Biseparation plots for conflict early warning systems in 2020 (country-year predictions).

In Fig. 4, we zoom in on systems that provide predictions at the monthly level. We plot the predicted probability from ViEWS, CF, and PRS in Ethiopia, Nigeria, and Sudan for each month in 2020. We chose these three countries because they comprise static (Sudan) and volatile (Ethiopia, Nigeria) violence levels that vary from relatively low to very high. In addition, the black lines show the count of fatal political violence in the countries. The black lines show that the countries experienced political violence in 2020, and the intensity varies within and
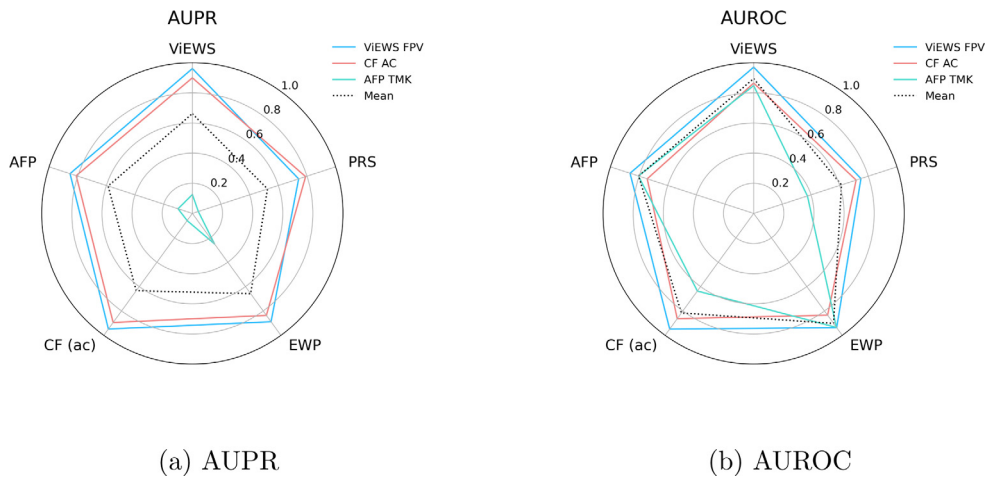
(a) AUPR   (b) AUROC

Fig. 3. Radar plots for conflict early warning systems in 2020 (country-year predictions).
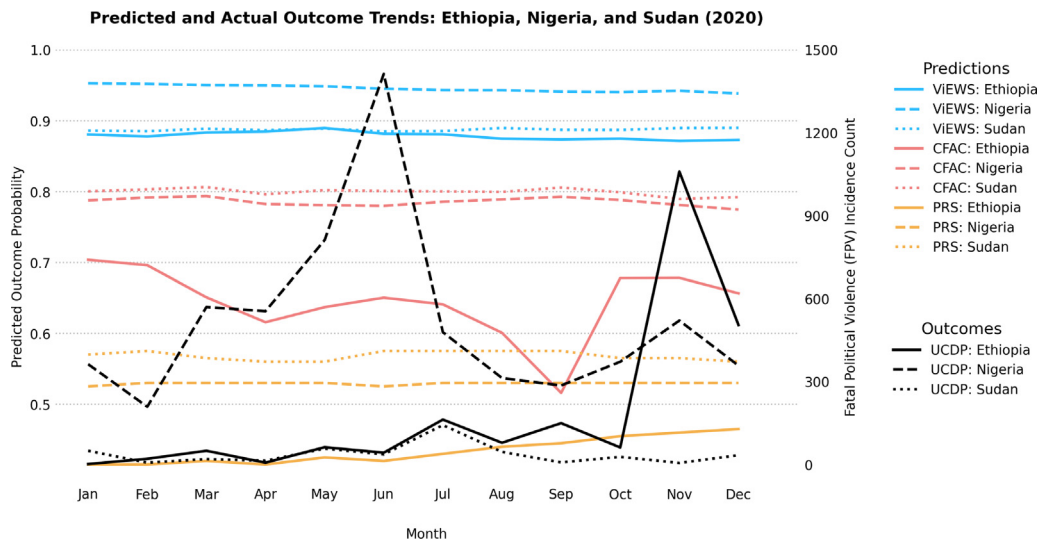

Fig. 4. Country-month predictions for Ethiopia, Nigeria, and Sudan in 2020 from ViEWS, CF, and PRS.

between them. Nigeria experienced more deadly violence than Ethiopia and Sudan, peaking in June. Violence in Ethiopia spiked in November, whereas Sudan sees consistent low-intensity violence throughout the year.

In Fig. 4, we can see that ViEWS assigns higher probabilities to violence occurrence than CF-AC and PRS in all three countries. Predictions for violence are approximately 0.95 for Nigeria and close to .9 for Ethiopia and Sudan. CF-AC is close to .8 for Nigeria and Sudan and between .5–.7 for Ethiopia. PRS predicted probabilities are below .6 for all countries. In these cases, ViEWS predicts the conflict occurrence better than the other systems, at least for Nigeria, where violence led to more than 25 fatalities every month. In contrast, the per-month predicted probability was too high for For Sudan and Ethiopia. At the same time, we see that the predicted probabilities are stable for all three countries. For a country like Nigeria, such a result is satisfactory: violence is ever-present, and

we anticipate a high likelihood of occurrence throughout. However, for Ethiopia and Sudan, we might expect more variation in predicted probability over time, but only the PRS adjusts its assessment, and only for Ethiopia. There is no doubt that the risk of violence is high. However, CEWS that aim to predict in fine-grained time windows, such as weeks, months, or quarters, should pick up situations in which the risk of violence is notably higher, such as in Ethiopia in November 2020, when massive fighting broke out in response to a contested regional election. Fig. 4 shows that CF-AC picks up on the dynamic and adjusts the probability of conflict from September to October. There is also a steady rise in predicted probabilities in Ethiopia by PRS. Apart from the varying probability in Ethiopia, most predictions are static over time. For CEWS with an ambition to predict short-term changes in conflict risk, allowing predicted probabilities to be more dynamic over time thus seems to be an area for development.

5. Conclusion

The article has comprehensively reviewed and compared conflict early warning systems. Three lessons can be drawn from the article. First, there is a need for improved transparency and accessibility of data and code. Most systems aim to prevent political violence by proliferating knowledge and providing accurate forecasts based on quality public data. In line with this objective, developing standards that promote transparency, accessibility, and inter-system cooperation can have considerable synergistic and reciprocal effects on collective knowledge proliferation and system development. Ultimately, more openness can ameliorate political violence.

Second, there is considerable variation in key parameters across systems. They can learn from each other's successful implementations, such as machine learning, ensemble methods, and evaluation using new metrics and tools. Strengthening collaboration through regular workshops and shared digital platforms could facilitate such learning. Finally, our comparison revealed significant overlaps in the countries with the forecasted highest risk. More research and improved cooperation are needed to understand the similarities and differences in forecasts across systems.

Given how different the CEWS we have reviewed are in terms of their aims and the outcomes they seek to forecast, this article has refrained from drawing firm conclusions regarding their relative performance. As noted in the ViEWS prediction competition, however, (Hegre, Vesco, & Colaresi, 2022; Vesco et al., 2022), it is clear that even the best models available struggle to predict *changes* in levels of political violence, and the emergence of new political violence in places that have historically been peaceful, in particular. To some extent, this is due to the intrinsic difficulty of such predictions. Many observers, for instance, note that Russia's decision to initiate a large-scale, formal attack on Ukraine in February 2022 was due to a miscalculation of both Russia's capabilities and Ukraine's willingness and ability to resist. Such miscalculation is almost impossible to predict – in other words, 'war is in the error term' (Fearon, 1995; Gartzke, 1999). However, it is also clear that the systems reviewed here are significant advances compared to what was available ten years ago. Furthermore, the efforts going into them mean progress is likely to continue. We hope that the review and comparison presented here will contribute to the further development of such systems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Qualitative reports

In Table A.1, we display the frequency of qualitative reports published by the conflict early-warning systems we review in the main text. We refer to the content in this table when discussing key features in the main text.

Appendix B. Forecasting history and publications by project

In Table A.2, we provide an overview of the previously published forecast for each of the conflict early-warning systems we review in the main text. We refer to the content in this table when discussing key features in the main text.

Appendix C. CEWS evaluation metrics by outcome

In Table A.3, we display the AUROC, AUPR, and Brier score for each conflict early warning system and outcome combination at the country-year level. The information in this table is used to generate Figs. 2 and 3. Table A.4 displays the AUROC, AUPR, and Brier score at the country-month level.

Appendix D. Country ranking

In Table A.5, we display the top 10 ranked countries for each conflict early warning system. The information in the table is used to generate Fig. 1 in the main text.

Appendix E. PRS prediction

Table A.6 shows how effectively PRS Political Risk rating explains the variation in PRS 18-month forecast. The result is used to test the relevance of using the Political Risk Rating as a surrogate for predicted probabilities in the quantitative forecast comparison.

Appendix F. Summaries of conflict early warning systems

F.1. Atrocity Forecasting Project (AFP)

F.1.1. Articles and resources
- https://politicsir.cass.anu.edu.au/research/projects/atrocity-forecasting
- AFP 2021 Report: Latest forecast (including model, methods, and data sources)
- Butcher et al. (2020): Dataset introduction paper
- Goldsmith and Butcher (2018): Evaluating past accuracy and discussing new forecasts

F.1.2. Accessibility and transparency
The Atrocity Forecasting Project is transparent regarding the data, model type, and data utilized to conduct their forecasts. However, while their Targeted Mass Killings (TMK) dataset is available for download, and their country risk rankings are available in PDF form, the complete datasets utilized for modeling and the code used to produce them are not publicly available.

F.1.3. Goals
In sum, the Atrocity forecasting project's proclaimed overarching goal is to enhance the capacity for forecasting

Table A.1
Publication of qualitative reports by CEWS.

Conflict early warning system	Reports
Atrocity Forecasting Project (AFP)	≈Every third year
Conflict Forecast (CF)	
Early Warnings Project (EWP)	Yearly
Global Conflict Risk Index (GCRI)	
Peoples Under Threat (PUT)	Yearly
Political Risk Services (PRS)	
PREVIEW	
ViEWS	Monthly
Volatility Risk Index (VRI) (ACLED)	
Water, Peace and Security (WPS)	Quarterly

Table A.2
Overview of CEWS forecast history and publication rate.

Conflict early warning system	Past Forecasts
Atrocity Forecasting Project (AFP)	≈Every third year 2021–2023 (4/2021) 2016–2020 (8/2017) 2011–2015 (8/2012)
Conflict Forecast (CF)	Monthly (t + 12 m) since 12/2020
Early Warnings Project (EWP)	Yearly 2020–2021 (12/2020) 2019–2020 (11/2019) 2018–2019 (10/2018) 2014 −2016 (same year)
Global Conflict Risk Index (GCRI)	
Peoples Under Threat (PUT)	Yearly (same year) since 2011
Political Risk Services (PRS)	
PREVIEW	
ViEWS	Monthly (t + 36) since 6/2018
Volatility Risk Index (VRI)	Short-term (NA)
Water, Peace and Security (WPS)	Quarterly (t + 12 Months) since 12/2019

Table A.3
CEWS evaluation metrics based on different outcomes (Africa 2020, country-year level).

CEWS	Metric	AFP TMK	CF AC	ViEWS FPV	Mean
AFP	AUROC	0.804348	0.743083	0.862963	0.803465
AFP	AUPR	0.100000	0.810533	0.853857	0.588130
AFP	Brier	0.025422	0.460739	0.377271	0.287810
CF (ac)	AUROC	0.636364	0.863354	0.948387	0.816035
CF (ac)	AUPR	0.058824	0.894591	0.948230	0.633881
CF (ac)	Brier	0.226351	0.133699	0.080544	0.146865
EWP	AUROC	0.934783	0.834448	0.937097	0.902109
EWP	AUPR	0.250000	0.836114	0.888237	0.658117
EWP	Brier	0.020232	0.451272	0.375339	0.282281
PRS	AUROC	0.375000	0.714912	0.748512	0.612808
PRS	AUPR	0.041667	0.790189	0.741828	0.524561
PRS	Brier	0.206504	0.230558	0.214612	0.217225
ViEWS	AUROC	0.847826	0.864907	0.970588	0.894440
ViEWS	AUPR	0.125000	0.899600	0.961075	0.661892
ViEWS	Brier	0.374700	0.152574	0.119697	0.215657

genocide and mass atrocities globally and in the Asia-Pacific region. The project pursues four primary aims:

1. Developing cutting-edge quantitative forecasting models

2. Improving the general understanding of the drivers behind conflict and political instability (which greatly increase the risk of genocide and mass atrocities)

Table A.4
CEWS evaluation metrics based on different outcomes (Africa 2020, country-month level).

CEWS	Metric	CF AC	ViEWS FPV	Mean
CF (ac)	AUROC	0.937326	0.942768	0.940047
CF (ac)	AUPR	0.837686	0.801800	0.819743
CF (ac)	Brier	0.085349	0.091358	0.088354
PRS	AUROC	0.839411	0.825180	0.832295
PRS	AUPR	0.734155	0.709773	0.721964
PRS	Brier	0.199082	0.198970	0.199026
ViEWS	AUROC	0.941508	0.969880	0.955694
ViEWS	AUPR	0.872913	0.899545	0.886229
ViEWS	Brier	0.174710	0.180814	0.177762

Table A.5
Predicted risk rankings (Africa 2020).

CEWS:	AFP	CF	EWP	PUT	PRS	ViEWS
Outcome	Genocide or Politicide onset	Conflict onset	Mass killing onset	Repression or mass killing incidence	Political risk rating	Fatal political violence incidence
Forecast (Published)	2016–2020 (8/2017)	2020 (2/2021)	2020 (11/2019)	2020 (6/2020)	2020 (12/2019)	2020 (1/2020)
1.	South Sudan	DR Congo	DR Congo	Somalia	Somalia	Nigeria
2.	Sudan	Sudan	Somalia	South Sudan	DR Congo & Sudan	DR Congo
3.	Nigeria	Somalia	Nigeria	DR Congo	Guinea & Nigeria	Somalia
4.	Somalia	Nigeria	Ethiopia	Sudan	Cameroon & Mali	Burkina Faso
5.	Libya	Mali	South Sudan	Libya	Cote d'Ivoire & Niger	Mali
6.	Mali	South Sudan	Sudan	Nigeria	Liberia	Sudan
7.	Cen. Afr. Rep.	Libya	Cen. Afr. Rep.	Cen. Afr. Rep.	Zimbabwe	Cameroon
8.	Egypt	Cameroon	Congo Republic	Ethiopia	Uganda	Ethiopia
9.	Algeria	Chad	Burundi	Cameroon	Libya & Togo	Kenya
10.	Ethiopia	Burkina Faso	Rwanda	Burundi	Algeria	Niger

Table A.6
PRS political risk rating and forecast correlation (2021).

	Dependent variable: 18 month forecast
	(1)
Political risk rating	−0.071***
	(0.004)
const	5.955***
	(0.295)
Observations	99
R^2	0.730
Adjusted R^2	0.727
Residual Std. Error	0.519 (df = 97)
F Statistic	262.413*** (df = 1.0; 97.0)

Note: *p<0.1; **p<0.05; ***p<0.01.

3. Improving the understanding of the causal chain between instability and genocide or mass atrocities.
4. Provide forecasts and reports that can serve as early warning tools to mitigate destructive outcomes and protect vulnerable populations.

F.1.4. Methods

The project utilizes forecasting techniques based on predictive modeling and machine learning that are intended to be used in combination with other qualitative and quantitative research methods (Butcher et al., 2020). The project's newest forecast utilizes a Generalized additive model (GAM) with a logit link, trained on data from 1946 to 2017 and applied to data up to the end of 2020 for the prediction.

AFP forecasting equation using a Generalized Additive Model (GAM):

$$GAM_{logit}(Genocide/Politicide\ onset_{t+1\ through\ t+3})$$
$$= \beta_0 + \beta_1(Instability) + f_1(Regime)$$
$$+ f_2(RegimeChange) + \beta_2(Election) + \beta_3(Election_{t+1})$$
$$+ \beta_4(Election_{t+2}) + \beta_5(Election_{t+3}) + \beta_6(Election_{t+4})$$
$$+ \beta_7(Human\ Rights) + f_3(Population)$$
$$+ f_4(InfantMortality) + \beta_8(MENA) + \beta_9(CS\_Asia)$$
$$+ \beta_{10}(RegionConflict) + \beta_{11}(EthnicFractionalization)$$
$$+ f_5(EthnicFractionalization) + \beta_{12}(GuerrillaTactics)$$
$$+ \beta_{13}(Assassinations) + \beta_{14}(InternationalizedCivilWar)$$
$$+ \beta_{15}(InterstateWar) + \beta_{16}(Targeted\ Mass\ Killing)$$

$$+ \ \beta_{17}(Ongoing\ Genocide/Politicide)$$
$$+ \ f_6(noGPyears) + f_7(Time)$$

F.1.5. Outcome of interest

The outcome of interest is *Genocide/Politicide Onset*, measured using the Targeted Mass Killing (TMK) dataset released by the AFP in 2020 (Butcher et al., 2020). The AFP considers a case as a targeted mass killing when an organized armed actor perpetrates the killings, and the following operational criteria are met:

1. 25 or more civilians are killed in a year
2. The actor deliberately targeted the victims
3. The victims were disproportionately associated with one or more ethnic, political, or religious group(s)
4. The group was specifically targeted to affect its political activity, reduce its numbers or expel its members

F.2. Conflict Forecast (CF)

F.2.1. Articles and resources
- https://conflictforecast.org/: Dashboard and Project Information
- Mueller and Rauh (2021): Comprehensive Project Methodology Article

F.2.2. Accessibility and transparency

Considerable transparency and access. Including an online dashboard and downloadable data for each iteration of the forecast. However, the input features used in the models and the code are not available to the public.

F.2.3. Goals

The overarching goal of the Conflict Forecast (CF) project is to provide data and machine learning methods that can accurately predict conflict to support various policy areas and decision-makers. The project particularly emphasizes the goal of providing better predictions for "*hard problems*" – that is, the outbreak of violence and conflict in previously peaceful countries, which models centered around past violent outcomes often fail to predict (Mueller & Rauh, 2021, p.1,3).

F.2.4. Methods

The Conflict Forecast utilizes a two-step machine learning process. First, a dynamic top model uses unsupervised learning for feature extraction to analyze text data composed of over 4 million documents collected from two news aggregators (LatinNews and BBC monitor) and three newspapers (The Economist, New York Times, Washington Post) on 30 topics. The share of topics for all countries is subsequently calculated for each month between 1989m1 and T. Finally, the shares are combined with a set of dummies (h_{it}) capturing post-conflict risk in a random forest model (mathematical definition on the next page) to forecast conflict outside of the sample. Where $\theta_i t$ is a vector summary of the topic shares and the log of the word counts; y_{it+1} is the onset of conflict in one month (Mueller & Rauh, 2021, p. 13–14).

To generate the non-linear model ($F_T(.)$), the CF tested predictions from adaptive boosting, k-nearest neighbor, neural networks, random forests, logit lasso regression combined with ensembles of the various models (Mueller & Rauh, 2021, p.14).

CF Training Model	**CF Prediction Model**
$y_{it+1} = F_T(h_{it}, \theta_{it})$	$\hat{y}_{iT+1} = F_T(h_{iT}, \theta_{iT})$

F.2.5. Outcome of interest

The Conflict Forecast project studies two outcomes: the onset of violence and the onset of armed conflict. The outcomes are operationalized and measured using the UCDP Georeferenced Event Dataset (GED) collapsed to a country–month level. All three categories (AC, OS, NS) are aggregated into one value representing the absence or presence of any fatality related to conflict (Mueller & Rauh, 2021, p.8).

1. **Violence Onset:**
   "An outbreak of any violence: a country goes from no fatalities to more fatalities in a month".
2. **Armed Conflict Onset:**
   "An outbreak of armed conflict: a country goes from having less than five fatalities per one million inhabitants in a month to more".

F.3. Early Warning Project (EWP)

F.3.1. Articles and resources
- https://earlywarningproject.ushmm.org/
- https://github.com/EarlyWarningProject
- Latest mass killing risk report EWP (2021) includes methods and data

F.3.2. Accessibility and transparency

The early warning project is entirely transparent and makes all the data and modeling available to the public on their website and GitHub. The project also provides comprehensive methodology articles and an online dashboard representing the data and models.

F.3.3. Goals

In sum, the overarching goals of the project are twofold. First, prevent future genocides by increasing the understanding of the processes that drive the phenomenon. Second, to produce forecasts that provide early warning signs and reliable risk assessments that can incentivize action and mitigate destructive outcomes.

F.3.4. Methods

The projects train and test various statistical algorithms on historical data to locate variables and models that effectively predict the risk of mass killings. The latest forecast and report utilizes a "logistic regression model with "elastic-net" regularization" that identifies predictive relationships between 30 variables and data on mass killings between 1945 and 2015. The model is subsequently applied to data from 2018 to produce the 2019–2020 forecasting assessments, which generates an

estimated risk of mass killings onset in percentage for each country.

F.3.5. Outcome of interest

Mass killing episodes are measured on a binary scale using the following operational criteria:

1. The episode is the consequence of the deliberate actions of an armed group to coerce, compel, or destroy the targeted group.
2. The episode resulted in the killing of at least 1000 noncombat civilians – which are defined as people that are not members of any irregular or formal military organization – within a year

F.4. Global Conflict Risk Index (GCRI)

F.4.1. Articles and resources
- https://knowledge4policy.ec.europa.eu/dataset/ds00151_en
- Halkia et al. (2020): Project methodology

F.4.2. Accessibility and transparency
NA, operational system.

F.4.3. Goals

The Global Conflict Risk Index (GCRI) is produced by the European Union's Disaster Risk Management Knowledge Centre (DRMKC). It forecasts the risk of violent conflict on a country level during the forthcoming 1–4 years. The index was developed to enhance the EU's conflict prevention capacities and understanding of short- and long-term conflict risk by generating robust, open-source-based quantitative evidence.

F.4.4. Methods

The index utilizes 25 variables along six dimensions – social, economic, political, security, geographical/environmental, and demographic – to calculate the probability of subnational conflicts (SN) and national power conflicts (NP). The project utilizes standard logistic regression models to assess the probability of SN/NP conflict incidences. Furthermore, the project conduct ten-fold cross-validation to assess the performance of the models (Halkia et al., 2020).

Mathematical Model Definition:
$$P = \frac{e^y}{1+e^y}; y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

F.4.5. Outcome of interest

The GCRI uses the *Incidence of Conflict* as their dependent variable. Their operational definition of conflict incidence is built on the UCDPs – Battle-Related Deaths (BRD), One-Sided Violence (OSV), and Non-State Conflict (NSC) – datasets and definitions (25 battle-related deaths per calendar year). The GCRI merges the UCDP datasets and separates the cases into what they define as either Subnational (SN) conflicts or National Power (NP) conflicts.

F.5. Peoples Under Threat (PUT)

F.5.1. Articles and resources
- General project information

  - https://peoplesunderthreat.org/about-the-peoples-under-threat-index/
  - PUT (2021): Latest system iteration, data, and methodology

F.5.2. Accessibility and transparency

The project is entirely transparent in terms of modeling and data. Furthermore, the data is available online in a table and dashboard format. However, the dataset is not available for download in its entirety.

F.5.3. Goals

The Peoples Under Threat (PUT) project and the Peoples Under Threat Index aim to provide information, risk assessment, and early warnings that can prevent systematic violent repression, genocide, and mass killings. The project aim's to achieve this objective by providing an early warning system based on conflict indicators that can identify the communities and countries that face the most significant risk of the previously mentioned destructive outcomes.

F.5.4. Methods

The Peoples Under Threat Index is produced through the aggregation and weighing of 10 indicators. These indicators are: (a) self-determination conflicts, (b) major armed conflict, (c) prior genocide/politicide, (d) flight of refugees and IDPs, (e) refugees and IDPs legacy of vengeance - group grievance, (f) rise of factionalized elites, (g) voice and accountability, (h) political stability, (i) the rule of law, (j) OECD country risk classification. In the latest data release, the index spans between 1.70 (lowest risk) and 29.03 (highest risk). It ranks the 115 countries where communities face the greatest risk of genocide, mass killings, or systematic violent repression. The following formula is utilized to weigh the indicators and produce the index:

Peoples Under Threat Index Formula
$$\left(\frac{a}{2}\right) + (b \cdot 1.25) + (c \cdot 2) + (d \cdot 10) + \left(\frac{e+f}{6}\right) + \left(\frac{g+h+i}{-1}\right) + (j \cdot 0.625)$$

F.5.5. Outcome of interest

The Peoples Under Threat project does not provide a statistical forecasting prediction model. Instead, the project utilizes the weighted indicators provided above to provide yearly relative risk measures and assessments of genocide, mass killings, and systematic violent repression. Since the different violent outcomes are bundled, it is unclear to what extent the index represents the risk of each separate outcome.

F.6. Political Risk Services (PRS)

F.6.1. Articles and resources
- https://www.prsgroup.com/

- https://epub.prsgroup.com/index.php/country-database/country-data
- PRS (n.d.) Methodology Paper

F.6.2. Accessibility and transparency

NA, commercial system.

F.6.3. Goals

The Political Risk Services proclaimed overarching goal is to provide their customers with timely, accurate, and completely unbiased forecasts of political risks to enable them to make informed business decisions and investments.

F.6.4. Methods

The Political Risk Services (PRS) chiefly provide customizable industry-specific (financial transfer, direct investment, and export market) forecasts. These forecasts are conducted on a country level in two stages – 18 months and five years – and are produced by analyzing the probability of future regime scenarios and how they might impact the level of political turmoil, government intervention, and ultimately the business climate in each specific country. The regime scenario scores are subsequently calculated and converted into letter grades between A+ and D- for the three industry-specific investment areas.

However, the PRS also produces The Political Risk Services Risk Index (PRSRI), which provides a general risk summary value based on the unique risk components of each generated forecast. The risk index is based on 17 variables/risk components which vary based on the time horizon (18 months or five years) and will be covered in-depth in the next section.

F.6.5. Outcome of interest

The PRSRI, as previously argued, assesses risk based on different risk components depending on the type of time horizon. Each component is given a risk value between 0 (lowest risk) and 4 (highest risk) and is then calculated and converted into the PRSRI with the following formula:

PRSRI Convertion Formula

$$\frac{RS-68}{-68} \cdot 100$$

18 Month Risk Components	5 Year Risk Components
1. Turmoil	1. Turmoil
2. Restrictions on equity	2. Investment restrictions
3. Restrictions on local operations	3. Restrictions of foreign trade
4. Taxation discrimination	4. Domestic economic problems
5. Repatriation restrictions	5. International economic problems
6. Exchange controls	
7. Tariff barriers	
8. Nontariff barriers	
9. Payment delays	

10. Expansionary economic policies
11. Labor Costs
12. Foreign debt

F.7. Preview

F.7.1. Articles and resources

- NA, operational system

F.7.2. Accessibility and transparency

NA, operational system.

F.7.3. Goals

PREVIEW is a capability produced by the German Federal Foreign Office to gather quality data and to produce early warnings and risk predictions. The system's primary aim is to guide German government practitioners and policymakers to make better and more informed decisions. Furthermore, PREVIEW's strategy to achieve this overarching goal is to provide regional and global prediction models and in-depth qualitative analysis to identify high-risk regions and countries at a precise spatial resolution (Manger et al., 2021, p. 1–2).

F.7.4. Methods

The regional prediction model is based on the PREVIEWS Quartile Conflict Model (PREVIEW QM), which identifies countries and regions at high risk over a short- to mid-term time horizon of up to 24 months. The predictive analytics and forecasts are combined with expert assessments and qualitative information to develop the final policy recommendations (Manger et al., 2021, p.1).

F.7.5. Outcome of iiterest

PREVIEW forecasts three outcomes based on data and operational criteria provided by ACLED. The three-count variables used by the project measure the number of incidences occurring that are related to one of three separate risk themes – FAT (fatalities), SRZ (security), and PRO (protests) – by measuring the logarithmic change (basis. exp) relative to the basic period (T0). First, the outcome variable FAT counts the number of conflict-related fatalities. Second, the outcome variable SRZ counts the number of security-related incidents – e.g., remote violence, one-sided violence, or battles. Third, the outcome variable PRO counts the number of protests and riots (Manger et al., 2021, p. 2–3).

F.8. Views: a political violence early-warning system

F.8.1. Articles and resources

- https://www.pcr.uu.se/research/views/: Project Website; including data, code, monthly qualitative reports etc.
- https://ucdp.uu.se/: Data
- Hegre et al. (2019): Article presenting ViEWS
- Hegre et al. (2021): Article revising and evaluating ViEWS

F.8.2. Accessibility and transparency

The majority of code and data are downloadable and available to the public. However, running the system requires installing the ViEWS system. Moreover, access to a supercomputer is necessary for the most computationally intense forecasts.

F.8.3. Goals

(From the ViEWS website:) The project will build a pilot for a worldwide system with uniform coverage and frequent updates to avoid blind spots, provide location- and actor-specific alerts, and, most importantly, be transparent, replicable, and publicly available, including public assessments of predictive performance.

F.8.4. Methods

Dynamic simulation (logistic regression), random forest, Ensemble Bayesian Model Averaging (Hegre et al., 2019, 2021).

F.8.5. Outcome of interest

ViEWS predicts state-based armed conflict, non-state conflict, and one-sided violence as defined by UCDP (Pettersson et al., 2021; Sundberg & Melander, 2013).

F.9. Volatility and Risk Index (VRI)

F.9.1. Articles and resources

- https://acleddata.com/early-warning-research-hub/volatility-and-risk-predictability-index/

F.9.2. Accessibility and transparency

A dashboard and limited methodology information are available online. The forecasting tool is currently in development, and in-depth method descriptions, code, and datasets are not available to the public.

F.9.3. Goals

The Volatility and Risk Index (VRI) primary goal is to produce accurate and practical information to monitor conflict environments by providing precise early warning signals and risk assessments. Specifically, the index emphasized deviations from the baseline levels of violence in each administrative division to achieve a relative and dynamic understanding of the potential intensity and frequency of surges in violence.

F.9.4. Methods and outcome of interest

The VRI predicts the volatility of political violence across administrative divisions – e.g., governorate, province, state – relative to a baseline level of violence to generate a dynamic risk level and assess the likelihood of future surges. Specifically, the system dynamically tracks positive deviations (increases) from the case-specific baselines of violent events. The project categorizes and operationalizes the baseline, volatility, and risk according to the matrix below.

Volatility/Baseline Matrix		
Volatility Baseline	Low Baseline (<1.5 events/week)	High Baseline (>1.5 events/week)
Low Volatility: Violence has spiked 2 or more standard deviations above the baseline **fewer** than 6 weeks during the past year.	Low Risk	Consistent Risk
High Volatility: Violence has spiked 2 or more standard deviations above the baseline **more** than 6 weeks during the past year.	Growing Risk	Extreme Risk

The baselines are calculated based on the average number of weekly violent events in the last three years, where 1.5 events and above is considered high. The volatility is calculated based on how often the violent events surpass two standard deviations above the specific baseline in an administrative district. An administrative district is, in turn, classified as highly volatile if six or more weeks within a year are characterized by such volatility.

F.10. Water, Peace, and Security (WPS)

F.10.1. Articles and resources

- https://waterpeacesecurity.org/info/about-wps
- https://waterpeacesecurity.org/info/methodology
- Kuzma et al. (2020): Methodology and presentation of forecasting system

F.10.2. Accessibility and transparency

The project is completely transparent in terms of its methodology, comprehensively presented in its latest methodology article (Kuzma et al., 2020). However, while most of the data is available in the project's online dashboard, the dataset and code are not available to the public.

F.10.3. Goals

The Water Peace and Security partnership was formed to mitigate and ameliorate security risks generated by water scarcity. The partnership's main objective is twofold. First, to provide knowledge about the linkages between water, security, and conflict. Second, to utilize that knowledge to provide tools and services that can help communities, societies, and countries to address water scarcities and their direct and indirect consequences effectively.

F.10.4. Methods

As previously stated, the WPS partnership aims to mitigate the consequences of water scarcity by providing innovative tools and services. The primary service is the partnership's conflict forecasting system that produces early conflict warnings by utilizing supervised learning techniques to assess the relationships between water scarcity and conflict. The project uses a random forest model, trained on data gathered between January 2004 and May 2016 and tested on data from January to June 2018.

F.10.5. Outcome of interest

The WPS project forecasts *conflict* using a binary scale which defines conflict as "Organized violence resulting in at least ten fatalities over 12 months". The measurement is based on the Armed Conflict Location and Event Database (ACLED) data and operational criteria (Kuzma et al., 2020).

References

Butcher, Charles, Goldsmith, Benjamin, Nanlohy, Sascha, Sowmya, Arcot, & Muchlinski, David (2020). Introducing the targeted mass killing data set for the study and forecasting of mass atrocities. *The Journal of Conflict Resolution*, *64*(7–8), 1524–1547.

Colaresi, Michael, & Mahmood, Zuhaib (2017). Do the robot: Lessons from machine learning to improve conflict forecasting. *Journal of Peace Research*, *54*(2), 193–214.

Eck, Kristine (2012). In data we trust? A comparison of UCDP GED and ACLED conflict events datasets. *Cooperation and Conflict*, *47*(1), 124–141.

EWP (2021). *Countries at risk for mass killing 2021–22: Early warning project statistical risk assessment results*. Simon-Skjodt Center for the Prevention of Genocide.

Fearon, James D. (1995). Rationalist explanations for war. *International Organization*, *49*(3), 379.

Gartzke, Erik (1999). War is in the error term. *International Organization*, *53*(3), 567–587.

Gleditsch, Kristian Skrede (2022). One without the other? Prediction and policy in international studies. *International Studies Quarterly*, *66*(3), sqac036.

Gleditsch, Nils Petter, Wallensteen, Peter, Eriksson, Mikael, Sollenberg, Margareta, & Strand, Håvard (2002). Armed conflict 1946–2001: A new dataset. *Journal of Peace Research*, *39*(5), 615–637.

Goldsmith, Benjamin, & Butcher, Charles (2018). Genocide forecasting: Past accuracy and new forecasts to 2020. *Journal of Genocide Research*, *20*(1), 90–107.

Halkia, Matina, Ferri, Stefano, Schellens, Marie, Papazoglou, Michail, & Thomakos, Dimitrios (2020). The global conflict risk index: A quantitative tool for policy support on conflict prevention. *Progress in Disaster Science*, *6*, Article 100069.

Hegre, Håvard, Allansson, Marie, Basedau, Matthias, Colaresi, Mike, Croicu, Mihai, Fjelde, Hanne, et al. (2019). ViEWS: A political violence early warning system. *Journal of Peace Research*, *56*(2), 155–174.

Hegre, Håvard, Bell, Curtis, Colaresi, Michael, Croicu, Mihai, Hoyles, Frederick, Jansen, Remco, et al. (2021). ViEWS$_{2020}$: Revising and evaluating the ViEWS political violence early-warning system. *Journal of Peace Research*, *58*.

Hegre, Håvard, Colaresi, Mike, Croicu, Mihai, Hoyles, Frederick, Olafsdottir, Gudlaug, Petrova, Kristina, et al. (2018). ViEWS: A political violence early warning system. Typescript, Uppsala University. www.uu.se/pcr/research/views.

Hegre, Håvard, Vesco, Paola, & Colaresi, Michael (2022). Lessons from an escalation prediction competition. *International Interactions*, *48*(4), 521–554.

Kuzma, Samantha, Kerins, Peter, Saccoccia, Liz, Whiteside, Cayla, Roos, Hannes, & Iceland, Charles (2020). *Leveraging water data in a machine learning–based model for forecasting violent conflict*: *Tech. rep.*, Water Peace and Security Partnership.

Manger, Nicole, Mayer, Thomas, & Fischer, Benjamin (2021). *Conflict trends at subnational level for the Central Sahel Region in the context of climate-change induced weather events*: *Tech. rep.*, German Federal Foreign Office, PREVIEW.

Melander, Erik, Pettersson, Thérése, & Themnér, Lotta (2016). Organized violence, 1989–2015. *Journal of Peace Research*, *53*(5), 727–742.

Molnar, Christoph (2021). *Interpretable machine learning. A guide for making black box models explainable*.

Morgan, Richard K., Beger, Andreas, & Glynn, Adam (2019). *Varieties of forecasts. Predicting adverse regime transitions*. V-Dem Institute.

Mueller, Hannes, & Rauh, Christopher (2021). *The hard problem of prediction for conflict prevention*: *Cambridge working papers in economics*, (p. 67). INET WORKING PAPERS.

Pettersson, Thérése, Davies, Shawn, Deniz, Amber, Engström, Garoun, Hawach, Nanar, Högbladh, Stina, et al. (2021). Organized violence 1989–2020, with a special emphasis on Syria. *Journal of Peace Research*, *58*(4), 809–825.

PRS (n.d.). PRS methodology. The Political Risk Services Group.

PUT (2021). *People under threat 2020*. Minority Rights Group International and Ceasefire Centre for Civilian Rights.

Raleigh, Clionadh, Hegre, Håvard, Karlsen, Joakim, & Linke, Andrew (2010). Introducing ACLED: An armed conflict location and event dataset. *Journal of Peace Research*, *47*(5), 651–660.

Sundberg, Ralph, & Melander, Erik (2013). Introducing the UCDP georeferenced event dataset. *Journal of Peace Research*, *50*(4), 523–532.

Vesco, Paola, Colaresi, Michael, Hegre, Håvard, Jansen, Remco, Lo, Adeline, Reisch, Gregor, et al. (2022). United they stand: Findings from an escalation prediction competition. *International Interactions*, *48*(4), 860–896.

Weidmann, Nils B., & Rød, Espen Geelmuyden (2015). Making uncertainty explicit. Separating reports and events in the coding of violence and contention. *Journal of Peace Research*, *52*(1), 125–128.