

COMMENTARY

The promise of machine learning in violent conflict forecasting

Max Murphy¹ , Ezra Sharpe^{1,2} and Kayla Huang³

¹School of Engineering and Applied Sciences, Harvard University Graduate School of Arts and Sciences, Cambridge, MA, USA

²School of Government, Harvard University Graduate School of Arts and Sciences, Cambridge, MA, USA

³Department of Computer Science, Harvard College, Cambridge, MA, USA

Corresponding author: Max Murphy; Email: maxmurphy1000@gmail.com

Received: 22 February 2024; **Revised:** 02 May 2024; **Accepted:** 10 May 2024

Keywords: artificial intelligence; forecasting; machine learning; peacemaking; violent conflict

Abstract

In 2022, the world experienced the deadliest year of armed conflict since the 1994 Rwandan genocide. Much of the intensity and frequency of recent conflicts has drawn more attention to failures in forecasting—that is, a failure to anticipate conflicts. Such capabilities have the potential to greatly reduce the time, motivation, and opportunities peacemakers have to intervene through mediation or peacekeeping operations. In recent years, the growth in the volume of open-source data coupled with the wide-scale advancements in machine learning suggests that it may be possible for computational methods to help the international community forecast intrastate conflict more accurately, and in doing so reduce the rise of conflict. In this commentary, we argue for the promise of conflict forecasting under several technical and policy conditions. From a technical perspective, the success of this work depends on improvements in the quality of conflict-related data and an increased focus on model interpretability. In terms of policy implementation, we suggest that this technology should be used primarily to aid policy analysis heuristically and help identify unexpected conflicts.

Policy Significance Statement

This commentary advocates for more rigorous empirical and policy examinations of the relationship between machine learning methodologies and conflict forecasting for civil wars. Rather than presuming machine learning as a panacea for conflict prediction, we advocate for a greater focus on interpretability during modeling and model heuristics in policymaking. If these conditions are met, we argue that predictive conflict models can help improve peacebuilding efforts by mitigating security risks facing peacekeeping operations, enabling more timely and judicious troop allocation, and testing different outcomes for negotiations during crisis diplomacy efforts.

1. Introduction

The world is witnessing a dramatic rise in the frequency and intensity of violent conflict. The year 2022 marked the bloodiest year for armed conflict since the Rwandan genocide in 1994, with the miasma of war, predominately inter-state conflicts, spreading from Mali to Myanmar (Institute for Economics and Peace, 2023). Importantly, this swelling of violence cannot be attributed solely to any particular conflict,

such as Russia's invasion of Ukraine—the year before the onset of this war saw over 100,000 conflict-related deaths (Institute for Economics and Peace, 2023).

Many of these internal conflicts are particularly pernicious because of a failure of forecasting. When a conflict is not adequately anticipated by the international community, relevant actors cannot take the necessary steps required—whether through greater humanitarian aid, a peacebuilding mission, or other support—to reduce the risk of a country's drift into violence. Consider the resurgence of violence in Ethiopia in 2020, which was widely and fatally unanticipated by the international community. After becoming Prime Minister in 2018, Abiy Ahmed implemented scores of liberal reforms, which included securing a peace deal with Eritrea, a country with which Ethiopia had shared a tense and hostile history (Mokaddem, 2019, 1; Soliman and Demissie, 2019). The international community was generally under the impression that conflict had subsided, so much so that Ahmed was awarded the Nobel Peace Prize, with the US celebrating his “extraordinary efforts” to “advance peace and end conflict in our world” (U.S. Embassy in Ethiopia, 2019, np). Then, in November 2020, Ahmed put an end to this simulacrum of peace by sending troops into the Tigray region, paving the way for some of the most intense violence in recent Ethiopian history. The episode has been heralded as a “cautionary tale of how the West, desperate to find a new hero in Africa, got this leader spectacularly wrong” (Walsh, 2021, np).

Of course, the recent spread of violent conflict should not be wholly reduced to the failure of forecasting. Attention to the risk of violence is only one part of the problem; sometimes, actors are aware that a conflict is likely and still fail to respond. The widespread apathy before the onset of the 1994 Rwandan genocide is illustrative of this (Dallaire, 2009). However, even in the case of Rwanda, there is evidence of international actors' failure to anticipate violence. Many organizations lacked the capacity necessary for prediction, with the Joint Evaluation of Emergency Assistance to Rwanda's report claiming that the UN had “poorly-developed structures for systematically collecting and analyzing information in a manner relevant to preventive diplomacy and conflict management” despite ample evidence (Eriksson et al., 1996, np).¹

Is this failure of forecasting—or, at the very least, the failure to improve forecasting—a corollary of the extreme complexity of conflict (Tetlock, 2005; Chadeaux, 2017a)? Conflicts are sociopolitical events defined by highly convoluted, overlapping, and nonlinear factors, so one could argue that it is fundamentally impossible to accurately predict them. The self-immolation of a Tunisian vegetable store owner triggered the largest and most rapid score of political protests the world has seen in centuries in the Arab Spring. The painting of graffiti on school walls by a group of teenage boys in Dara'a in 2011, and their subsequent torture by the Syrian authorities sparked huge protests and eventually a full-scale civil war. Indeed, “the unexpected” is not only ubiquitous in—but seems to be built into the fabric of—world politics, and perhaps this is inherently unpredictable (Taleb, 2010; Seybert and Katzenstein, 2011).

Yet, largely due to the twinned developments of the explosion in the volume of data and developments in machine learning (ML) (Buchanan, 2020), it seems increasingly feasible to forecast conflicts more accurately than ever before. A growing literature suggests that there are sturdier, more repeatable patterns undergirding conflict than was previously thought, findings that can then be used to produce promising predictions (Guo et al., 2018). Ostensibly “random” onsets or inflammations of conflict may, in fact, be foreseeable, and this offers a glimmer of hope for reducing the rise of modern conflicts.

2. Overview of the literature and developments of the technology

Conflict *forecasting*—that is, the area of research surrounding methods in anticipating conflict—is similar to, but substantively different from, conflict *modeling*, which instead aims to identify the causal relationships between particular features of a nation or population and its risk of civil violence. At its

¹ Actors with a huge financial interest in accurate predictions of conflict have also failed to consistently forecast these events. By analyzing government bond yields, Chadeaux (2017b) demonstrates that financial actors widely underestimate the risk of conflict and commonly exhibit signs of surprise in the immediate aftermath.

most extreme, the predictive part of this field has been dismissed as “unscientific” or “pointless” (Chadefaux, 2017a, 8). More commonly, though, forecasting studies are critiqued for conflating correlation with causation (Shmueli, 2010). Unpacking the theories and causal drivers of conflict is necessary, but as Chadefaux (2017a, 8) argues “both explanation *and* prediction are needed to generate and test theories”; in the last two decades, prediction has begun to assume a more prominent position as an actionable and meaningful objective. This commentary will focus primarily on the concept of civil warfare (or civil conflicts, used interchangeably here), as popularized by Fearon and Laitin (2003), who proposed a commonly accepted framework in which a civil conflict occurs when 1000 total deaths occur with at least one state force involved and is sustained at a minimum of 100 deaths each year (Blair and Sambanis, 2020).

Forecasting models are trained on data stretching from event-based data², which detail daily events (such as riots and strikes), to variables that are theoretically understood to drive conflict, including democratic indices, measures of economic inequality, and climate variables like temperature change.³ This information is then used to train a model to predict a selected outcome variable, such as the number of conflict-related fatalities per month. Deciding what, precisely, to predict—and when—has taken on many forms. Many of these decisions are limited by the nuances of the chosen dataset—for instance, Armed Conflict Location & Event Data Project (ACLED) is updated weekly, with event-based data and fatality counts (Raleigh and Kishi, 2019). This enables researchers to model fatality counts with minimal lag time.

There has been research on predicting outcomes at various phases of conflict, from the onset of violence to its termination (Kerins and Burke, 2019; Arana-Catania et al., 2022). In this sense, the field is underpinned by the notion of negative peace and tends to focus less on how ML methods can bring about positive measures (Galtung, 1969). Several studies aim to evaluate the *change* in death rates in conflict zones, too—this class of models focuses on forecasting the potential exacerbation of current at-risk areas (Vesco et al., 2022). Others predict multi-class “conflict states” (Hegre et al., 2013), binary assumptions of war or peace (Ward et al., 2010), or individual political events (Libel, 2022).

Conflict forecasting has attracted oscillating levels of skepticism and excitement over the last half-century. We draw upon Hegre et al. (2017) to briefly chronicle the field, and highlight how it has been driven by changes in government interest, data ubiquity, technological sophistication, and computation capabilities.

2.1. First wave of interest (the 1960s)

In 1963, the foundational Correlates of War was developed by Singer to accumulate quantitative knowledge about patterns of warfare (Correlates of War, 2022), and datasets produced by the project are referenced heavily (Lagazio and Russet, 2001; Gleditsch and Ward, 2013). This was bookended by work on the mathematics of war by Richardson (1960), Wright’s (1965) prediction equations, Sorokin’s (1957) “Social and Cultural Dynamics,” and Boulding’s theories on conflict from 1962 (Rummel, 1979). These academics are generally considered to be pioneers of the scientific analysis of conflict and have helped usher in a new wave of data collection and quantitative study. These efforts continued more sporadically throughout the 1970s and 1980s (Gurr and Lichbach, 1986), but received considerably less attention.

2.2. Emergence of machine learning and the establishment of Political Instability Task Force (1990s)

In the late 1980s and early 1990s, the field received a revitalization from researchers like Schrodtt who proposed several papers on the use of neural networks and other ML methods to predict interstate conflicts

² Popular open-source datasets include: ACLED; the UCDP’s GED; ViEWS; Conflict and Peace Data Bank; Mapping Militants Project; GDELT; Correlates of War project data.

³ Other general indicators that have been used in analysis include data from: the Afro barometer; Arab barometer; Ethnic Power Relations; Freedom House; Global Terrorism Database; Puarde Formal Bilateral Influence Capacity; perceived mass atrocities dataset; political terror scale; Polity score; quality of government dataset; VDEM; world bank world development indicators.

(Schrodt and Mintz, 1988). In 1989, King also published several pieces of work on extending event count variables to continuous models. From there, one of the notable pushes for further innovation came in 1994 with the creation of the Political Instability Task Force (PITF) under the CIA—a body of scholars from universities around the United States was convened to advise the federal government on states vulnerable to failure and instability (George Mason University, 2006). The group published several reports over the next decade outlining the causes of state failure and its implications for forecasting techniques (Goldstone and Gurr, 2000; Goldstone et al., 2010). In addition to drawing more attention to the field, the PITF reports also sparked an array of publications in response, some of which proposed more accurate statistical procedures (King and Zeng, 2001).

2.3. *The modern generation (2010s–2020s)*

In the early 2000s, scholars like Beck et al. (2000) and Lagazio and Russett (2001) continued to explore the use of neural networks in analyzing *interstate* conflicts. However, the greatest growth in the number of papers published began in the 2010s, beginning with widely cited papers like Ward et al. (2010), which challenged the assumption that statistically significant explaining variables were well suited for use in prediction models and helped launch a shift in focus to localized indicators and sources of data.

The landscape of forecasting *methods* has experienced significant growth over the last 20 years, with a greater focus on model explainability (Baillie et al., 2021; Attina et al., 2022). Importantly, the period brought about a diversification of approaches. The introduction of sophisticated statistical techniques and computational models greatly expanded the purview of conflict researchers, including Monte Carlo methods, as discussed by Ward and Gleditsch (2002), and agent-based modeling (Epstein, 2012). At the same time, the deployment of ML algorithms, such as random forest and gradient-boosted trees by Muchlinski et al. (2016), has enriched predictive capabilities through more complex, supervised learning methods. In terms of addressing the temporal nature of conflict data, Markov-switching processes have been employed (Brandt et al., 2014) to better capture the dynamics of time series, reflecting the nonlinearities in many forecasting domains. Similarly, the adoption of natural language processing techniques has expanded, with Mueller and Rauh (2022) applying topic modeling to navigate the challenges of class-imbalanced data and Besse et al. (2012) utilizing N-gram models for deriving forecasts from sequential event data, among others. The integration of these diverse methods into larger ensemble models has also been a notable trend as well, often resulting in significant improvements in accuracy (Ettensperger, 2021). In more recent work, though nascent, researchers have also been investigating whether or not forecasting, in general, might benefit from advances in transformers and large language models (LLMs). For instance, Google's TimesFM paper introduced a transformer-based time series model with promising zero-shot performance on times series data (Das et al., 2024), and ensembles of LLMs have also been recently shown to outperform groups of human forecasters in a variety of forecasting tournaments (Schoenegger et al., 2024).

2.4. *State-of-the-art performance and major players*

Beyond strictly technical advancements, a small group of specialized institutions has played a critical role in fostering research in conflict forecasting. Namely, the ACLED publishes forecasts and general analysis on a monthly basis with regular, regional briefs (Raleigh and Kishi 2019). Since 2021, the Peace Research Institute of Oslo has led competitions focused on predicting changes in death rates in unstable areas and other larger research initiatives, like ViEWS (Hegre et al., 2021). On the government front, the United States' State Department Bureau of Conflict and Stabilization Operations maintains comprehensive conflict forecasting and monitoring projects like the Conflict Observatory and the Instability Monitoring & Analysis Platform (United States Department of State, 2024). Over the last decade, these efforts have spurred a wealth of research, elevated the field's legitimacy, and generated several state-of-the-art models.

3. Technical bottlenecks and limitations

Despite the improvements in conflict forecasting systems, there remain some critical bottlenecks that must be addressed to improve the efficiency of these models. The first set of problems relates to data quality and the second to the interpretability of models.

3.1. Conflict data

While there are a host of issues that affect the quality of conflict data, this commentary cannot do justice to all of them and will focus on the primary concerns. First, key data are missing for many countries, which significantly hinders the ability to train performant models (Cederman and Weidmann, 2017). For example, Attina et al. (2022, 11) were unable to make forecasts for one-third of the countries in Africa because of “missing observations in the training set.” Similarly, researchers often find themselves unable to include certain types of data that could be potentially important determinants of violence (Racek et al., 2024)—almost one in four countries around the world has not had an agricultural census for 15 years, which forces them to omit agricultural data in their models (Burke et al., 2021).⁴

What makes this issue particularly difficult is that the discrepancies in the quality of the data tend to follow the chasm between wealthy and poor states. Since poorer states experience disproportionately higher levels of violent conflict than wealthier ones (Braithwaite et al., 2016) and are often caught in the “conflict trap,” this problem is especially acute (Collier and Nicholas, 2002). In half of all states in Africa, the average time taken between nationally representative livelihood surveys is 6.5 years, whereas in most wealthy countries, it happens several times each year (Burke et al., 2021). The cost of undertaking national surveys is extremely high, and some leaders may be skeptical about carrying them out to conceal a lack of economic progress (Burke et al., 2021).⁵ The process of collecting data relevant to conflict forecasting is also logistically difficult because of poor infrastructure and defective communications structures (Marivoet and De Herdt, 2014).

Reporting bias often leads to missing data or measurement errors. For example, the type of news outlet can have a large influence on what gets reported (Demarest and Langer, 2018). Herkenrath and Knoll (2011) found that in Argentina, Mexico, and Paraguay, the difference between national and international newspaper coverage was huge—with the latter reporting significantly fewer protest events than the former. While international sources may be less susceptible to partisan pressures, they exhibit a bias toward covering urban events over rural ones (Miller et al., 2022). Several studies detail bias from local sources too (Croicu and Kreutz, 2017), albeit with some disagreement on the effect of factors like press freedom (Drakos and Gofas, 2006; Urlacher, 2009).

Bias also stems from the methodological decisions made by the data collectors. Consider the discrepancy between how the ACLED and the Uppsala Conflict Data Program’s Georeferenced Event Dataset (UCDP-GED) reported civilian deaths in Mexico in 2021. The former claimed that 6739 civilians had been killed, whereas the latter identified 28. This fissure stems from a methodological difference: ACLED counts unnamed armed groups, whereas UCDP-GED does not (Raleigh et al., 2023).

In order to deal with the vast volume of event-based data—even a small segment of which would be incredibly challenging for humans to analyze—many organizations have opted to automate the process. While it has sped up the ingestion process, this approach has exacerbated a fourth complication: misclassification. GDELT has published billions of data points and releases new batches every 15 minutes, but with greater automation comes more coding errors (Demarest and Langer, 2022). For example, ICEWS, GDELT, and Phoenix, which use machine-coded data, have loose “inclusion criteria,” casting an

⁴ Beyond the omission that comes from the data collection process, data are also excluded by researchers. Often, researchers’ bias can manifest through the exclusion of certain types of data, which warps the observational architecture and, ultimately, the performance of the models (Firchow, 2018, 8).

⁵ Various techniques have been developed to address some of the challenges around missing data. Fariss et al. (2020) imputed estimates of one-sided killings in countries that were deficient in data. Radford et al. (2023) have done valuable work into measuring conflict-related fatalities across multiple reports.

extremely wide net when searching for results (Raleigh and Kishi, 2019). Consequently, many events with little to no relevance are included in the dataset. For example, in June 2019, ICEWS classified 25 events between the United States and Iran as being of comparable severity to a nuclear war (based on the “CAMEO” ontology), when in fact most of the events “capture Iran engaging in a ‘war of words’ without any conflict or threats with the US” (Raleigh et al., 2023, 6). Even more strikingly, Phoenix’s system inaccurately classified an article discussing a hippopotamus attack at Victoria Falls as a conflict between the United States and Zimbabwe (Raleigh and Kishi, 2019).

Finally, there are difficulties around duplication. At times, duplicate results can provide valuable insights as impactful events are likely to be discussed more widely. This can be a valuable signal, but when attempting to “measure changes of ‘ground-truth’ behavior,” duplication introduces real difficulties (Schrodt and Yonamine, 2012, 8). There have been efforts to address deduplication, but these have generally proven insufficient. For example, GDELT checks to see whether any other documents have the same titles, but only does so by searching for 15 minutes around the time the post was first seen and the sources it appeared in (Raleigh and Kishi, 2019). The proliferation of misinformation and disinformation, in part fueled by the same technology underpinning conflict forecasting, has heightened this problem (Bontridder and Poulet, 2021). With a higher volume of false information online, the likelihood of these machine-based data collection picking up on erroneous reports of events increases, resulting in misrepresentation of actions on the ground (Miller et al., 2022).

3.2. Interpretability

In addition to the data quality challenges, several problems can arise during the modeling phase. When neural networks are used, there is often a trade-off between accuracy and interpretability (Deng and Ning, 2021). Indeed, this is a problem that is being addressed by researchers in many fields, such as AI-driven medical work and image classification (Goebel et al., 2018; Frasca et al., 2024). There have been promising developments in making these “black boxes” more interpretable (Deng et al., 2021), and an entire field known as Explainable AI has emerged to deal with the challenge (Dwivedi et al., 2023). Nevertheless, knowing precisely what variables these systems lean on the most to make their predictions still tends to be hidden away (Brandt et al., 2022). Amarasinghe et al. (2023) distinguish between intrinsically interpretable and opaque models, the latter of which can become more explainable if post hoc methods are introduced. This poses a problem across a range of policy domains where ML is used, and the field of conflict forecasting is no different. If policymakers cannot understand *why* a country is likely to experience changes in the levels of violence—that is, changes in which variables are causing the prediction—there will be insufficient trust and minimal political will to take action (Sunstein, 2023).

However, it seems increasingly possible to balance the trade-off such that the forecasts are simultaneously accurate and understandable. Montgomery et al. (2012) make the case for using ensemble Bayesian model average (EBMA) in social sciences—an approach that pools and then averages predictions across multiple models—arguing that it improves out-of-sample forecasting. Ward and Beger (2017) utilized EBMA to generate 1- and 6-month predictions of conflict, and were able to simultaneously interpret the conflict drivers while also achieving an AUC score of 0.823, although the quality of this measure depends on the specifications of the target variable. Colaresi et al. (2016) also applied EBMA to the challenge of conflict forecasting and found that while the precision was slightly lower than that of the best-performing model, their chosen model was able to balance true positives and negatives better. Their method performed well at forecasting several spikes in conflict in January 2012, namely in South Sudan and Somalia. Still, EBMA’s effectiveness in explainability remains relatively contingent on the transparency of its constituent models, suggesting the need for further work.

More recently, in terms of interpretability, Attina et al. (2022) used dynamic elastic net to predict the number of fatalities caused by state-based conflict each month. The value of this adaptive approach is that each country was modeled individually, and the model is able to select the most relevant variables (out of 700 available). It was then possible for countries with similar conflict drivers to be grouped together, further improving the heuristic function of the research. This did come with a slight sacrifice in terms of

accuracy: DynENet performed seventh best out of the tested models in terms of Mean Squared Error. However, it remained “well above the median performance of competing models on 12 out of 13 evaluation metrics,” demonstrating the possibility of balancing these two considerations (Attina et al., 2022, 13).

While acknowledging that correlation does not prove causation, the deployment of these models has still provided some valuable heuristic insights about the contours of conflict. First, at a basic level, these models have lent credence to the “conflict trap,” which is the notion that countries or regions that have experienced conflict have a high likelihood of experiencing more conflict in the future (Collier and Nicholas, 2002). Simply put: conflict brings about conditions that are conducive to further conflict. Mueller and Rauh (2022) find that when one episode of conflict ends, the likelihood that another will start again immediately after is 30%, but after 10 years of not experiencing any conflict, the risk of a conflict breaking out is 0.5%. The corollary of this is that in the vast majority of cases, outbreaks of conflict can be accurately predicted simply by analyzing recent levels of conflict—but predicting these cases is less valuable as they are more generally anticipated by policymakers. It is the instances of conflict that exist *outside* of the conflict trap—which happen without clear and recent precedents of violence, like the violence in Tigray in 2020—that are “very unlikely and hard to forecast,” but possess the most potential policy impact because they are the most destabilizing cases (Mueller and Rauh, 2022, 2447).

The deployment of these models has also revealed the continued importance of the relationship between physical geography and conflict, perhaps pushing back against those claiming that war has entered a “post-physical” era (Gregory, 2011; Möllers, 2021). For instance, Aquino et al. (2019) mapped the geospatial network of cities in an attempt to predict conflict. Since they used a dynamical model, they were able to see which connections between geographic nodes were most responsible for changes in violence (i.e., get some sense of the main driving factors). For example, they found that in Somalia—where their model predicted new violent events with 95% accuracy—the city of Burr Gaabo went from being in the state of “war” to “peace” when the connection with Kenia, another geographic node, went from “enemy” to “ally.” Put differently: it was the hostile relationship between the two locations was largely responsible for the high risk of violence.

The final way in which these more interpretable models have served a heuristic function is through detailing the complex relationship between environmental factors and conflict. Scheffran et al. (2023) yoke together two strands of research—“tipping points” in risk/conflict and cooperation under conditions of climate stress—to investigate how climate factors could alter the risk of conflict. They deduce that poor quality governance exposes countries to “climate-induced tipping,” using the case of Lake Chad as an example, whereas having a robust civil society can act as a bulwark against climate-driven conflict (Scheffran et al., 2023, 12). This type of insight might help peacemakers who are operating in areas that are being acutely impacted by climate change, such as the Sahel. Overall, the field of conflict forecasting must continue to emphasize interpretability in addition to accuracy. It is only through *understanding* the determinants of conflict—as well as generating the predictions themselves—that stand to significantly support the work of policymakers and peacebuilders.⁶

4. Risks and opportunities in policy implementation

Conflict forecasting is becoming a tangible reality in the policy domain, especially for the United Nations Peacekeeping Operations (PKOs). The Secretary General stressed in his 2020–2022 Data Strategy report that predictive peacekeeping would bolster forecasts of armed violence, enable more accurate strategic decision-making, and encourage timely deployments of boots-on-the-ground (United Nations, 2020). Yet the state of the field remains somewhat opaque since early warning is only effective insofar that action is taken over words alone. Broadly, the impact of predictive ML for PKOs should be gauged not only

⁶We encourage the work done by ViEWS, whose annual conflict forecasting competition judged its entrants along five dimensions, one of which was “interpretability/parsimoniousness”—see Appendix A of Vesco et al. (2022).

through how interpretable the analysis is, but also through the direct effect that this analysis has on enduring peacekeeping missions (Druet, 2021).

In this regard, much of the potential for conflict forecasting in PKOs stems from the use of The Situational Awareness Geospatial Enterprise (SAGE) database, which serves as the backbone event and incident tool for most UN peacekeeping missions (Druet, 2021). The use of this data would be particularly useful for training predictive ML models; rather than training on a corpus of irrelevant data, which drains both labor and resources, models can be refined using mission-specific data with significantly greater predictive capacity. The UN's Joint Mission Analysis Centre database in Darfur, for instance, contains troves of high-quality data on troop movements, anecdotal evidence from local informants, new rebel splits, environmental factors, and even positive measures for peace such as ceasefire talks and diplomatic engagements (Galtung, 1969; Duursma and Karlsrud, 2019). To this end, the UN PKOs in Mali (MINUSMA) have succeeded in deploying predictive awareness analysis with mortar detection equipment to preemptively address threats against troops in many of Kidal's most violent hotspot regions (Druet, 2021). Here, lies one of the key opportunities of using predictive forecasting: the ability to identify threats against PKOs and strengthen camp security. The risk of insurgency against UN PKOs remains one of the key reasons for operational inefficiency, with 13 out of 24 UN civil conflict PKOs being attacked by rebel groups from 1989 to 2003 and peacekeeping troop fatalities drastically rising in MINUSMA to become the second most deadly mission in UN history (Salverda, 2013; Henke, 2016; Rietjens and de Waard, 2017; United Nations Peacekeeping, 2024). Underpinned by contextualized SAGE data, there is vast potential for conflict forecasting to reduce operational risks during PKOs and ultimately pave the way for more effective and safe peacekeeping endeavors in the future.

Monitoring highly dynamic conflicts through forecasting is also important for making prudent choices about troop deployment. PKOs are often blamed for impotence due to their lack of presence beyond military bases, but they are also faced with deployment issues when conflicts spill over into noncontiguous geographic space (Duursma and Karlsrud, 2019). Consider, for instance, that just under half of all insurgency attacks in Darfur take place over 100 kilometers from the nearest peacekeeping camp (Duursma and Karlsrud, 2019). Evidently, operational range is an issue that plagues UN PKOs, and consequently predictive geospatial data can be particularly useful while making decisions on dynamically reallocating troops (Tuvdendarjaa, 2022)—especially because the “when” question is just as important as the “where” question in conflict prediction, as discussed by the International Crisis Group in a report on PKOs in Sudan (International Crisis Group, 2005).

The third avenue of opportunity for ML in conflict forecasting is for applications in “deep conflict resolution”, a term coined by Olsher (2015) to encapsulate more holistic approaches to peacebuilding involving local knowledge, social psychology, and stakeholder values. At present, there continues to be cognitive limitations in conflict resolution that often prevent the realization of dreams of lasting stability, namely insufficient expertise, groupthink within military communities, and ethnocentric biases—particularly in contexts where PKOs do not have the due time to learn the cultural specificities of the regions within which they operate (Olsher, 2015). Duffey (2000, 165), for example, argues that a lack of cultural and linguistic understanding contributed to the failure of the UN Operation in Somalia II mission, and that “improved efforts must be made toward understanding the cultural issues at all levels of interpersonal interaction and process implementation.” In such cases, the continued development of predictive ML tools like Olsher's (2015, 282) *cogSol* can be used to simulate and forecast different stakeholder views of the world based on field experts' cultural models and real-time conflict data, allowing peacekeeping personnel to “find negotiation win-wins..., avoid offense, provide peacekeeping decision tools, and protect emergency responders' health.” Importantly, this should not be used to further an at-a-distance foreign policy, where PKOs might be more inclined to manage from afar and therefore lessen “the ability to interact, understand and empathize with local populations” (Duursma and Karlsrud, 2019, 13). By working in line with gradual reduction in tensions theory, predictive negotiation tools can be used by civil society NGOs, UN civil affairs officers, and international diplomats to work alongside local communities in culturally and politically complex environments to maximize outcomes and foster a lasting stability and peace (Duursma and Karlsrud, 2019).

Some key risks and obstacles remain a barrier for the successful implementation of ML tools in peacebuilding. The growing focus on open-sourcing data is particularly instructive in the context of UN PKOs since a secondary analysis of mission success could also involve sharing data internally between missions to learn and analyze best practices and failures (Druet, 2021). Coupled with developments in reinforcement learning algorithms, a strong culture of data sharing could help deliver critical insights for optimizing missions and resources. However, many mission leaders continue to create internal friction when asked to provide data to the UN headquarters, protesting that performance metrics lack context or could leak sensitive information (Druet, 2021). This “paradox of information ownership and sharing” within UN PKOs continues to be impedimentary for the deployment of ML in effectively forecasting and responding to conflict (Druet, 2021, 17).

The researcher must also conduct dimension reduction at some point in the model life cycle, introducing the risk that personal biases become baked into model taxonomies. Most researchers often try to follow the principle of Ockham’s razor—that the best model is the one with the fewest assumptions made—and attempt to achieve this by reducing the number of features, and thus assumptions on causality (Wainwright and Mulligan, 2013; Piasini et al., 2023). In doing so, many issues can arise in the context of conflict forecasting, as seen in the case when analysts working on the UN PKO in the Congo (MONUSCO) attempted to amalgamate a large number of Mayi-Mayi militia groups under a single title, which caused issues further downstream in attributing perpetrators of attacks and wrongfully accusing communities of insurgency (Druet, 2021; United Nations, 2024).

Perhaps the most discussed risk in the literature of conflict forecasting is that of information security and adversarial actors. Despite the size of human resources available, the UN and other peacebuilding bodies are not well-resourced enough to mitigate against intrusions from highly sophisticated cyber adversaries (Druet, 2021). As a supranational organization comprised of many competing national interests, there is always a latent risk that training alone cannot obviate different national allegiances, and resultantly “states do not wish to share secrets with all the countries in the world and refuse to allow other states to send troops to spy on their own governance” (Martin-Brûlé, 2021, 494). This problem becomes especially acute in situations where misinformation and disinformation campaigns arise to splinter factions in UN PKOs, such as in the Central African Republic, which risks confidential data on informants and predictive analysis being leaked to direct adversaries who intend to further promote conflict (Druet, 2021). The logical extension in addressing this risk is to truly underscore the importance of data privacy concerns, even more so given the fact that a significant amount of SAGE data is collected from local informants whose personal security is always jeopardized by insurgency groups (Druet, 2021).

The final risk in policy implementation is expectation management. Many of the challenges associated with using forecasting models in policy are rooted in the unique expectations placed upon the field of geopolitics, as analysts are often counted on to give a prophetic view into the future and to generate a wholly deterministic “upstream” understanding of events that have not yet happened (Gentry and Gordon, 2019). Politically speaking, this creates a danger that conflict forecasting can *only* be appreciated for its validation—that is, the predictive accuracy of models—rather than verification, which is “the process by which the model is checked to make sure that it is solving the equations correctly” (Clifford and Valentine, 2003, 286; Baillie et al., 2021). Gentry and Gordon (2019) discuss this risk with the “batting-average” metaphor used in intelligence communities, where analysts are measured on their overall frequency of “hits” rather than their analytic rigor—which only helps to foster negligence by focusing solely on accuracy instead of notions of error and interpretability (Jervis, 2010). To address this, Mueller and Rauh’s (2022) ML cost-based intervention framework uses a conflict weighting system to allocate degrees of emphasis on some potential false positive scenarios over others based on their possible scale of harm, allowing decision-makers to take fuller stock of developing cases that could rapidly escalate. Therefore, a sufficiently cautious application of forecasting can avoid the dangers of wrongfully viewing predictive models as a panacea for intelligence forecasting (Musumba et al., 2021). If used in concert with human agency while attempting to address the risks of implementation, forecasting models can be decisive tools for predicting, preventing, and responding to violent conflict.

5. Discussion

Even if comprehensive and granular data encompassing a range of the drivers of civil war existed, it would be impossible to perfectly predict conflict since it is a complex sociopolitical phenomenon laden with interlocking and nonlinear variables. However, in this commentary, we argue that this technology can help forecast violent conflict with a meaningful degree of accuracy, which then can—and should—be used to inform foreign policy and peacebuilding decisions. While we advocate for greater use of AI systems in conflict forecasting, we also encourage caution by emphasizing some critical considerations spanning both the technical process of building these systems and the policy implementation stage. The data used to train these models—including both data that capture the determinants of conflict, such as economic indicators and climate variables, and event-based data—remain the key barrier to progress. Clausewitz’s claim, made 200 years ago, that “casualty reports on either side are never accurate, seldom truthful, and in most cases deliberately falsified” continues to hold true (Clausewitz, 1976 [1832], 234). Moreover, when building these models, there should be a strong emphasis on interpretability—even if it comes at a slight cost of accuracy. Understanding *why* a model presents specific predictions facilitates one of the key benefits of the deployment of this technology: the heuristic function. The use of these ML models has already highlighted some key characteristics of conflict in the 21st century, such as the continued presence of the conflict trap, the importance of physical geography, and the complex relationship between environmental factors and conflict. This improved understanding of conflict can then inform the protection of PKO personnel, troop deployment, and deep conflict resolution. What forecasting efforts can teach policymakers and peacebuilders about the character of conflict in the 21st century is of comparable value to the predictions themselves.

Provenance. This article was submitted for consideration for the 2024 Data for Policy Conference to be published in *Data & Policy* on the strength of the Conference review process.

Acknowledgments. The authors are grateful for the valuable feedback provided by Dr. Michael Kenwick (Rutgers University) and Dr Juan Luis Manfredi Sánchez (Georgetown University).

Data availability statement. No original data was used in the production of this article.

Author contribution. Conceptualization: M.M; E.S; K.H. Methodology: M.M; E.S; K.H. Writing original draft: M.M; E.S; K.H. All authors approved the final submitted draft.

Competing interest. The authors declare no competing interests exist.

References

- Amarasinghe K, Rodolfa KT, Lamba H and Ghani R (2023) Explainable machine learning for public policy: Use cases, gaps, and research directions. *Data & Policy* 5, e5.
- Aquino G, Guo W and Wilson A (2019) Nonlinear dynamic models of conflict via multiplexed interaction networks. Preprint. arXiv:1909.12457.
- Arana-Catania M, van Lier F-A and Procter R (2022) Supporting peace negotiations in the Yemen war through machine learning. *Data & Policy* 4, e28. <https://doi.org/10.1017/dap.2022.19>
- Attina F, Carammia M and Iacus SM (2022) Forecasting change in conflict fatalities with dynamic elastic net. *International Interactions* 48(4), 649–677.
- Baillie EJ, Howe PD, Perfors A, Miller T, Kashima Y and Beger A (2021) Explainable models for forecasting the emergence of political instability. *PLoS One* 16, e0254350. <https://doi.org/10.1371/journal.pone.0254350>
- Beck NN, Diego S, King G and Zeng L (2000) Improving quantitative studies of international conflict: A conjecture. *American Political Science Review* 94, 21–35. <https://doi.org/10.1017/S0003055400220078>
- Besse C, Bakhtiari A and Lamontagne L (2012) Forecasting conflicts using N-grams models. The Florida AI Research Society. Available at <https://www.semanticscholar.org/paper/Forecasting-Conflicts-Using-N-Grams-Models-Besse-Bakhtiari/018fad5c571cf0dc5861afa9620c373a99e0567e>.
- Blair R and Sambanis N (2020) Forecasting civil wars: Theory and structure in an age of “big data” and machine learning. *Journal of Conflict Resolution* 64(10), 1885–1915. <https://doi.org/10.1177/0022002720918923>
- Bontridder N and Poulet Y (2021) The role of artificial intelligence in disinformation. *Data & Policy*. 3, e32. <https://doi.org/10.1017/dap.2021.20>
- Braithwaite A, Dasandi N and Hudson D (2016) Does poverty cause conflict? Isolating the causal origins of the conflict trap. *Conflict Management and Peace Science* 33(1), 45–66. <https://doi.org/10.1177/0738894214559673>

- Brandt PT, D'Orazio V, Khan L, Li YF, Osorio J and Sianan M** (2022) Conflict forecasting with event data and spatio-temporal graph convolutional networks. *International Interactions* 48, 800–822.
- Brandt PT, Freeman JR and Schrodt PA** (2014) Evaluating forecasts of political conflict dynamics. *International Journal of Forecasting* 30, 944–962. <https://doi.org/10.1016/J.IJFORECAST.2014.03.014>
- Buchanan B** (2020) *The AI Triad and What it Means for National Security Strategy*. Center for Security and Emerging Technology (CSET).
- Burke M, Driscoll A, Lobell DB and Ermon S** (2021) Using satellite imagery to understand and promote sustainable development. *Science* 371, 1219.
- Cederman LE and Weidmann NB** (2017) Predicting armed conflict: Time to adjust our expectations? *Science* 355, 474–476. <https://doi.org/10.1126/science.aal4483>
- Chadefaux T** (2017a) Conflict forecasting and its limits. *Data Science* 1(1–2), 7–17. <https://doi.org/10.3233/DS-170002>
- Chadefaux T** (2017b) Market anticipations of conflict onsets. *Journal of Peace Research* 54(2), 313–327. <https://doi.org/10.1177/0022343316687615>
- Clausewitz CV** (1976) The engagement in general continued. In Howard M and Paret P (eds.), *On War*. Princeton, NJ: Princeton University Press.
- Clifford N and Valentine G** (2003) *Key Methods in Geography*. London, UK: Sage Publications Inc.
- Colaresi M, Hegre H and Nordkvelle J** (2016) Early ViEWS: A prototype for a political violence early-warning system. In *American Political Science Association Annual Meeting 2016, Philadelphia*.
- Collier P and Nicholas S** (2002) Understanding civil war: A new agenda. *Journal of Conflict Resolution* 46, 3–12.
- Correlates of War** (2022) About. Correlates of War. Available at <https://correlatesofwar.org/history/> (accessed 29 January 2024).
- Croicu M and Kreutz J** (2017) Communication technology and reports on political violence: Cross-national evidence using African events data. *Political Research Quarterly* 70, 19–31.
- Dallaire R** (2009) *Shake Hands with the Devil: The Failure of Humanity in Rwanda*. Canada: Vintage Canada.
- Das A, Kong W, Sen R and Zhou Y** (2024) A decoder-only foundation model for time-series forecasting. Preprint. arXiv:2310.10688.
- Demarest L and Langer A** (2018) The study of violence and social unrest in Africa: A comparative analysis of three conflict event datasets. *African Affairs* 117(467), 310–325.
- Demarest L and Langer A** (2022) How events enter (or not) data sets: The pitfalls and guidelines of using newspapers in the study of conflict. *Sociological Methods & Research* 51(2), 632–666. <https://doi.org/10.1177/0049124119882453>
- Deng S and Ning Y** (2021) A survey on societal event forecasting with deep learning. Preprint. arXiv:2112.06345.
- Deng S, Rangwala H and Ning Y** (2021) Understanding event predictions via contextualized multilevel feature learning. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. New York, NY: ACM.
- Drakos K and Gofas A** (2006) The devil you know but are afraid to face: Underreporting bias and its distorting effects on the study of terrorism. *Journal of Conflict Resolution* 50(5), 714–735. <https://doi.org/10.1177/0022002706291051>
- Druet D** (2021) *Enhancing the Use of Digital Technology for Integrated Situational Awareness and Peacekeeping-Intelligence*. DPO Peacekeeping Technology Strategy.
- Duffey T** (2000) Cultural issues in contemporary peacekeeping. *International Peacekeeping* 7(1), 142–168.
- Duursma A and Karlsrud J** (2019) Predictive peacekeeping: Strengthening predictive analysis in UN peace operations. *Stability: International Journal of Security & Development* 8(1), 1–19.
- Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P and Ranjan R** (2023) Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys* 55(9), 1–33.
- Epstein JM** (2012) Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences of the United States of America* 99(Suppl 3), 7243–7250. <https://doi.org/10.1515/9781400842872.247>
- Eriksson J, Adelman H, Borton J, Christensen H, Kumar K, Suhrke A, Tardif-Douglin D, Villumstad S, and Wohlgenuth L** (1996) The international response to conflict and genocide: Lessons from the Rwanda experience: Synthesis report. Joint Evaluation of Emergency Assistance to Rwanda.
- Ettensperger F** (2021) Forecasting conflict using a diverse machine-learning ensemble: Ensemble averaging with multiple tree-based algorithms and variance promoting data configurations. *International Interactions* 48, 555–578. <https://doi.org/10.1080/03050629.2022.1993209>
- Fariss CJ, Kenwick MR and Reuning K** (2020) Estimating one-sided-killings from a robust measurement model of human rights. *Journal of Peace Research* 57(6), 801–814. <https://doi.org/10.1177/0022343320965670>
- Fearon JD and Laitin DD** (2003) Ethnicity, insurgency, and civil war. *The American Political Science Review*, 97(1), 75–90.
- Firchow P** (2018) *Reclaiming Everyday Peace: Local Voices in Measurement and Evaluation after War*. Cambridge, UK: Cambridge University Press.
- Frasca M, La Torre D and Pravettoni G** (2024) Explainable and interpretable artificial intelligence in medicine: A systematic bibliometric review. *Discover Artificial Intelligence* 4, 15. <https://doi.org/10.1007/s44163-024-00114-7>
- Galtung J** (1969) Violence, peace, and peace research. *Journal of Peace Research* 6(3), 167–191. <https://doi.org/10.1177/002234336900600301>
- Gentry JA and Gordon JS** (2019) *Strategic Warning Intelligence: History, Challenges, and Prospects*. Washington, D.C.: Georgetown University Press.
- George Mason University** (2006) Internal wars and failures of governance, 1955–2005. Available at <https://web.archive.org/web/20061208000556http://globalpolicy.gmu.edu/pitf/> (accessed 2 February 2024).

- Gleditsch KS and Ward MD** (2013) Forecasting is difficult, especially about the future. *Journal of Peace Research* 50, 17–31. <https://doi.org/10.1177/0022343312449033>
- Goebel R, Chandler A, Holzinger K, Lecue F, Akata Z, Stumpf S and Holzinger A** (2018) Explainable AI: The new 42?. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction* (pp. 295–303). Cham: Springer.
- Goldstone J and Gurr T** (2000) Executive summary: State failure task force report, phase III findings. Instability Task Force, University of Maryland. Available at <https://gsdrc.org/document-library/state-failure-task-force-report-phase-iii-findings/>.
- Goldstone JA, Bates RH, Epstein DL, Gurr TR, Lustik M, Marshall MG, Ulfelder J and Woodward M** (2010) A global model for forecasting political instability. *Political Economy: Structure & Scope of Government eJournal* 54, 190–208. <https://doi.org/10.1111/J.1540-5907.2009.00426.X>
- Gregory D** (2011) The everywhere war. *Geographical Journal* 177(3), 238–250.
- Guo W, Gleditsch K and Wilson A** (2018) Retool AI to forecast and limit wars. *Nature* 562, 331–333.
- Gurr TR and Lichbach MI** (1986) Forecasting internal conflict. *Comparative Political Studies* 19, 3–38. <https://doi.org/10.1177/0010414086019001001>
- Hegre H, Karlsen J, Nygård HM, Strand H and Urdal H** (2013) Predicting armed conflict, 2010–2050. *International Studies Quarterly* 57, 250–270. <https://doi.org/10.1111/ISQU.12007>
- Hegre H, Metternich NW, Nygård H M and Wucherpfennig J** (2017) Introduction: Forecasting in peace research. *Journal of Peace Research* 54(2), 113–124.
- Hegre H, Nygård H M and Landsverk P** (2021) Can we predict armed conflict? How the first 9 years of published forecasts stand up to reality. *International Studies Quarterly* 65(3), 660–668. <https://doi.org/10.1093/isq/sqaa094>
- Henke ME** (2016) Has UN peacekeeping become more deadly? Analyzing trends in UN fatalities. In *Providing for Peacekeeping No. 14*. New York: International Peace Institute.
- Herkenrath M and Knoll A** (2011) Protest events in international press coverage: An empirical critique of cross-national conflict databases. *International Journal of Comparative Sociology* 52(3), 163–180.
- Institute for Economics & Peace** (2023) Global Peace Index: Measuring Peace in a Complex World, Sydney, June 2023. Available at <http://visionofhumanity.org/resources>.
- International Crisis Group** (2005) The AU's Mission in Darfur: Bridging the Gaps. Africa Briefing No. 28.
- Jervis R** (2010) *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War*. Ithaca: Cornell University Press.
- Kerins P and Burke S** (2019) Forecast Cloudy: A Case Study in Predicting Conflict Risk, New America. Available at <https://www.newamerica.org/resource-security/reports/conflict-prediction-case-study/> (accessed 3 February 2024).
- King G and Zeng L** (2001) Improving forecasts of state failure. *World Politics* 53, 623–658. <https://doi.org/10.1353/wp.2001.0018>
- Lagazio M and Russett BM** (2001) A neural network analysis of militarized disputes, 1885–1992. Temporal Stability and Causal Complexity. Available at <https://www.semanticscholar.org/paper/A-NEURAL-NETWORK-ANALYSIS-OF-MILITARIZED-DISPUTES%2C-Lagazio-Russett/2532ded44a8be70833e908ed69ec613d50bde715>.
- Libel T** (2022) Lesson (un)replicated: Predicting levels of political violence in afghan administrative units per month using ARFIMA and ICEWS data. *Data & Policy* 4, e32. <https://doi.org/10.1017/dap.2022.26>
- Marivoet W and De Herdt T** (2014) Reliable, challenging or misleading? A qualitative account of the most recent national surveys and country statistics in the DRC. *Canadian Journal of Development Studies/Revue canadienne d'études du développement* 35(1), 97–119.
- Martin-Brulé S-M** (2021) Competing for trust: Challenges in United Nations peacekeeping-intelligence. *International Journal of Intelligence and Counterintelligence* 34(3), 494–524.
- Miller E, Kishi R, Raleigh C and Dowd C** (2022) An agenda for addressing bias in conflict data. *Scientific Data* 9, 593. <https://doi.org/10.1038/s41597-022-01705-8>
- Mokaddem S** (2019) Abiy Ahmed's 'Medemer' reforms: Can it ensure sustainable growth for Ethiopia and what are the challenges facing the new government? Policy Center for The New South, PB-19/08.
- Möllers N** (2021) Making digital territory: Cybersecurity, techno-nationalism, and the moral boundaries of the state. *Science, Technology, & Human Values* 46(1), 112–138.
- Montgomery JM, Hollenbach F M and Ward M D** (2012) Improving predictions using ensemble Bayesian model averaging. *Political Analysis* 20(3), 271–291. <http://www.jstor.org/stable/23260318>.
- Muchlinski DA, Siroky DS, He J and Kocher MA** (2016) Comparing random Forest with logistic regression for predicting class-imbalanced civil war onset data. *Political Analysis* 24, 87–103. <https://doi.org/10.1093/pan/mpv024>
- Mueller M and Rauh C** (2022) The hard problem of prediction for conflict prevention. *Journal of the European Economic Association* 20(6), 2440–2467. <https://doi.org/10.1093/jeea/jvac025>.
- Musumba M, Fatema N and Kibriya S** (2021) Prevention is better than cure: Machine learning approach to conflict prediction in sub-Saharan Africa. *Sustainability* 13(13), 7366. <https://doi.org/10.3390/su13137366>
- Olsher DJ** (2015) New artificial intelligence tools for deep conflict resolution and humanitarian response. *Procedia Engineering* 107, 282–292.
- Piasini E, Liu S, Chaudhari P, Balasubramanian V and Gold JI** (2023) How Occam's razor guides human decision-making. bioRxiv [Preprint].
- Racek D, Thurner PW, Davidson BI, Zhu XZ and Kauermann G** (2024) Conflict forecasting using remote sensing data: An application to the Syrian civil war. *International Journal of Forecasting* 40(1), 373–391.
- Radford BJ, Dai Y, Stoehr N, Schein A, Fernandez M and Sajid H** (2023) Estimating conflict losses and reporting biases. *Proceedings of the National Academy of Sciences* 120(34), e2307372120.

- Raleigh C and Kishi R** (2019) *Comparing Conflict Data: Similarities and Differences across Conflict Datasets*. WI, USA: ACLED.
- Raleigh C, Kishi R and Linke A** (2023) Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices. *Humanit Soc Sci Commun* 10, 74. <https://doi.org/10.1057/s41599-023-01559-4>
- Richardson LF** (1960) Statistics of deadly quarrels, 1809–1949. Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/ICPSR05407.v1>
- Rietjens S and de Waard E** (2017) UN peacekeeping intelligence: The ASIFU experiment. *International Journal of Intelligence and CounterIntelligence* 30(3), 532–556.
- Rummel RJ** (1979) Understanding conflict and war: Vol. 4: War, power, peace. University of Hawaii. Available at https://www.hawaii.edu/powerkills/WPP.CHAP13.HTM#* (accessed 2 February 2024).
- Salverda N** (2013) Blue helmets as targets?: A quantitative analysis of rebel violence against peacekeepers, 1989–2003. *Journal of Peace Research* 50(6), 707–720.
- Scheffran J, Guo W, Krampe F and Okpara U** (2023) Tipping cascades between conflict and cooperation in climate change. *EGUsphere* 2023, 1–27.
- Schoenegger P, Tuminauskaite I, Park PS and Tetlock PE** (2024) Wisdom of the silicon crowd: LLM ensemble prediction capabilities rival human crowd accuracy. Preprint. [arXiv:2402.19379](https://arxiv.org/abs/2402.19379).
- Schrodt PA and Mintz A** (1988) The conditional probability analysis of international events data. *American Journal of Political Science* 32(1), 217–230. <https://doi.org/10.2307/2111318>
- Schrodt PA and Yonamine J** (2012) Automated coding of very large scale political event data. In *New Directions in Text as Data Workshop*, Harvard.
- Seybert LA and Katzenstein PJ** (2011) Protean power and control power: Conceptual analysis. In Katzenstein PJ and Seybert LA (eds.), *Protean Power: Exploring the Uncertain and Unexpected in World Politics*. Cambridge Studies in International Relations. Cambridge: Cambridge University Press, pp. 3–26.
- Shmueli G** (2010) To explain or to predict? *Statistical Science* 25(3), 289–310. <https://doi.org/10.1214/10-STS330>
- Soliman A and Demissie AA** (2019) Can Abiy Ahmed Continue to Remodel Ethiopia? Chatham House. Available at <https://www.chathamhouse.org/2019/04/can-abiy-ahmed-continue-remodel-ethiopia>.
- Sorokin P** (1957) *Social and Cultural Dynamics*. Boston, MA, USA: Porter Sargent.
- Sunstein CR** (2023) The use of algorithms in society. *Review of Austrian Economics*, 1–22.
- Taleb NN** (2010) *The Black Swan: The Impact of the Highly Improbable*. New York, NY, USA: Random House Trade Paperbacks.
- Tetlock PE** (2005) *Expert Political Judgment*. Princeton, NJ: Princeton University Press.
- Tuvdendarjaa M** (2022) Challenges of the United Nations Peacekeeping Operations. DPI APCSS Security Nexus, vol. 23. Available at https://dkiapcss.edu/nexus_articles/challenges-of-the-united-nations-peacekeeping-operations/.
- U.S. Embassy in Ethiopia** (2019) Congratulations on the Awarding of the Nobel Peace Prize to Prime Minister Abiy. Available at <https://et.usembassy.gov/congratulations-on-the-awarding-of-the-nobel-peace-prize-to-prime-minister-abiy/> (accessed 1 February 2024).
- United Nations** (2020) Data Strategy of the Secretary-General for Action by Everyone, Everywhere with Insight, Impact and Integrity 2020–22.
- United Nations** (2024) Escalating Violence in Democratic Republic of Congo Exacerbating Humanitarian Crisis, Special Representative Warns Security Council, Urging Durable Political Solution | Meetings Coverage and Press Releases. Available at <https://press.un.org/en/2024/sc15596.doc.htm>.
- United Nations Peacekeeping** (2024) Fatalities. Available at <https://peacekeeping.un.org/en/fatalities>.
- United States Department of State** (2024) Instability Monitoring and Analysis Platform (IMAP). State Department. Available at www.state.gov/about-us-bureau-of-conflict-and-stabilization-operations/instability-monitoring-and-analysis-platform/ (accessed 2 April 2024).
- Urlacher BR** (2009) Wolfowitz conjecture: A research note on civil war and news coverage. *International Studies Perspectives* 10(2), May, Pages 186–197. <https://doi.org/10.1111/j.1528-3585.2009.00369.x>
- Vesco P, Hegre H, Colaresi M, Jansen RB, Lo A, Reisch G and Weidmann NB** (2022) United they stand: Findings from an escalation prediction competition. *International Interactions* 48(4), 860–896.
- Wainwright J and Mulligan M** (2013) *Environmental Modelling*. Hoboken, NJ, USA: John Wiley & Sons.
- Walsh D** (2021) The Nobel Peace Prize That Paved the Way for War. *New York Times*, 15 December 2021.
- Ward MD and Beger A** (2017) Lessons from near real-time forecasting of irregular leadership changes. *Journal of Peace Research* 54(2), 141–156. <https://doi.org/10.1177/0022343316680858>
- Ward MD and Gleditsch KS** (2002) Location, location, location: An MCMC approach to modeling the spatial context of war and peace. *Political Analysis* 10, 244–260. <https://doi.org/10.1093/pan/10.3.244>
- Ward MD, Greenhill B and Bakke KM** (2010) The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research* 47, 363–375. <https://doi.org/10.1177/0022343309356491>
- Wright Q** (1965) *A Study of War*. Chicago: University Of Chicago Press.

Cite this article: Murphy M, Sharpe E and Huang K (2024). The promise of machine learning in violent conflict forecasting. *Data & Policy*, 6: e35. doi:10.1017/dap.2024.27